Power Systems

Naser Mahdavi Tabatabaei
Nicu Bizon *Editors*

Numerical Methods for Energy Applications

Springer

Power Systems

Electrical power has been the technological foundation of industrial societies for many years. Although the systems designed to provide and apply electrical energy have reached a high degree of maturity, unforeseen problems are constantly encountered, necessitating the design of more efficient and reliable systems based on novel technologies. The book series Power Systems is aimed at providing detailed, accurate and sound technical information about these new developments in electrical power engineering. It includes topics on power generation, storage and transmission as well as electrical machines. The monographs and advanced textbooks in this series address researchers, lecturers, industrial engineers and senior students in electrical engineering.

Power Systems is indexed in Scopus

More information about this series at http://www.springer.com/series/4622

Naser Mahdavi Tabatabaei · Nicu Bizon
Editors

Numerical Methods for Energy Applications

Springer

Editors
Naser Mahdavi Tabatabaei
Department of Electrical Engineering
Seraj Higher Education Institute
Tabriz, Iran

Nicu Bizon
Faculty of Electronics
Communications and Computers
University of Pitesti
Pitesti, Romania

ISSN 1612-1287 ISSN 1860-4676 (electronic)
Power Systems
ISBN 978-3-030-62190-2 ISBN 978-3-030-62191-9 (eBook)
https://doi.org/10.1007/978-3-030-62191-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to
all our teachers and colleagues
who enabled us to write this book,
and our family and friends
for supporting us all along

Foreword

The demand for energy is steadily growing, in spite of the efforts made to reduce industrial and domestic consumption. Expanding the use of Variable Renewable Energy (VRE) technologies, such as wind and photovoltaic, in power systems is vital for reducing pollutant emissions and limiting the impact on climate change while continuing to meet this demand. With their costs constantly reducing, these sustainable energy technologies, which are strongly supported by governmental policies, have been more and more widely deployed in recent years. However, the inherent variability of wind and solar PV energy resources represent a major challenge for power systems operators and regulators. The use of distributed generation in interconnected microgrids is the best way to meet the variable energy demand. Optimal solutions should be found for the allocation of such distributed energy sources and associated storage devices, as well as for the interconnection of the microgrids.

The aim of this book is to present the appropriate numerical methods that could be used to address this kind of novel challenges that power engineers and scientists have to face, and that cannot be solved by known exact formulas. Thus, the book provides the reader with both basic and advanced knowledge about the techniques of modeling, simulation, design, control, and optimization of the sustainable energy systems of the future.

Numerical methods are able to promptly provide approximate solutions to the mathematical problems associated with the design, control, and optimization of an energy system. The most widely used numerical methods in the literature are introduced in the first part of the book. By reading the nine chapters of this part, Ph.D. students, as well as the engineers and scientist less familiar with the field, will quickly gain the necessary knowledge on the algorithms and the numerical approximation strategies generally used in engineering, and more specifically in the study of electric power systems.

The eight chapters of the second part of the book address the specific problems of advanced energy systems. The advantages and disadvantages of different numerical methods are comparatively highlighted so that to enable the specialists in the field to select the most appropriate one for a specific energy application.

The last part of the book includes 18 chapters that exemplify the application of the numerical methods discussed in the first part for the study of the energy systems presented in the second part.

The 35 chapters of the books are structured in three parts for pedagogical reasons. The readers who master the numerical methods and the principles of electric systems can skip the first and second parts, while the others will appreciate the possibility of gradually acquire the expertise needed for solving their specific problems.

All those who contributed to this book are recognized experts in the various fields of power engineering. They ensured the high-quality content of each chapter they wrote. Not only master and Ph.D. students, young researchers, and engineers but also senior practitioners will find this book useful for their studies or professional activities related to the modeling, control, and optimization of modern energy systems.

Angouleme, France
August 2020

Prof. Dr.-Ing. Habil. Sci. mDHC Lucian Dascalescu
IEEE Fellow

Preface

Energy demand has increased exponentially in recent years due to industrialization and population growth and is one of the main indicators showing the level of economic and industrial growth for a country. Consequently, more attention must be paid to how energy is generated and transmitted to the consumers. Because fossil fuel reserves are limited and energy production based on these resources is polluting, sustainable energy development based on renewable energy is needed, which in the next decade will help reduce pollutant emissions and the impact on climate change.

The purpose of this book is indirectly related to the issue mentioned above because the book's chapters address the energy applications that engineers, specialists, and scientists face every day using numerical methods to solve these mathematical problems in energy when they cannot be solved by known exact methods. So the book will address numerical calculation methods in general but will exemplify how they apply especially to energy systems. Thus, reading the book, the reader will gain knowledge about the techniques of modeling, simulation, design, control, and optimization of energy systems.

Any of the techniques mentioned above (modeling, simulation, design, control, and optimization) can be approached by numerical methods that provide a quick approximation solution for the defined mathematical problem in the case of an energy application. Although the result obtained is not exact but close to the exact solution, the algorithms specific to numerical methods are continuously improved to obtain an imposed error even in real-time applications. In addition, computing power has increased and allows the implementation of these complex algorithms based on a complete and clear set of rules and procedures that estimate computable errors. So, current energy systems (classified into mini-grids, micro-grids, nano-grids, and pico-grids according to power level, complexity, and connectivity) include high-speed digital computers that allow real-time communication between them, thus making it possible to obtain optimal or suboptimal solutions for many complex problems that could not be solved efficiently in centralized energy systems.

The numerical examples included in this book's chapters show that we need these numerical methods to solve some problems that are not analytically solvable. The obtained results prove that the algorithms proposed in the literature work in a remarkable way so that the unsolvable problems can quickly become solvable.

The comparative analysis of the most well-known numerical algorithms offers the readers of this book (whether Ph.D. students or specialists) a strong background on high-performance numerical algorithms and an up-to-date perspective on numerical approximation strategies generally used in all fields of engineering. The numerical examples presented in this book range from simple ones (which introduce numerical methods to be understood) to advanced ones that are addressed especially to specialists in the field of energy systems. Thus, potential solutions are offered for many practical problems that require more or less complex numerical methods.

Therefore, going through and understanding the techniques presented in this book will help doctoral students to prepare thoroughly, giving them the necessary knowledge to identify when it is necessary to use a certain numerical method to solve an energy application, taking into account the specifics of this problem (which can be control, optimization, etc.) and its complexity.

Certainly, the specialists will find in the 35 chapters of the book the solutions they need to solve many of the problems encountered in their daily work, which otherwise they would have had to look for in many other books and journals.

For pedagogical reasons, the approach is gradual and it is recommended for doctoral students to read the chapters in order, starting with those in part I, which lay the foundations of Advanced Numerical Methods.

The numerical examples in Part II address specific Advanced Energy Systems problems using numerical methods from simple to complex, in order to comparatively highlight the advantages of the latter, or methods based on different numerical calculation algorithms (in principle equally competitive), in order to highlight comparatively why one method or another does not work well for a specific energy application, which is very important from the point of view of a specialist.

The numerical examples in Part III exemplify how the application of the methods discussed in Part I for Energy Systems to few selected energy applications considered representative for understanding. So, this book comprises 35 chapters structured in three parts as mentioned above.

The first part called *Advanced Numerical Methods* introduces in nine chapters the main numerical methods used in engineering and physical science in general, highlighting the advantages and disadvantages of each class of methods, and the recommended fields of application. A short introduction for readers on the content of the nine chapters will be presented below.

Chapter "Advanced Numerical Methods for Equations, Systems Equations and Optimization" presents an overview of advanced numerical methods sustained with many examples in Matlab to help understand these methods and the errors offered by each of them in a finite number of epochs.

The analysis of partial differential equations in time-dependent problems using Finite Difference Methods (FDM) is performed in Chapter "Analysis of Partial Differential Equations in Time Dependent Problems using Finite Difference Methods and the Applications on Electrical Engineering," where it is highlighted that these old and simple numerical methods are still used successfully in solving differential equations encountered in boundary value problems, zero-stability and convergence of initial value problems, absolute stability of boundary value equations, stiff ordinary

differential equations, diffusion equations, and parabolic problems. So, this chapter presents the theoretical and practical approaches of the finite difference method based on electrical engineering energy applications.

A deep theoretical analysis of the Finite Element Method (FEM) is approached in Chapter "Theoretical Approaches of Finite Elements Method (FEM)," highlighting that the procedure to use the FEM is different from that for FDMs. The differences between the two FEM methods called Galerkin and Ritz methods, which are usually used in numerical electrostatics and magnetostatics analysis, AC analysis, transient analysis, and 2D or 3D geometry analysis, are highlighted with clear examples included in this chapter.

Chapter "Advanced Numerical Methods Based on Artificial Intelligence" approaches the numerical methods based on artificial intelligence and tests a Genetic Algorithm–based method in determining the optimal horizontal model of equivalent soil and a Neural Network–based method for the evaluation of the impedance matrix regarding stratified soil with three vertical layers, which is necessary, for example, to evaluate the inductive coupling between overhead high-voltage transmission lines and metallic gas transmission pipeline.

Different iterative methods to solve a system of nonlinear equations are described and discussed methods in Chapter "Numerical Methods for Solving Nonlinear Equations," including the class of conjugate gradient methods, multi-step methods, and Newton-like methods, which are usually used to provide an accurate prediction of a natural phenomenon or a good and real-time optimization of systems, at relatively low computational costs and effort involved.

It is worth mentioning that the numerical methods such as FEM and FDM, or other numerical methods are based on the predefined topological map, generally called "mesh," such as Finite Volume Method (FVM) and Boundary Value Method (BVM), are very eminent for solving the physical problems in the engineering and science, but the problems with mesh-based methods are as follows: (i) they require the qualitative mesh, which is somewhat tedious, time-consuming and messy task; (ii) meshing and re-meshing for a large computational domain is time-consuming, tedious, and costly task also requires the skills; (iii) in very complex computational domains, the mesh-based method fails in terms of accuracy; (iv) glass hour and shear locking phenomena generally found in traditional FEM. So, Chapter "Theoretical Approach to Element Free Galerkin Method and Its Mathematical Implementation" presents a new class of numerical methods known as mesh-free methods, such as Element Free Galerkin (EFG) method. The advantages of EFG method with mesh-based methods are highlighted on two elastostatic numerical problems: 1D problem of bar with body forces and 2D Timoshenko cantilever beam with traction at tip. The numerical results have been evaluated and compared with exact results as well.

Theoretical approach of Chebyshev Spectral Collocation (CSC) method and its mathematical implementation are detailed in Chapter "Theoretical Approach to Chebyshev Spectral Collocation Method and Its Mathematical Implementation." The stability analysis of the incompressible viscous flow between the two concentric counter-rotating vertical cylinders is selected as a representative example to apply the CSC method. The governing stability equations for disturbance flow quantities

are derived in cylindrical polar coordinates by coupling the energy equation with the Navier-Stokes equations, discretized using CSC method, modeled with appropriate boundary conditions form as a general eigenvalues problem, and analyzed for different Reynolds numbers, taking into account the effect of viscous heating, radius ratio, and buoyancy functionality.

In Chapter "Advanced Numerical Methods Based on Optimization" the unconstrained and constrained optimization algorithms for numerical methods are analyzed for fundamental problems in energy systems. Recent heuristic algorithms used in power supply systems have been presented and implemented, helping the reader to understand these optimization methods.

The ill-posed inverse electromagnetic and power engineering problems are introduced in Chapter "Ill-Posed Inverse Problems in Electrical Engineering Applications," both theoretically and applied, by detailing the numerical solution of case studies for several regularization techniques of energy systems.

The second part called *Advanced Energy Systems* makes in eight chapters a presentation of energy systems based on an advanced concept such as energy hubs, multi-energy sources, smart systems, and so on, respectively, the numerical analysis of energy systems used in different industrial sectors. The content of the eight chapters will be briefly presented below.

Chapter "Advanced Energy Systems Based on Energy Hub Concept" explains the concept of energy hub and makes a comprehensive overview of different applications of energy hubs in different energy consumption sectors, including residential, industrial, agricultural, and commercial. By integrating energy systems such as electricity, natural gasoline thought an energy hub, more benefits appears in optimal planning, control, and management of the energy sources.

Therefore, there has been a rapid movement toward multi-energy sources that have approached in Chapter "Sustainable Energy Systems Based on the Multi-energy Sources," where some examples of sustainable energy systems with different energy sources are given as well.

Smart homes can be also a good example of sustainable energy systems operating with different energy sources. Modeling, control, optimization, and technical-economic analysis of a smart home are presented in Chapter "Modeling of Energy Systems for Smart Homes."

One energy source that is recently used as a backup energy source in hybrid power system based on variable multi-energy sources such as Renewable Energy Sources (RES) is the fuel cell system, which is an electrochemical device that generates energy and water based on air and hydrogen. To analyze the phenomena of interests (fluid flow, heat and mass transfer, chemical reactions, etc.) in order to obtain optimized geometries and adequate operating parameters for different materials, the Finite Volume Method (FVM) is proposed in Chapter "Finite Volume Method Used for Numerical Investigations of Electrochemical Devices" to improve the performance of the electrochemical device. The chapter provides a general overview of the Finite Volume Method and Computational Fluid Dynamic (CFD), being applied to electrochemical devices used in energy systems in order to give insights into understanding the influence of different configurations (channel patterns, width and

depth, layer thickness), of operating conditions (flow rate, pressure, temperature), and of material characteristics and properties (catalyst microstructure, porosity, permeability) on the performance and durability. The numerical model has been experimentally validated revealing that the FVM is a useful tool that can lead to technology optimization and costs reduction.

Chapter "Night Operation of a Solar Chimney Integrated with Spiral Heat Exchanger" presents another energy application using the CFD commercial software package ANSYS Fluent. The study *focuses essentially* on the effects of the geometrical parameters of the collector, the meteorological conditions as well as the effectiveness of the heat exchanger on the air mass flow rate, the temperature rise within the collector, and the overall performance of the solar-geothermal hybrid power system. A geothermal heating device is used as a backup energy source to guarantee a continuous and stable operation during night hours.

Thus, the stable operation of energy systems is approached in Chapter "Incorporating of IPFC in Multi-machine Power System Phillips-Heffron Model" incorporating of Interline Power Flow Controller (IPFC) in Multi-Machine Power System Phillips-Heffron Model.

Note that the IPFC and Unified power flow controller (UPFC) are embedded devices of the latest generation from class of Flexible AC Transmission Systems (FACTS) devices used to improve the transfer capacity and controllability of energy systems. The numerical results obtained with MATLAB for the dynamic simulations of the energy system show improved stability by using IPFC.

If a backup power source is not available or is not used due to excessive costs of using it, then an energy storage system (ESS) must be used to ensure stable operation of the energy system by compensating energy flow balance for variable RES power. Chapter "Techno-Economical Analysis of Energy Storage Systems in Conventional Distribution Networks" presents a technical-economic analysis of energy storage systems in conventional distribution networks. The ESS management is performed in order to minimize the total cost of daily energy loss and energy supply of the system. For this, the optimization function includes as variable energy price, storage utilization duration, amount of load demand, power loss of the system, costs, limits and characteristics of storage system.

Proton Exchange Membrane Fuel Cell (PEMFC) is currently used in stationary and mobile applications. FC vehicles operate under high dynamic load conditions, so the energy management strategy must be validated using a mathematical model from Simulink/MATLAB in the Real Time (RT)-LAB platform such as OPAL-RT technology. The advantages of real-time numerical simulation are highlighted in Chapter "OPAL-RT Technology Used in Automotive Applications for PEMFC."

The third and last part, called *Numerical Energy Applications* highlights in 18 chapters the application of numerical methods for the control, optimization, and protection of energy systems. A brief presentation of the applications covered by each of the 18 chapters will be presented below.

Chapter "Theoretical Techniques for the Exploration of Piezoelectric Harvesters" analyzes the design and optimization of piezoelectric harvester systems using Finite Element Analysis (FEA) and time integration techniques from MatLab toolbox and

Maxwell 3D package. Readers are provided with useful information to create original and efficient harvesters based on the magnetostatic and magnetodynamic relations presented in this chapter.

In general, the electromagnetism problems can be formulated using Maxwell equations, but analytical solutions are easily obtained only for some objects with standard geometric shapes, such as sphere, infinite plane, elementary antennas, etc. Chapter "Numerical Analysis of Electromagnetic Fields" presents the numerical methods that can accurately estimate the imposed electromagnetic field for practical problems. These methods can be classified into three types: ray-tracing, surface discretization, and volume discretization. Depending on the achieved precision, these methods are classified into full-wave and asymptotic methods, also called low and high-frequency methods. Methods based on volumetric discretization, such as Finite-Difference Time-Domain (FDTD) and frequency domain finite-element method (FEM), have the advantages of easy modeling of non-homogeneous media. The methods based on the discretization of surfaces using Surface Integral Equations (SIE) are computationally more efficient than the volumetric ones. The optimized variant of the method-of-moments (MoM) such as fast multipole method (FMM) and multilevel fast multipole algorithm (MLFMA) and the physical optics (PO) are analyzed in this chapter by numerical examples.

Wireless Energy Transfer or Wireless Power Transfer (WPT) is a new technology, which transfers energy through electromagnetic field to inaccessible places or distributed systems, so at lower costs, because wiring is eliminated. Chapter "Optimization Methods for Wireless Power Transfer" analyses the efficiency of three types of WPTs (radiant transfer, inductive transfer, and resonant coupling transfer), that strongly depends on the resonator parameters (L—self-inductance, M—mutual inductance, C—parasitic capacitance, and R—Ohmic resistances) of the two magnetically coupled coils, placed at different distances and angles in assemblies with several configurations. The mutual inductances computed with MATLAB utilizing the integration, the numerically calculated ones using ANSYS Q3D Extractor, and respectively those obtained through measurements show close values, indicating a consistency regarding all three methods of parameter determination.

Numerical Assessment of Electromagnetic Energy and Forces in Non-Destructive Measurement Devices is analyzed in Chapter "Numerical Assessment of Electromagnetic Energy and Forces in Non-destructive Measurement Devices." It is known that *non-destructive testing* is vital to ensure that pieces of a particular structure perform their specific functions for a predetermined amount of time. The specialists in the non-destructive testing field have created and implemented tests to characterize the materials or to detect, localize, and measure the flaws (defects), which can cause plane crashes, nuclear power plant explosions, dam breakage, train derailment, fires, and other less visible events, but with dangerous consequences and high impact.

Optimal integration of electric vehicles in smart grid energy flow is approached in Chapter "Optimal Integration of Electric Vehicles in Smart Grid Energy Flow," including modeling, simulation, and result comparison.

Numerical Approaches of Biomass Plants Efficiency are discussed in Chapter "Numerical Approaches of Biomass Plants Efficiency." The chapter includes a

description of the three distinct phases of the technological process of producing energy from biogas, starting with the type and quantities of the raw materials used and how much energy the Cogeneration Power Plant (CPP) can produce. The data acquisition system is part of a sophisticated automated system called "DIANE," which permanently monitors, coordinates, and controls all the operations in the cogeneration power plant. Following the measurement of many parameters of the electricity production process due to each generator, the relationship between electricity production and biomass consumed in the form of polynomial functions was obtained.

The answer to the question "what is the state in which conservative systems consume less power or energy?" is approached in Chapter "Power and Energy Flow in Cvasi-Stationary Electric and Magnetic Circuits," where power and energy flow in cvasi-stationary electric and magnetic circuits is analyzed. Therefore, advanced numerical analysis is proposed to find the extreme point of power or energy functionals for electric and magnetic circuits in the quasi-stationary state, where energy functional is the total energy of a certain system, as a functional of the system's state. Lagrange multipliers and the variational method in Hilbert space have demonstrated the existence of the minimum of the functionals. Several examples implemented in PSPICE prove the theoretical result of minimum consumed power principle (PMCP).

Chapter "Numerical Methods for Analysis of Energy Consumption in Drying Process of Wood" presents general aspects regarding the electromagnetic field in radio frequency and microwaves, the thermal field, mass problems in radio frequency drying, and the numerical analysis of high-frequency drying using FEM-BEM.3D-RFmove_term_masa software in radio frequency field and Comsol Multiphysics in microwave field.

After the optimization and control of PEMFC system shown in chapters "Finite Volume Method Used for Numerical Investigations of Electrochemical Devices" and "OPAL-RT Technology Used in Automotive Applications for PEMFC," the design and energy efficiency analysis of a Fuel Cell Hybrid Electric Vehicle (FCHEV) is presented in Chapter "Design and Energy Analysis for Fuel Cell Hybrid Electric Vehicle." FCHEV combines the advantage offered by PEMFC as the main energy system by using an efficient energy management strategy (EMS) for the hybrid battery/ultracapacitor ESS to operate the PEMFC system safely. The FCHEV performance obtained in simulation using standardized load cycles is validated by taking into account a real experimental speed profile and numerical analysis of the acquired data. The efficiency obtained by FCHEV and the electric vehicle (using only battery) is presented as well.

Chapter "Finite Element Solutions for Magnetic Shielding Power Applications" presents the applications of the Finite Element Method (FEM) to magnetic shielding for cables in power energy applications, including FCHEVs. The topic is important from an electromagnetic compatibility point of view as well as for people health hazards. It is known that energy systems such as power lines, underground cables, low/medium voltage substations, and building electrical distribution systems are electromagnetic noise sources. Therefore, in order to comply with the standards for

maximum permissible magnetic field levels, the use of shielding devices is required and the harmonic level must be monitored continuously.

Also, the decision-making in the electric distribution systems is based on data collected from consumers and the various measurement points located in the network (transformer substations, supply points, branch points, etc.) through the Supervisory Control and Data Acquisition (SCADA) system. Chapter "Regression Analysis-Based Load Modelling for Electric Distribution Networks" proposes a regression analysis-based load modeling for electric distribution networks based on available data that help the Distribution Network Operator (DNO) to accurately estimate the state of the supervised system. The approaches refer to estimation of the active and reactive powers from the LV/MV (low voltage/medium voltage) electric substations with a mixt load structure (residential, commercial, and industrial) at the peak load of the system and the required load of residential consumers which represent the highest percentage from the load structure fed from the LV/MV electric substations. The proposed approaches were tested in real operation conditions of MV distribution networks from Romania.

It is known that the Overhead High-Voltage Power lines (OHVPLs) are major sources of extremely low frequency electric and magnetic fields, which can induce electric currents in the human body. The Finite Element Method (FEM) is proposed in Chapter "Finite Element Analysis of Electromagnetic Fields Emitted by Overhead High-Voltage Power Lines" to compute and analyze—from the perspective of public exposure—both electric and magnetic fields associated with typical configurations of OHVPLs used in the Romanian power system. Compliance with the exposure limits established by the International Commission on Non-Ionizing Radiation Protection (ICNIRP) for the general public is obtained in all cases: the calculated magnetic fields being below the ICNIRP limit of 100 μT, while the electric field levels exceed the ICNIRP limit of 5000 V/m only in limited areas beneath the 400 kV lines.

The Finite Element Method (FEM) is also proposed in Chapter "Design and Finite Element Analysis of Permanent Magnet Synchronous Generator for Wind Turbine Application" to Design and optimize a permanent magnet synchronous generator for wind turbine application.

The Modified Tellegen Principle is proposed in Chapter "Power and Energy System Modeling Based on Modified Tellegen Principle" to model the energy systems. The instantaneous calculation of power for real linear and nonlinear systems allows obtaining results that are correct both mathematically and physically.

The design procedure of a Machine Learning (ML)-based yaw control strategy for a Horizontal Axis Wind Turbine (HAWT) is presented in Chapter "Self-Tuning Yaw Control Strategy of a Horizontal Axis Wind Turbine Based on Machine Learning." The proposed yaw control strategy is based on the interaction of three different Artificial Intelligence (AI) techniques to design an ML system: Reinforcement Learning (RL), Artificial Neural Networks (ANN), and metaheuristic optimization algorithms. The implementation of a metaheuristic optimization algorithm, in this case, a Particle Swarm Optimization (PSO) algorithm, allows calculation of the optimal yaw control action that responds to the compromise between the generated power increment and the mechanical loads increase due to the yaw actuation.

Various algorithms of Gauss-Seidel, Newton-Raphson, Fast-Decoupled Load Flow are presented in Chapter "Numerical Methods of Electric Power Flow in Interconnected Systems" for interconnected power systems in different case of buses (load, voltage controlled, and slack buses) and various scenarios: with PQ buses, with PV (generators) buses, with PV buses, taking into consideration the limitation of the generated reactive power and so on.

Numerical Methods in Selecting the Location of Distributed Generation in Energy Network is presented in Chapter "Numerical Methods in Selecting Location of Distributed Generation in Energy Network." The advantages and disadvantages of using Distributed Generators (DGs) in a distribution network are highlighted. DG placement methods based on intelligent algorithms and numerical optimization methods are analyzed. The obtained results highlight a better performance of numerical optimization methods in terms of execution time and reduction of power losses, so their use in high dimensional networks is recommended.

Beside the Interline Power Flow Controller (IPFC) analyzed in chapter "Incorporating of IPFC in Multi-machine Power System Phillips-Heffron Model," other Flexible AC Transmission Systems (FACTS) such as Static VAR Compensator (SVC), Thyristor Controlled Series Compensation (TCSC), and Unified Power Flow Controller (UPFC) are analyzed in Chapter "Numerical Methods for Power System Analysis with FACTS Devices Applications." First, the mathematical modeling of SVC, TCSC, and UPFC devices are described. Then, applications of FACTS devices to improve transient stability, small-signal stability, voltage profile, and reduce power losses and inrush currents caused by transformers are presented.

Firstly, the mathematical modeling of SVC, TCSC and UPFC devices is described. Then, applications of FACTS devices to improve the transient stability, small signal stability and voltage profile, and reduce power losses and inrush currents caused by transformers are presented. Simulations performed on the IEEE 14 bus system using Newton Raphson method to solve equations of the optimization functions confirm that the optimal placement of FACTS devices improves most power system specifications.

In conclusion, the book includes enough funds to understand the modeling of an energy application and selecting the most appropriate method from the set of possible numerical methods to solve it, presenting one or more numerical methods for solving the energy applications included in the book and providing many numerical examples to understand the proposed method. As a large number of electrical engineers have participated in the writing of this book, we are confident that it will be helpful for young electrical researchers and practitioners working in or for energy systems to optimize the operation of these systems.

Tabriz, Iran — Naser Mahdavi Tabatabaei
Pitesti, Romania — Nicu Bizon

Contents

Contributors

Masoud Alilou Energy Management Research Center, University of Mohaghegh Ardabili, Ardabil, Iran

Horia Andrei Doctoral School of Engineering Sciences, University Valahia, Targoviste, Romania

Paul Cristian Andrei Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania

Livia Bandici Department of Electrical Engineering, University of Oradea, Oradea, Romania

Rameshkumar Bhoraniya Marwadi University, Rajkot, India

Javier Bilbao Applied Mathematics Department, University of the Basque Country, Bilbao, Spain

Nicu Bizon Faculty of Electronics, Communications and Computers, Department of Electronics, Computers and Electrical Engineering, University of Pitesti, Pitesti, Romania

Lavinia Bobaru Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania

Eugenio Bravo Applied Mathematics Department, University of the Basque Country, Bilbao, Spain

Elena Carcadea National Research and Development Institute for Cryogenics and Isotopic Technologies, ICSI, Rm. Valcea, Romania

Stefan Castravete CAELYNX-Europe, Craiova, Romania

Dumitru Cazacu University of Pitesti, Pitesti, Romania

Andrei Ceclan Numerical Methods Research Center, Department of Electrical Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Milad Chanddel Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran

Keyvan Choopani Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran

Simina Coman Department of Computers and Information Technology, University of Oradea, Oradea, Romania

Levente Czumbil Numerical Methods Research Center, Department of Electrical Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Sorin Deleanu School of Applied Sciences and Technology, Northern Alberta Institute of Technology (NAIT), Edmonton, Canada

Tugce Demirdelen Electrical and Electronic Engineering Department, Adana Alparslan Turkes Science and Technology University, Adana, Turkey

Amel Dhahri Research Unit: Materials, Energy and Renewable Energies, College of Sciences, University of Gafsa, Gafsa, Tunisia

Emil Diaconu Faculty of Electrical Engineering, Electronics and Information Technology, University Valahia, Targoviste, Romania

Hidir Duzkaya Electrical and Electronics Engineering Department, Engineering Faculty, Gazi University, Ankara, Turkey;
Ampre Laboratory CNRS UMR, Ecole Centrale de Lyon, Ecully, France

Reza Effatnejad Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran

Firat Ekinci Energy Systems Engineering Department, Adana Alparslan Turkes Science and Technology University, Adana, Turkey

Alexandru Enescu Doctoral School of Engineering Sciences, University Valahia, Targoviste, Romania

Burak Esenboga Electrical and Electronic Engineering Department, Adana Alparslan Turkes Science and Technology University, Adana, Turkey

Unai Fernandez-Gamiz Nuclear Engineering and Fluid Mechanics Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain

Amin Foroughi Nematollahi Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran

Marian Gaiceanu Department of Automatic Control and Electrical Engineering, University Dunarea de Jos Galati, Galati, Romania;
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania

Theodora Gaiceanu Faculty of Automatic Control and Computer Engineering, Gheorghe Asachi Technical University of Iasi, Iasi, Romania;

Department of Automatic Control and Applied Informatics, Gheorghe Asachi Technical University of Iasi, Iasi, Romania

Olatz Garcia Applied Mathematics Department, University of the Basque Country, Bilbao, Spain

Iulian Ghenea Doctoral School of Fundamental and Engineering Sciences, University Dunarea de Jos of Galati, Galati, Romania

Gheorghe Grigoras Electrical Engineering Faculty, Power System Department, "Gheorghe Asachi" Technical University of Iasi, Iasi, Romania

Mahdi Hedayati Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran

Valeriu Manuel Ionescu University of Pitesti, Pitesti, Romania

Mihai Iordache Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania

Rajendrasinh Jadeja Marwadi University, Rajkot, India

Ramdevsinh Jhala Marwadi University, Rajkot, India

Erol Kurt Department of Electrical and Electronics Engineering, Technology Faculty, Gazi University, Ankara, Turkey

Hatice Hilal Kurt Department of Physics, Faculty of Sciences, Gazi University, Ankara, Turkey

Teodor Leuca Department of Electrical Engineering, University of Oradea, Oradea, Romania;
Academy of Romanian Scientists, Bucharest, Romania

Jose Manuel Lopez-Guede Automatic Control and System Engineering Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain

Eduard Lunca Department of Electrical Measurements and Materials, "Gheorghe Asachi" Technical University of Iasi, Iasi, Romania

Ioana Manta National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenic and Isotopic Technologies, Ramnicu Valcea, Romania;
Faculty of Power Engineering, Polytechnic University of Bucharest, Bucharest, Romania

Adriana Marinoiu National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenics and Isotopic Technologies ICSI Rm. Valcea, Valcea, Romania

Daniel Mayer Department of Electrical and Computational Engineering, University of West Bohemia, Plzen, Czechia

Shahram Mehdipour-Ataei Faculty of Polymer Science, Iran Polymer and Petrochemical Institute, Tehran, Iran

Dan D. Micu Numerical Methods Research Center, Department of Electrical Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Abas Moghadasi Department of Electrical Engineering, Tehran South Branch, Islamic Azad University, Tehran, Iran

Amir Mohammad Moghaddam Energy Management Research Center, University of Mohaghegh Ardabili, Ardabil, Iran

Maryam Mohammadi Faculty of Polymer Science, Iran Polymer and Petrochemical Institute, Tehran, Iran

Narges Mohammadi Faculty of Mechanical Engineering, Amirkabir University of Technology, Tehran, Iran

Bhaumik Nagevadiya Marwadi University, Rajkot, India

Bogdan Constantin Neagu Power System Department, Electrical Engineering Faculty, Gheorghe Asachi Technical University of Iasi, Iasi, Romania

Dragos Niculae Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania

Ahmed Omri Research Unit: Materials, Energy and Renewable Energies, College of Sciences, University of Gafsa, Gafsa, Tunisia

Jamel Orfi Department of Mechanical Engineering, King Saud University, Riadh, Saudi Arabia

Pinank Patel Marwadi University, Rajkot, India

Maria Simona Raboaca National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenic and Isotopic Technologies, Ramnicu Valcea, Romania

Mircea Raceanu National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenics and Isotopic Technologies ICSI Rm. Valcea, Valcea, Romania

Abolfazl Rahiminejad Department of Electrical and Computer Science, Esfarayen University of Technology, Esfarayen, North Khorasan, Iran

Jose Antonio Ramos-Hernanz Electrical Engineering Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain

Gabriela Rata Faculty of Electrical Engineering and Computer Science, Stefan Cel Mare University of Suceava, Suceava, Romania

Mihai Rata Faculty of Electrical Engineering and Computer Science, Stefan Cel Mare University of Suceava, Suceava, Romania

Carolina Rebollar Applied Mathematics Department, University of the Basque Country, Bilbao, Spain

Aitor Saenz-Aguirre Nuclear Engineering and Fluid Mechanics Department, University of the Basque Country, Eibar, Spain

Hossein Shayeghi Energy Management Research Center, University of Mohaghegh Ardabili, Ardabil, Iran

Vasile Solcanu Doctoral School of Fundamental and Engineering Sciences, University Dunarea de Jos of Galati, Galati, Romania

Marilena Stanculescu Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania

Milan Stork Department of Electronics and Information Technology/RICE, University of West Bohemia, Plzen, Czechia

Nemat Talebi Department of Electrical Engineering, Tehran South Branch, Islamic Azad University, Tehran, Iran

M. Cengiz Taplamacioglu Electrical & Electronics Engineering Department, Engineering Faculty, Gazi University, Ankara, Turkey

Suleyman Sungur Tezcan Electrical and Electronics Engineering Department, Engineering Faculty, Gazi University, Ankara, Turkey

Behrooz Vahidi Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran

Concepcion Varela Applied Mathematics Department, University of the Basque Country, Bilbao, Spain

Mihai Varlam National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenics and Isotopic Technologies ICSI Rm. Valcea, Valcea, Romania

Elena Otilia Virjoghe University Valahia, Targoviste, Romania

Silviu Vornicu Department of Electrical Measurements and Materials, "Gheorghe Asachi" Technical University of Iasi, Iasi, Romania

Abdurrahman Yavuzdeger Energy Systems Engineering Department, Adana Alparslan Turkes Science and Technology University, Adana, Turkey

Mehmet Zile Mersin University, Mersin, Turkey

Ekaitz Zulueta Automatic Control and System Engineering Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain

Advanced Numerical Methods

Advanced Numerical Methods for Equations, Systems Equations and Optimization

Horia Andrei, Dan D. Micu, Marian Gaiceanu, Marilena Stanculescu, and Paul Cristian Andrei

Abstract In this chapter an overview of advanced numerical methods is presented. After errors are defined, the readers are initiated with the principles of approximating functions, numerical methods for solving equations and systems equations and optimization methods. The theoretical basics of polynomial interpolation, numerical differentiation and numerical integration are presented, and, as a natural sequel, for each method some examples are given. Numerical methods for linear and nonlinear equations and system equations, numerical methods for computing eigenvalues and eigenvectors are discussed and examples are given. The main goal of the optimization methods is to find the maximum or minimum of the objective function by using linear or nonlinear programming as shown in this chapter. Finally, an introduction in Matlab is done to put in evidence the easy-to-use and attractiveness of this popular software. Each method is accompanied by examples to help understanding.

Keywords Errors · Approximating functions · Numerical methods for solving equations and systems equations · Optimization methods · Matlab applications

H. Andrei (✉)
Doctoral School of Engineering Sciences, University Valahia, Targoviste, Romania
e-mail: paul.andrei@upb.ro; hr_andrei@yahoo.com

D. D. Micu
Numerical Methods Research Center, Technical University of Cluj-Napoca, Cluj-Napoca, Romania
e-mail: Dan.Micu@ethm.utcluj.ro

M. Gaiceanu
Department of Control Systems and Electrical Engineering, University Dunarea de Jos Galati, Galati, Romania
e-mail: marian.gaiceanu@ugal.ro

M. Stanculescu · P. C. Andrei
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania
e-mail: marilena.stanculescu@upb.ro

P. C. Andrei
e-mail: paul.andrei@upb.ro; hr_andrei@yahoo.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_1

Nomenclatures

A.	**Acronyms**
Matlab	Matrix Laboratory
OS	Operating Systems
Mac OS X	Macintosh Operating System
B.	**Symbols/Parameters**
$\tilde{x}$	The approximate value
x	The true (unknown) quantity
ε	The error
$\|\varepsilon\|$	The absolute error
ε_r	The relative error
f(x)	The function of real variable x
A	The Vandermonde matrix
tk	The k-moment of time
p	The p-step of numerical algorithm
xr	The exact solution
xo	The first solution
$\Re$	The real set
T	The matrix of the coefficient
X	The vector the unknowns
G	The column vector of free terms
det$T$	The determinant of matrix T

1 Introduction

The challenge of this first chapter of the book is to synthesize in a few pages thousands pages of research and studies. More problems from different domains as all fields of engineering sciences, medicine, physics sciences (meteorology, pollution, astronomy, etc.), chemistry, civil engineering, life and social sciences, business, arts, etc. can be evaluated and approximated by mathematical model. In modeling, simulation, design and optimization procedures the numerical analysis and methods have the most important contribution. Also, the data acquisition systems and forecasting methods are based on evolved numerical algorithms. Starting from mathematical principles and from the development of computer sciences the numerical methods give correct solutions to many real and difficult problems.

Generally numerical methods can be classified in direct methods which compute the solution in a specified number of iterations and deliver the correct solution when they were performed in infinite arithmetic precision, respectively the more common iterative methods that, start from an initial state, develop successive steps of approximations which converge to the correct solution as a limit of a sequence of values

obtained from iterations. The study of errors represents an important part of numerical analysis. Thus the generation and propagation of errors in numerical approaches are presented in first section of chapter.

The significant growth both of hardware architectures and programming languages has led to the development of advanced numerical methods. It deals with the approximation of functions and provides the different methods for solving linear and nonlinear equations and systems, calculation of eigenvalues and eigenvectors, determination of optimum point of a defined objective function. Therefore in Sect. 2 the approximating functions are defined. The theoretical basics of polynomial interpolation, numerical differentiation and numerical integration are presented, and, as a natural sequel, for each method some examples are given.

Another important part of advanced numerical methods that can be applied to real-world problems is represented by methods for solving equations and systems equations. Numerical methods for linear and nonlinear equations and system equations, numerical methods for computing eigenvalues and eigenvectors are discussed and examples are given.

Optimization methods solve more practical in various areas such as engineering sciences, financial, aeronautical and terrestrial routs, and environmental pollution. Starting from defining an objective functions which satisfy some constraints, the main goal of these methods is to find the maximum or minimum of the objective function by using linear or nonlinear programming as shown and exemplified in Sect. 5.

Section 6 refers to Matlab as being the most widely used programming environment in which numerical algorithms are implemented. Several examples put in evidence the easy-to-use and attractiveness of this popular software.

The main conclusions and the bibliographic references are presented in the last two sections of the chapter.

2 Generation and Propagation of Errors

Any measuring procedure for determining an arbitrary quantity is subjected errors, so the numerically measured or determined value does not represent its true or exact value. Therefore, the errors appear and are defined in the measurement processes and in the numerical computation procedures, which, often, can be implemented in the same computation system [1].

In general, if we denote by $\tilde{x}$ the approximate value of the true (unknown) quantity x, then the error ε of determination by measurement and/or by numerical computation of it is defined as [2]:

$$\varepsilon = x - \tilde{x} \tag{1}$$

It is obvious that in relationship (1) it is known, by measurement or by numerical computation, only $\tilde{x}$ such that for a precise determination very close to the reality, of the true quantity x, an upper limit of the error must be specified. This limitation of the error gives mathematical meaning to the relationship (1).

The absolute error $|\varepsilon|$ is defined as:

$$|\varepsilon| = |x - \tilde{x}| \tag{2}$$

and the relative error ε_r is the ratio between the absolute error and the approximate value module:

$$\varepsilon_r = \frac{|\varepsilon|}{|\tilde{x}|} = \frac{|x - \tilde{x}|}{|\tilde{x}|} \tag{3}$$

The errors that appear in the measurement processes and in the numerical computation procedures are classified as follows [3]:

- *rounding errors* which are due to floating point computation performed by the computer;
- *inherent errors* that occur as a result of the measuring process and the accuracy of the measuring instrument, errors of the reading and data entering into the computer (these are random errors characterized by the normal or Gaussian distribution function) or error resulting from the necessity of representing in the computer a value using a finite number of digits (for example the numbers $\pi = 3.14192654\ldots$, $\sqrt{2} = 1.41421356\ldots$, $e = 2.73\ldots\ldots$ are approximated using a finite number of digits although they contain an infinite number of digits);
- *truncation errors* which are the result of the used numerical calculation algorithm, such as, for example, stopping (truncating) the development of a Taylor power series at a certain term for computing numerically the function in a specified point.

In numerical computational methods, the errors propagate naturally. This is because, as shown above, the numerical determination of an arbitrary quantity x is done with a certain error and then, obviously, each of the classical operations: (+), (−), (×) and (/) performed by the computer introduce errors, which are propagated by applying them on several operands. There are proved [4] propagation formulae for each type of error as they are defined in the relationships (1), (2) and (3).

For example, let us consider two random quantities, of the same physical type and having the dame unit of measure, x and y, whose approximate values are $\tilde{x}$ and $\tilde{y}$, with absolute and relative errors ε_x, $\varepsilon_{r,x}$ respectively ε_y, $\varepsilon_{r,y}$. Under these conditions, using the relationship (2) in which, for simplification, it is considered that all the values are positive, the absolute error, obtained when one uses the addition operation between two numbers, propagates as the sum of the absolute errors of determination of the two quantities, i.e. [5, 6].

$$|\varepsilon_+| = \varepsilon_+ = |(x + y) - (\tilde{x} + \tilde{y})| = (x + y) - (\tilde{x} + \tilde{y}) = \varepsilon_x + \varepsilon_y \tag{4}$$

In a similar manner, for multiplication operation, the error propagation is given by the following relationship:

$$|\varepsilon_\times| = \varepsilon_\times = |(x \times y) - (\tilde{x} \times \tilde{y})| = (x \times y) - (\tilde{x} \times \tilde{y}) = \varepsilon_x \tilde{y} + \varepsilon_y \tilde{x} \tag{5}$$

where it was considered the product $\varepsilon_x \varepsilon_y \approx 0$, taking into account that the upper imposed limits for the two absolute errors must be as small as possible, for a better computation accuracy. As follows, using the relationships (3) and (4) it is determined the propagation error, relative to the addition operation of two operands, which is:

$$\varepsilon_{r,+} = \frac{|\varepsilon_+|}{\tilde{x} + \tilde{y}} = \frac{\varepsilon_+}{\tilde{x} + \tilde{y}} = \frac{\tilde{x}}{\tilde{x} + \tilde{y}} \varepsilon_{r,x} + \frac{\tilde{y}}{\tilde{x} + \tilde{y}} \varepsilon_{r,y} \tag{6}$$

If the relationships (3) and (5) are used, then the propagation of the error relative to the multiplication operation becomes:

$$\varepsilon_{r,\times} = \frac{|\varepsilon_\times|}{\tilde{x} \times \tilde{y}} = \frac{\varepsilon_\times}{\tilde{x} \times \tilde{y}} = \varepsilon_{r,x} + \varepsilon_{r,y} \tag{7}$$

The two examples of computing the propagation of the absolute and relative errors for two essential arithmetic operations in numerical calculation, show, on one hand the necessity of using measuring equipment with as small as possible inherent errors and, on the other hand, the necessity of implementing some efficient numerical methods which uses symbolic algorithms, with few calculations, but also modern computers with high computation power.

3 Approximating Functions

3.1 Polynomial Interpolation

The interpolation problem consists in determining some functions which allow the evaluation of a function defined by tables of values in any point of the definition domain, not only in the network's nodes [7–9].

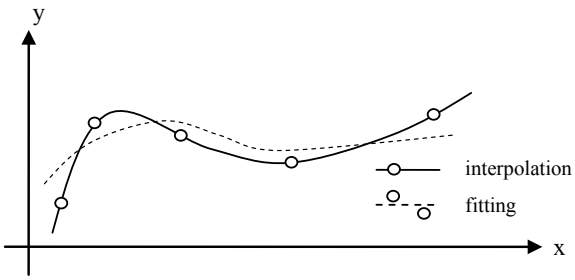


Fig. 1 The difference between function interpolation and fitting

Unlike interpolation, function approximation or function fitting consists in establishing some analytical expressions for approximating, as well as possible, the behaviour of a function defined by its values, without imposing the restriction that the function must pass through all the points in which the given function is defined (Fig. 1). These kinds of problems are encountered in practice in establishing some empirical formulae for approximating some experimentally defined functions. The solution of such problems in not unique, it depends on the functions' class in which a solution is sought, as well as for the error estimation criteria.

As follows, only interpolation problems will be addressed.

Consider a real function $y = f(x)$ of real variable x, defined by the distinct values:

x	x_0	x_1	x_2	…	x_n
y	y_0	y_1	y_2	…	y_n

It is required to estimate $f(x)$ for other values of variable x than those from the table of values. This can be realized by determining a function $g(x)$ that satisfies the *interpolation conditions*, i.e. to pass through the given points:

$$g(x_i) = y_i, i = 0, \ldots, n \tag{8}$$

The problem solution is not unique. For this reason, to the interpolation function $g(x)$ some supplementary conditions are imposed. It is suspected that, from here, many approaches can be followed, function of the specific conditions. In principle, the interpolation function can be represented as follows:

$$g(x) = \sum_{k=0}^{n} c_k \varphi_k(x) \tag{9}$$

where $\varphi_k(x)$ are $n + 1$ independent linear functions, that define a base. There are two main interpolation method classes:

1. Global interpolation, by which it is determined only one interpolation function for the entire function definition domain.

2. Interpolation on portions, by functions defined on each of the interval defined by the nodes. These functions are connected such that the certain continuity conditions must be fulfilled. Beside the continuity of the values, which is an explicit condition, supplementary conditions can be imposed, such as the continuity of the 1st order derivatives (the connection does not create "corners"), etc.

An essential aspect must not be forgotten: the functions defined by experimental values can be affected by errors, but the precision of the given values cannot exceed some practical limitations. Therefore, it should not be exaggerated to present results that suggests greater accuracy than the one given by the available data (for example, the presentation of the results with many significant digits, although in practice 3 digits are reasonable).

3.1.1 Global Interpolation

In global interpolation, all base functions $\varphi_k(x)$ are explicitly defined, by only one expression, for the whole function definition domain. Function of the choice of the base, there are various interpolation methods. In most cases, these functions are polynomial type.

(i) *Direct polynomial interpolation*

The functions $\varphi_k(x)$ are chosen of the following type:

$$\varphi_k(x) = x^k, k = 0, \ldots, n \tag{10}$$

Therefore, the polynomial interpolation function is:

$$g(x) = \sum_{k=0}^{n} c_k x^k \tag{11}$$

The problem is reduced to the determination of the coefficients c_k, imposing the interpolation condition (8). In this respect the obtained system of $n + 1$ equation becomes:

$$g(x_i) = \sum_{k=0}^{n} c_k x_i^k = y_i,\ i = 0, \ldots, n \tag{12}$$

The matrix expression of the system is:

$$\mathbf{Ac} = \mathbf{y} \tag{13}$$

in which:

$$\mathbf{A} = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix}; \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix}; \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} \tag{14}$$

The matrix **A**, called *Vandermonde* matrix, is non-singular, if the values x_i are distinct.

This method is uninteresting, because it requires a significant computational effort (solving a system of order n), and the matrix **A** is, in general, ill-conditioned, which can lead to big errors.

(ii) *Newton interpolation*

The functions $\varphi_k(x)$ are chosen to have the following form:

$$\begin{cases} \varphi_0(x) = 1 \\ \dots \\ \varphi_k(x) = (x - x_0)(x - x_1) \dots (x - x_{k-1}); k = 1, \dots, n \end{cases} \tag{15}$$

By imposing the interpolation condition, it results the following system of equations:

$$\begin{cases} g(x_0) = y_0 = c_0 \\ g(x_1) = y_1 = c_0 + c_1(x_1 - x_0) \\ g(x_2) = y_2 = c_0 + c_1(x_2 - x_0) + c_2(x_2 - x_0)(x_2 - x_1) \\ \dots \\ g(x_n) = y_n = c_0 + c_1(x_n - x_0) + c_2(x_n - x_0)(x_n - x_1) \end{cases} \tag{16}$$

This is *lower-triangular* type system, whose solution is easily determined:

$$\begin{cases} c_0 = y_0 \\ c_1 = (y_1 - c_0)/(x_1 - x_0) = (y_1 - y_0)/(x_1 - x_0) \\ c_1 = (y_2 - c_0 - c_1(x_2 - x_0))/(x_2 - x_0)/(x_2 - x_1) \\ \dots \dots \dots \end{cases} \tag{17}$$

Newton method presents a series of advantages. First, the determination of the coefficients implies a reduced computation effort, the system is relatively well-conditioned, new points can be added, with the possibility of partial recovery of the old results, etc.

(iii) *Lagrange interpolation. Lagrange interpolation of a function defined on the following interval:*

$$x_0 < x_1 < \dots < x_k < \dots < x_n \tag{18}$$

Is based on the Lagrange polynomials:

$$l_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x - x_i}{x_k - x_i}; k = 0, \ldots, n \tag{19}$$

It is found that these polynomials have the properties:

$$l_k(x_i) = \delta_{ki} = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases} \tag{20}$$

3.1.2 Interpolation on Portions

Global interpolation can lead to results far from those expected in case of functions with a more specific behaviour.

In such situations, it is advisable to use some methods of interpolation on portions, on which are sought different expressions of the interpolation functions on the intervals defined by the network of the definition points of the function to be interpolated. In principle, this can be done by adopting a base function $\varphi_k(x)$ distinct on the intervals defined by the definition network of the function to be interpolated.

(i) *Linear interpolationon portions*

Under this name it is hidden the most trivial interpolation method, which consists in adopting an interpolation function defined by straight segments which pass through the definition points of the function to be interpolated (Fig. 2).

In these points the considered interpolation function is continuous, but not derivable.

From the formal point of view, one can define a set of base functions of Lagrange polynomial type, defined on sintervals, according to the relationships (Fig. 3):

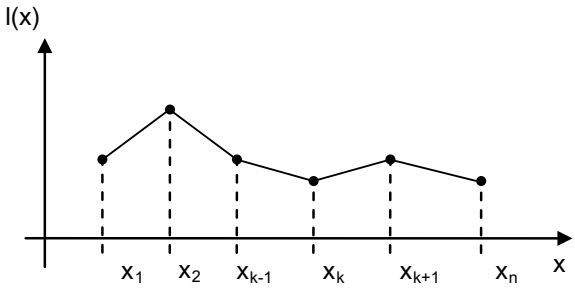


Fig. 2 Linear piecewise interpolation

Fig. 3 Lagrange polynomials type

$$l_1(x) = \begin{cases} \dfrac{x - x_2}{x_1 - x_2} & x \in [x_1, x_2] \\ 0 & x \notin [x_1, x_2] \end{cases}$$

$$l_k(x) = \begin{cases} \dfrac{x - x_{k-1}}{x_k - x_{k-1}} & x \in [x_{k-1}, x_k] \\ \dfrac{x - x_{k+1}}{x_k - x_{k+1}} & x \in [x_k, x_{k+1}], \quad k = 2,..,n-1 \\ 0 & x \notin [x_{k-1}, x_{k+1} \end{cases}$$

$$l_k(x) = \begin{cases} \dfrac{x - x_{n-1}}{x_k - x_{k-1}} & x \in [x_{n-1}, x_n] \\ 0 & x \notin [x_{n-1}, x_n] \end{cases}$$

(ii) *Interpolation using Hermite polynomials*

If it is necessary for the interpolation function to be not only continuous, but also derivable, the interpolation on portions using Hermite polynomials can be adopted. It is considered an interval in the definition domain of the function to be interpolated *y(x)*:

$$x_k \leq x \leq x_{k+1}; k = 1, \ldots, n-1 \tag{21}$$

and a local variable:

$$s = x - x_k \in [0, h_k]; h_k = x_{k+1} - x_k; k = 1, \ldots n-1 \tag{22}$$

It is also defined (Fig. 4):

$$\delta_k = \frac{y_{k+1} - y_k}{x_{k+1} - x_k} = \frac{y_{k+1} - y_k}{h_k}; k = 1, \ldots, n-1 \tag{23}$$

It is considered as auxiliary quantities the values of the derivatives of the interpolation polynomials in the extremities of the segments on which they are defined:

$$d_k = P'(x_k) \tag{24}$$

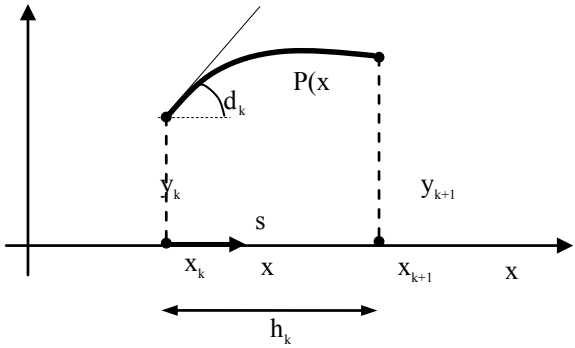


Fig. 4 Polynomial interpolation on portions

Most often, the interpolation is done using cubic polynomials (of degree 3), which offer enough flexibility to impose suitable interpolation conditions. Such a polynomial is (there for simplicity, we denoted $h = h_k$):

$$P(x) = \frac{3hs^2 - 2s^3}{h^3} y_{k+1} + \frac{h^3 - 3hs^2 + 2s^3}{h^3} y_k + \frac{s^2(s-h)}{h^2} d_{k+1} + \frac{s(s-h)^2}{h^2} d_k \tag{25}$$

It can be verified that this polynomial satisfies the *interpolation conditions*:

$$\begin{aligned} &s = 0 \Rightarrow P(x_k) = y_k \\ &s = h \Rightarrow P(x_{k+1}) = y_{k+1} \end{aligned} \tag{26}$$

To determine these polynomials (Hermite) it is necessary to know the derivatives d_k in the extremities of the segments. The derivative of a polynomial (as function of s) is:

$$P'(x) = \frac{6hs - 6s^2}{h^3} y_{k+1} + \frac{-6hs + 6s^2}{h^3} y_k + \frac{3s^2 - 2sh}{h^2} d_{k+1} + \frac{3s^2 - 4sh + h^2}{h^2} d_k \tag{27}$$

Consequently:

$$\begin{aligned} &s = 0 \Rightarrow P'(x_k) = d_k \\ &s = h \Rightarrow P'(x_{k+1}) = d_{k+1} \end{aligned} \tag{28}$$

But these values of the derivatives in the interpolation points are not given, and they must be determined by using a supplementary procedure, that must ensure, besides the continuity of the interpolation function (which results implicitly), also the continuity of order 1 and eventually order 2 derivatives.

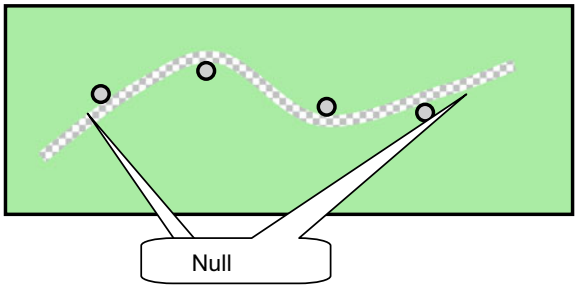


Fig. 5 Drawing a curve using the spline

(iii) *Spline interpolations*

A first variant consists in "*spline*" interpolation. The name of this method comes from an instrument used to draw curves (Fig. 5).

As follows, from the great variety of possibilities, only ***cubic spline*** functions will be treated. The second order derivative of the interpolation polynomial on the considered segment is:

$$P''(x) = \frac{(6h - 12s)\delta_k + (6s - 2h)d_{k+1} + (6\mathrm{s} - 4h)d_k}{h_k^2} \tag{29}$$

The values in the two extremities are:

$$\begin{aligned} x = x_k \Rightarrow s = 0 \Rightarrow P''(x_k^+) &= \frac{6\delta_k - 2d_{k+1} - 4d_k}{h_k} \\ x = x_{k+1} \Rightarrow s = h_k \Rightarrow P''(x_k^-) &= \frac{-6\delta_k + 4d_{k+1} + 2d_k}{h_k} \end{aligned} \tag{30}$$

We denote by $P''(x_k{}^+)$, respectively by $P''(x_{k+1}{}^-)$ the left, respectively the right limit value of the 2nd order derivative in the two extreme points of the interval. In a similar manner, we proceed for the interval (x_{k-1}, x_k). It results:

$$P''(x_k^-) = \frac{-6\delta_{k-1} + 4d_k + 2d_{k-1}}{h_{k-1}} \tag{31}$$

To ensure the continuity of the 2nd order derivative in the point x_k it is necessary that:

$$P''(x_k^+) = P''(x_k^-) \Rightarrow \frac{-6\delta_k + 4d_{k+1} + 2d_k}{h_k} = \frac{-6\delta_{k-1} + 4d_k + 2d_{k-1}}{h_{k-1}} \tag{32}$$

It result the following equations:

$$h_k d_{k-1} + 2(h_k + h_{k+1}) + h_{k-1}d_{k+1} = 3(h_k\delta_{k-1} + h_{k-1}\delta_{\mathrm{k}}); k = 2, \ominus, n-1 \tag{33}$$

To determine the n unknown values d_k, two more equations must be added. These equations can be obtained by imposing supplementary conditions in the two extremities of the definition interval (x_1, respectively x_n). A first approach consists in adopting only one interpolation polynomial on the first interval $x \in [x_1, x_3]$, respectively the last one $x \in [x_{n-2}, x_n]$, method called *not a knot end condition.* Another possible approach, which defines the so-called ***natural spline*** functions, consists of adopting null values for the 2nd order derivative in the extremities of the segment, equivalent to adopting null curves in these extremities (suggested also in the Fig. 5). In this approach:

$$P''(x_k^+) = 0 \Rightarrow 2d_1 + d_2 = 3\delta_1 \tag{34}$$

$$P''(x_n^-) = 0 \Rightarrow d_{n-1} + 2d_n = 3\delta_{n-1} \tag{35}$$

To preserve the symmetry property of the matrix **A**, the two conditions are multiplied with convenient coefficients, obtaining:

$$P''(x_k^+) = 0 \Rightarrow (2d_1 + d_2 = 3\delta_1)(^*h_2) \Rightarrow 2h_2d_1 + h_2d_2 = 3h_2\delta_1 \tag{36}$$

$$P''(x_n^-) = 0 \Rightarrow (d_{n-1} + 2d_n = 3\delta_{n-1})(^*h_{n-2}) \Rightarrow h_{n-1}d_{n-1} + 2h_{n-1}d_n = 3h_{n-2}\delta_{n-1} \tag{37}$$

The following system of equation is obtained:

$$\mathbf{A}\,\mathbf{d} = \mathbf{g} \tag{38}$$

The coefficients matrix **A** is three-diagonal and symmetrical.

3.2 Numerical Differentiation

In principle, the derivative of a function $x(t)$ at moment t_k can be expressed function of the previous values (already known at moment t_k) of the variable, by using relationships of the type:

$$\left.\frac{d}{dt}x(t)\right|_{t_k} = \sum_{i=0}^{p} \alpha_n x(t_{k-i}) \tag{39}$$

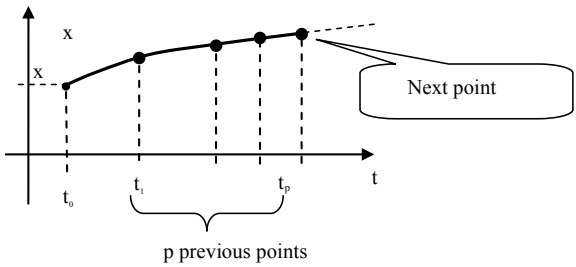


Fig. 6 The approximation of the derivative

The coefficients are determined, in principle, by approximating the previous behaviour of the function by using suitable functions, most often by using polynomials with a convenient degree, as presented in Fig. 6.

For a constant discretization step (h), the previous formulae become:

$$\left.\frac{d}{dt}x(t)\right|_{t_k} = \frac{1}{h}\sum_{i=0}^{p} c_n x(t_{k-i}) = \underbrace{\frac{1}{h}c_0 \underbrace{x(t_k)}_{\text{unknown}} + \underbrace{\frac{1}{h}\sum_{i=1}^{p} c_n \underbrace{x(t_{k-i})}_{\text{known}}}_{\text{explicit method}}}_{\text{implicit method}} \tag{40}$$

These formulae present the following particularities.

- They define a method with «p steps», or of order p, because there are used the already known values of the function in (p) steps of previous computation, to determine the new value $x(t_k)$.
- The methods with 1 step (order 1) present the advantage of an easy start, because they do not require supplementary information other than those provided by the initial conditions.
- The method is «implicit», because the value yet unknown $x(t_k)$ also appears in the right member of the approximation formula.
- The method can be transformed into an explicit one, if the lower limit of the summing index (i) is 1 instead of 0. In this way, the unknown value $x(t_k)$ no longer appear in the approximation formula of the derivative at moment t_k, such that the derivative will be computed function of the previous (known) values of the function.

4 Numerical Methods for Solving Equations and Systems Equations

4.1 Numerical Methods for Solving Equations

Algebraic equations, of order higher than four, and the transcendent ones belong to the nonlinear equations' category that requires for their computation numerical solving methods. Numerical methods are exclusively iterative, and the solutions of theses equations are obtained as limits of some convergent strings. As follows, we briefly present a few of the newest numerical methods [10–12].

(1) *Bisection method or half interval method*

This is one of the simplest method for solving, in the interval (a, b), the nonlinear equation $f(x) = 0$. If $f(a) \cdot f(b) < 0$ then compulsory exists the unique solution x of the equation in the interval (a, b), [3]. The first iteration is computed:

$$x_{m0} \equiv x_0 = \frac{a+b}{2} \tag{41}$$

the half of the interval $[a, b]$. If x_0 is not a solution of the equation, then the verification can be done by comparing

$$|f(x_0)| < \varepsilon \tag{42}$$

and then it is evaluated among the two sub-intervals (a, x_0) and (x_0, b), the one which contains the solution as it is shown in Fig. 7. The evaluation takes place according to the sign of the product of the function' values at the ends of the sub-intervals, i.e.:

$$f(a) \cdot f(x_0) < 0 \tag{43}$$

If the inequality (43) is satisfied, then the solution is searched in the interval (a, x_0), the algorithm continuing by halving the interval:

$$x_{m1} \equiv x_1 = \frac{a+x_0}{2} \tag{44}$$

If the inequality (43) is not satisfied, then the solution is in the interval (x_0, b), and as a consequence the algorithm will continue by halving the interval:

$$x_{m1} \equiv x_1 = \frac{x_0+b}{2} \tag{45}$$

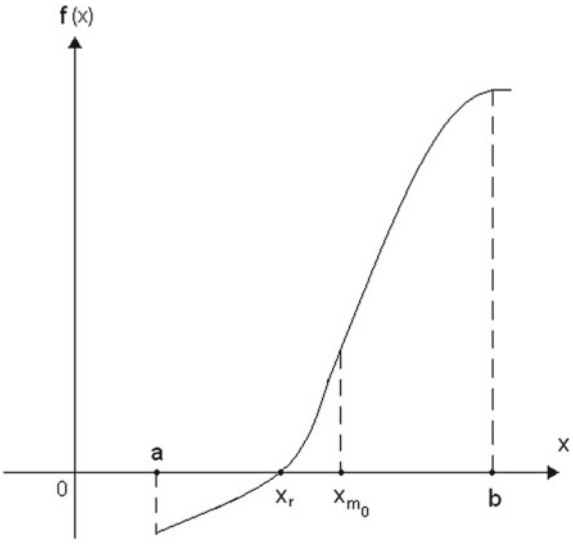


Fig. 7 The first iteration from the half interval method

The series of the successive values of the halves of the sub-intervals: $x_0, x_{m1}, x_{m2}, \ldots$ is convergent to the exact solution x_r. The iterative halving algorithm stops when the length of the sub-intervals obtained in the iteration becomes smaller than the value of the imposed error ε:

$$|b_n - a_n| < \varepsilon \tag{46}$$

The half interval method is easy to use and is surely convergent, but it requires a big number of function evaluations.

(2) *Secant method (proportional parts method).*

This method solve the equation $f(x) = 0$ in interval (a, b) when $f(a) \cdot f(b) < 0$, by dividing the considered interval in two proportional parts with $|f(a)|$ and $|f(b)|$.

So, in a first approximation, it is determined the abscise corresponding to the intersection between the chord which connects the points of coordinates $(0, f(a))$ and $(b, f(b))$ with $0x$-axis:

$$x_1 = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)} \tag{47}$$

The iterative inverse of proportional division of the interval continues, function of the position of the chord with respect to the figure. If the chord is placed to the left of the figure (Fig. 8), i.e. if:

$$f(a) \cdot f(x_1) < 0 \tag{48}$$

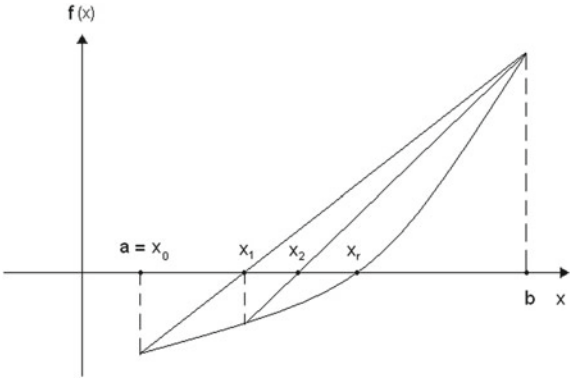


Fig. 8 First iterations of the secant method

then the sub-interval $(a,\ x_1)$ is again divided in proportional parts and the approximate solution for the next step is:

$$x_2 = \frac{a \cdot f(x_1) - x_1 \cdot f(a)}{f(x_1) - f(a)} \tag{49}$$

In this case, the starting point of the approximation is the point $a = x_0$, the point b remains fix and the movement of the successive approximations takes place form left to right. It the chord is to the right of the graphic, i.e.:

$$f(x_1) \cdot f(b) < 0 \tag{50}$$

then the sub-interval $(x_1,\ b)$ is again divided in proportional parts and the approximate solution for the next step is:

$$x_2 = \frac{x_1 \cdot f(b) - b \cdot f(x_1)}{f(b) - f(x_1)} \tag{51}$$

In this case, the starting point of the approximations is the point $b = x_0$, the point a remains fix and the movement of the approximations takes place from the right to the left. Finally, the value x_k is the solution of the given equation, so $f(x_k) < \varepsilon$, where ε is the error-imposed value. The convergence condition (called Fourier condition) $f(x_0) \cdot f''(x_0) < 0$ leads to the optimal choice of the starting point x_0 of iteration.

(3) *Newton method*

This method is based on the linearization procedure, in which there are retained only the first two terms from the function expansion $f(x)$, $f : [a, b] \to \Re$, continuous on $[a, b]$ in Taylorseries around the point x_0. Therefore, the initial $f(x) = 0$ is replaced by the linear equation:

$$f(x_0) - (x - x_0) \cdot f'(x_0) = 0 \tag{52}$$

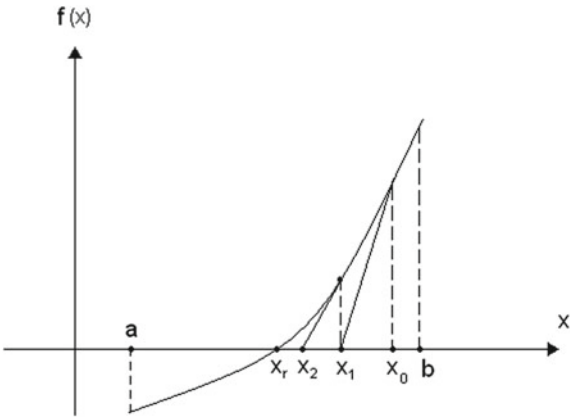


Fig. 9 First iterations in Newton- Raphson method

whose solution is:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{53}$$

From geometrical point of view (Fig. 9), in Newton method, x_1 represents the abscise of the intersection point of tangent to the $f(x)$ graphic and the point of coordinates $(x_0, f(x_0))$ with $0x$-axis. It results the recursive relationship for the approximation of the solution $k + 1$:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{54}$$

Similarly to the above secant method the optimal choice of the initial point x_0 leads to a good convergence of Newton method. In this case the condition is $f(x_0) \cdot f''(x_0) > 0$ which leads to the determination of the order two derivative of the function. More, at each step of the iteration, the first order derivative of the function must be evaluated. This implies a huge computation effort.

4.2 Numerical Methods for Solving the System of Equations

The computation of the algebraic systems is one of the main problems in the numerical methods. Many problems are being reduced, finally, to systems of algebraic equations whose number of equations can be very big, which leads to a significant computational effort [13–16].

(1) *Gauss elimination method*

Gauss method is used to solve the linear equation systems that have non-singular square matrix ($\det\mathbf{T} \neq 0$), using an algorithm of successive elimination of the unknowns. Therefore, if the first equation of the system is multiplied

$$\mathbf{T}\,\mathbf{X} = \mathbf{G} \tag{55}$$

one at a time with the factors $-T_{i1}/T_{11}$, for $i = 2, N$ and then it is added to the equation with number i; so, the unknown x_1 will be eliminated from each equation, obviously excepting the first equation. The following equivalent system is being obtained:

$$\begin{pmatrix} T_{11}X_1 + & T_{12}X_2 & +.. & +T_{1N}X_N \\ & T_{22}^{(1)}X_2 & +.. & +T_{2N}^{(1)}X_N \\ & & & \\ & T_{N2}^{(1)}X_1 & +.. & +T_{NN}^{(1)}X_N \end{pmatrix} = \begin{pmatrix} G_1 \\ G_2^{(1)} \\ \\ G_N^{(1)} \end{pmatrix} \tag{56}$$

where we used the following notations: $\mathbf{T}$ is the matrix of the coefficient and is a matrix of dimension $N*N$, $\mathbf{X}$ is the vector with N elements of the unknowns, $\mathbf{G}$ is the column vector of N elements with free terms, and the new coefficients obtained in this first computation step are indicated by the upper index $^{(1)}$. Also, it is obvious that, in the first equation of the system, the coefficients do not modify. The algorithm continues by multiplying the second equation of the system with $-T_{i2}/T_{22}$, for $i = 3, N$ and then by adding at each iteration i of the equation, the unknown x_2 will be eliminated. The following equivalent system will be obtained:

$$\begin{pmatrix} T_{11}X_1 + & T_{12}X_2 & +.. & +.. & +T_{1N}X_N \\ & T_{22}^{(1)}X_2 & +.. & +.. & +T_{2N}^{(1)}X_N \\ & & T_{23}^{(2)}X_3 & +.. & +T_{2N}^{(2)}X_N \\ & & T_{N3}^{(2)}X_1 & +.. & +T_{NN}^{(2)}X_N \end{pmatrix} = \begin{pmatrix} G_1 \\ G_2^{(1)} \\ \\ G_N^{(2)} \end{pmatrix} \tag{57}$$

The elimination procedure of the unknowns continues until the system becomes:

$$\begin{pmatrix} T_{11}X_1 + & T_{12}X_2 & +.. & +.. & +T_{1N}X_N \\ & T_{22}^{(1)}X_2 & +.. & +.. & +T_{2N}^{(1)}X_N \\ & & T_{33}^{(3)}X_2 & +.. & +T_{3N}^{(3)}X_N \\ & & & & T_{NN}^{(N-1)}X_N \end{pmatrix} = \begin{pmatrix} G_1 \\ G_2^{(1)} \\ \\ G_N^{(N-1)} \end{pmatrix} \tag{58}$$

that represents the conversion of $\mathbf{T}$ to an upper triangular matrix, then the system becomes:

$$\begin{pmatrix} T_{11}+ & T_{12} & +.. & +.. & +T_{1N} \\ & T_{22}^{(1)} & +.. & +.. & +T_{2N}^{(1)} \\ & & T_{33}^{(3)} & +.. & +T_{3N}^{(3)} \\ & & & & T_{NN}^{(N-1)} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \\ X_N \end{pmatrix} = \begin{pmatrix} G_1 \\ G_2^{(1)} \\ \\ G_N^{(N-1)} \end{pmatrix} \tag{59}$$

where the upper indices indicate the step in which the respective element was obtained. This step of transforming the matrix **T** to a triangular shape is called elimination or triangularization, and each of the elements T_{ii}, with $i = 1, N$ from the main diagonal is called pivot. The unknowns are determined after $(N - 1)$ elimination steps, one at a time, starting from the last one:

$$\begin{aligned} X_N &= \frac{G_N^{(N-1)}}{T_{nn}^{(n-1)}} \\ X_{N-1} &= \frac{G_{N-1}^{(N-2)} - T_{N-1}^{N-2} \cdot X_N}{T_{N-1,N-1}^{(N-2)}} \\ X_i &= \frac{G_i^{(i-1)} - \sum_{j=i+1}^{N} T_{i,j}^{(i-1)} \cdot X_j}{T_{ii}^{(i-1)}} \quad ; j = N-2, \ldots, 2 \\ X_1 &= \frac{G_1 - \sum_{j=2}^{N} T_{1,j} \cdot X_j}{T_{11}} \end{aligned} \tag{60}$$

This step of eliminating the unknowns is called retro-substitution. The two steps can be cumulated, if the coefficients matrix is extended with a supplementary column, which contains the free terms of the system (eventually, by adding supplementary columns, one can solve simultaneously more equation systems with same coefficients but various free terms, as in the example presented below).

It is obvious that, along the algorithm, in order to operate, all the pivots must be non-zero. More, to reduce the rounding errors, it is recommended that the value of the pivot, in module, to be as big as possible. To solve these two requirements, the pivoting procedure is used, that is the permutation (change) of the lines (or of the lines and of the columns). The pivoting from a certain computation step does not modify the terms already obtained in the matrix **T**.

(2) *Gauss-Jordan method*

This method represents a modification of Gauss algorithm. It is applied to the extended matrix by adding the free term column, respectively by adding the free terms columns, in case of simultaneous computation or more systems with same coefficients, or of a unity matrix, to determine the inverse of the coefficients matrix [17].

The method has two steps.

- The matrix is brought to the "*staircase*" shape:

$$T_e \Rightarrow T_e^{(1)} = \left(\begin{array}{ccccc|ccc} 1 & x & x. & x & x & x & x & . . x \\ 0 & 1 & x & x & x. & x & x & . . x \\ 0 & 0 & 1 & x & x & x & x & . . x \\ 0 & 0 & 0 & 1 & x & x & x & . . x \\ 0 & 0 & 0 & 0 & 1 & 0 & x & . . x \end{array}\right) \tag{61}$$

All the coefficients from the diagonal of the initial matrix are set to the value 1, and the ones under the main diagonal are cancelled. The procedure is similar to one form Gauss method, the only difference coming from the diagonal coefficients.

- The matrix is brought to the *reduced "staircase"* shape*:* (62)

$$T_e^{(1)} \Rightarrow T_e^{(2)} = \left(\begin{array}{ccccc|ccc} 1 & 0 & 0 & 0 & 0 & y & y & . . y \\ 0 & 1 & 0 & 0 & 0. & y & y & . . y \\ 0 & 0 & 1 & 0 & 0 & y & y & . . y \\ 0 & 0 & 0 & 1 & 0 & y & y & . . y \\ 0 & 0 & 0 & 0 & 1 & 0 & x & . . y \end{array}\right) \tag{62}$$

All coefficients from the initial matrix diagonal are set to the value 1, and the other are cancelled. The procedure is similar to the retro-substitution from the Gauss method. Finally, the extended part of the matrix will contain, depending on the case:

– The solution of the system, for the case when the coefficients matrix has been extended with the free terms vector.
– The inverse coefficients, when the matrix was extended with a unity matrix.

5 Optimization Methods

The totality of the methods and techniques of determining and practical implementation of the best solutions is called *optimization.* The problem of optimization has met several development stages, marked by the emergence of some research areas. The problem of optimization is initiated from antiquity, continuing in the Renaissance period (in which remarkable progress was made in the development of the optimization problem by the appearance of the variational calculus, respectively its extension by the optimal command), culminating with the theory of optimal control of L. S. Pontryagin (1956). L. Euler provides *the necessary conditions* (of the first order), for determining a stationary solution of finding the minimum, leading to the equations called today, *the Euler–Lagrange equations.* A. M. Legendre first dealt with the second variation in 1786. At the basis *of the dynamic programming* developed by Bellman is *the Hamilton–Jacobi equation* (1838) [18–21].

The variational calculus deals mainly with the optimization of the problems having the following standard form:

$$\text{functional minimization } I = \int_a^b L(q, \dot{q}, t)dt$$
$$\text{where, the end conditions } q(a) = \overline{q} \text{ and } q(b) = \hat{q} \tag{63}$$

or the form:

$$\min\ I = \int_a^b L(q(t), u(t), t)dt, \text{ where } q(a) = \overline{q},\ q(b) = \hat{q} \text{ and } \dot{q}(t) = u(t),$$
$$\text{for}\quad a \leq t \leq b \tag{64}$$

The fundamental feature of these problems is that the index minimization (63) replaces the problem of minimization in the space of all curves, L being the *lagrangean of the* problem.

In optimal control problems there are at least two objects that give the problem structure: the dynamics f and the functional I. The optimal solution is obtained by minimizing the functional I. The problem of Bernoulli's minimum time, published in *Acta Eruditorum (1697, Groningen)*, is a real problem of minimum time such as those present, from the theory of optimal control. Bernoulli named the solution as the fastest way to find the *minimum time*. Moreover, this is the first problem that involves a dynamic system, with the explicit requirement of finding the optimal path. The optimal control of a dynamic system involves the synthesis of the solution candidates directly resulting from the first order derivation of the functional cost.

The following types of optimization can be highlighted:

A. *depending on the state of process operation:*

A.1—*stationary*, to which the selected performance criterion is not associated with a dynamic (the method of the least squares, gradient methods, etc.), so the operating state is the stationary one. It could be divided as follows:

A.1.1 *open loop*—when the mathematical model of the process is precisely known, neglecting the perturbations influence,

A.1.2 *closed loop*—when the parameters are varying in time.

A.2—*dynamic—with infinite final time*, it enters into the stabilization problems;

– *with finite final time*, involves the use of either the Hamilton–Jacobi-Bellman formalism governing optimal closed-loop control, or the Euler–Lagrange variational formalism.

Dynamic optimization can be:

A.2.1—*open loop*—appears as a solution of the optimization problem with constraints: temporal, control, evolution along a trajectory.

A.2.2—*closed loop*—is an optimal feedback control type behavior according to the state. This is the most used case because it reduces the sensitivity of the system to the variations of the parameters and to the access regarding the state of the system.

B. *depending on the imposed restrictions:*

- *Problems without constraints* or *free problem* with reference to control. In reality the signals are limited.
- *Problems with constraints* regarding the system situation require the knowledge of the initial situation, and the candidates for the optimal adjustment problems can only be the commands that generate the trajectories that are in the set of permissible states.
- *Problems with free time* are to achieve a set target moving state space at any point of time in the range of interval definition.

Conditions for optimum

The problem of determining and characterizing the extremes begins with the determination of the extremes for monovariable function, multivariable functions, of the functional extremes, until the presentation of the necessary optimal conditions derived from the Pontryagin minimum principle.

Based on some tools, such as Lagrange's *perturbation* (or variation) *method* and *the variational calculation method*, the minimum point is investigated, then based on sufficiently optimal conditions in terms of the Hamilton–Jacobi equation the minimum points will be separated from the maximum points.

Possible candidates for the optimal solution are provided by using the necessary conditions. By using sufficient conditions the optimal solution can be found.

Weierstrass's theorem ensures that a continuous function over a closed interval reaches its maximum and minimum at this interval and/or at the end of the interval.

The extremes of functions

The problem of determining the extremes of a function is to determine the corresponding minimum and/or maximum points. The problems raised in this subchapter are related to the existence of a minimum, its uniqueness, as well as the possibility of investigating a minimum point.

Let's consider a function defined on the set of real numbers (Fig. 10) with values in the set of real numbers. From the graphical representation it is observed that with the increase of the abscissa on the interval $(-\infty, M_1)$, the ordinate values increase, in the interval (M_1, m_1) the ordinate values decrease and in the interval (m_1, M_2), the ordinates increase monotonically. The point x_{M1} is a *local maximum*, because in a conveniently chosen neighborhood, the $f(x)$ values, for any x other than x_{M1}, are smaller than $f(x_{M1})$. The values $f(x)$ taken by the function for the x—axis of a conveniently chosen neighborhood of a minimum point, x_{m1}, are all greater than $f(x_{m1})$. It is said that the function admits at point x_{m1} is a *local minimum*. The mathematical transposition of the two conclusions is carried out as follows.

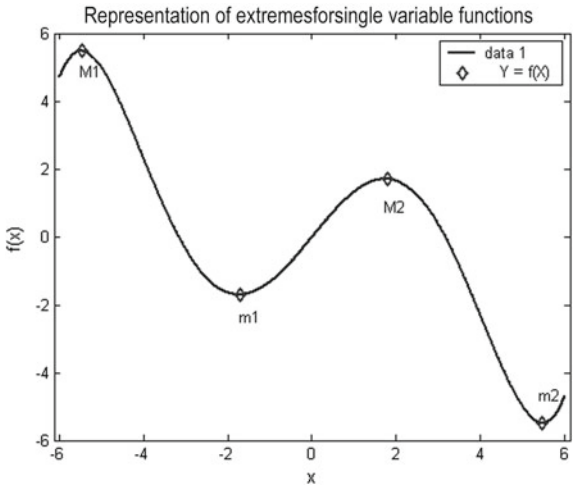


Fig. 10 Extreme points representation for a real function

Local minimum. Global minimum

A point x^* from the real space definition domain $\Re$ is called the *local minimum* of the function f if there is a small number $\varepsilon > 0$, so that for any x located near of x^*, that is:

$$\left|x - x^*\right| < \varepsilon \tag{65}$$

$f(x^*)$ does not exceed the value of $f(x)$: $f(x^*) \leq f(x)$.

In real set $\Re$, x^* is called the *global minimum* (*absolute*) of f if it meets the condition

$$f\left(x^*\right) \leq f(x) \text{ for all } x \text{ from } \Re. \tag{66}$$

Local maximum. Global maximum

A point x^* from the real space definition domain $\Re$ is called the *local maximum* of the function f if there is a small number $\varepsilon > 0$, so that for any x in the vicinity of x^*, that is:

$$\left|x - x^*\right| < \varepsilon \tag{67}$$

$f(x)$ does not exceed the value of $f(x^*)$: $f(x^*) \geq f(x)$.

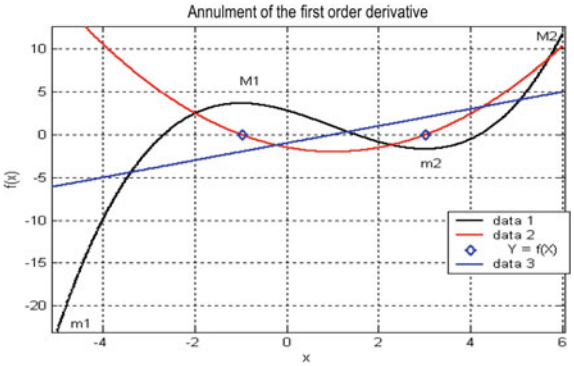


Fig. 11 Vanishing of the first order derivative of a real function

In real set $\Re$, x^* is called the *global minimum* (*absolute*) of f if it meets the condition

$$f(x^*) \geq f(x) \quad \text{for all } x \text{ in } \Re. \tag{68}$$

Necessary and sufficient conditions for determining an extreme point

It is known that the first order derivative of a continuous function after passing through an extreme point changes its sign (Fig. 11). Thus, for the maximum point M_1, the function is increasing monotonically for $x < x_{M1}$, the first derivative being positive to the left of the maximum point: $\lim_{x \to M_{1-}} f'(x) > 0$, and decreasing for $x > x_{M1}$, the first order derivative becoming negative to the right of this point: $\lim_{x \to M_{1+}} f'(x) < 0$.

Analogously, for a minimum point, m_1, the function is decreasing monotonically for $x < x_{m1}$, the first order derivative being negative to the left of the minimum point: $\lim_{x \to m_{1-}} f'(x) < 0$, and descending monotonically $x > x_{\mathrm{m1}}$, the derivative being positive to the right of this point: $\lim_{x \to m_{1+}} f'(x) > 0$.

In conclusion, the derivative of the first order must be canceled when passing through the points of local maximum or minimum local (Fig. 11):

$$\begin{aligned} f'(M_1) &= f'(M_2) = 0 \\ f'(m_1) &= f'(m_2) = 0 \end{aligned} \tag{69}$$

The relation (69), which is only the condition of vanishing of the first order derivative, represents *the required condition* for the presence of an extreme point. The values of the abscissa at the extreme points (Fig. 11) are as follows:

$x_1 =$
-5.47 for M_1
-1.70 for m_1
1.80 for M_2
5.47 for m_2.

The values of the ordinates corresponding to the extreme points are as follows:

$f(x_1) =$
5.47 For M_1
-1.69 for m_1
1.69 For M_2
-5.47 for m_2.

For a real variable function (Fig. 10) the variation of the first order derivative is studied with the help of the derivative of the second order. If the derivative of second order is negative when passing the function through an extreme point, then the extreme point (x_{M1}) is a *maximum* (Fig. 12), $f''(x_{M1}) < 0$.

If the derivative of second order is positive when passing the function through an extreme point, then the extreme point (x_{m2}) is a *minimum* (Fig. 12), $f''(x_{m2}) > 0$.

The above inequalities represent the sign of the second order derivative, and are used to separate the extreme points into *maxima* and *minima* being *the sufficient*

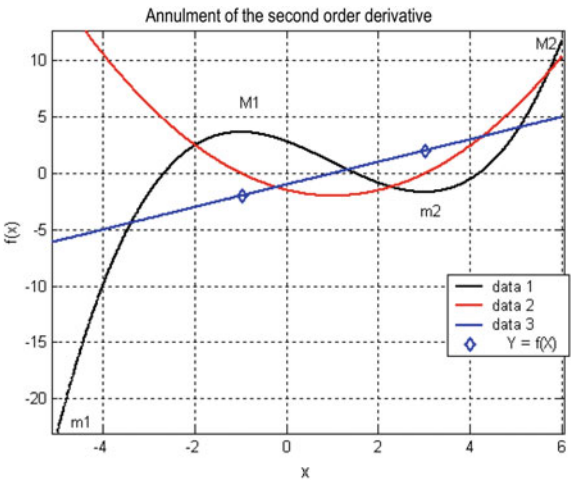


Fig. 12 Vanishing of the second order derivative

conditions, guaranteeing the existence of a maximum or a minimum when they are fulfilled.

In conclusion, for the case where the first order derivative exists in all points of the domain $\Re$:

- the point x_{m2} is a *local minimum* of the $f(x)$, if the necessary and sufficient conditions are fulfilled;
- the point x_{m2} este un *minim global*, is a *global minimum*, if for any other x, x_{m2} is the smallest local minimum: $f(x_{m2}) \leq f(x)$.

Implementation in the Matlab language:

Example 1. Matlab program for determining the minimum of the components of a vector [0 3 2 −1 5 −99 6 5 −1 2 4].

Solution:

Minimal components.

vectort = [0 3 2 −1 5 −99 6 5 −1 2 4];

(a) *minimal vector.*

Enter the Matlab command line:
data_min = min(vectort).
Follows as:
data_min =
−99

(b) If it is desired to extract the minimum of a positive element of the given vector, type the following into Matlab command line:

data_min = min(vectort (vectort > 0)).
Follows as
data_min =
2

(c) If it is desired to extract the minimum of a negative element of the given vector:

data_min = min(vectort (vectort < 0)).
Results
data_min =
−99

Example 2. Computing the minimum of a real function.

Let's consider the function $p(x) = x^2 + 3x - 4$.

a) functieminscript.m function is created
functieminscript.m

```
% To minimize the function:
%%%%%%%%%%%
%       p(x)=x^2+3x-4
%%%%%%%%%%%
%p=[1 3 -4];
function p = functieminscript(x)
p=x.^2+3*x-4
```

b) the built function:
functiemin.m

```
% has the role of both displaying the function and determining the minimum.
% p(x)=x^2+3x-4
p=[1 3 -4];
x=[-5:.1:5];
px=polyval(p,x);
plot(x,px)
xlabel('x')
ylabel('p=f(x)')
title('The minimum of a single variable function ')
%finding  minimum value over an interval  (-5,5)
xvec = fminbnd('functieminscript',-5,5)
% Function value at the minimum point
p=functieminscript(xvec)
xvec =
   -1.5000
p =
   -6.2500
```

Results are shown in Fig. 13.

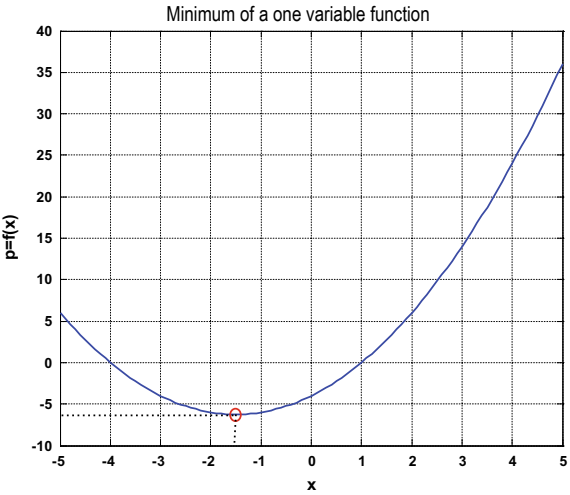


Fig. 13 The minimum of a function

6 Application to Matlabs

Matlab represents the reference software for studying the numerical methods and for solving some practical problems [21–25]. It is available on the classic operating systems (OS) Windows, Linux, Mac OS X. As examples of the numerical methods presented in the previous sections, implemented in Matlab, we will present Newton-method for finding the solutions of the equations and the Gauss method for solving the systems of equations.

(i) *Newton method implemented in Matlab*

```
function x=zeronewton(f, fprim, x, xprev)
% Solving the equation f(x)=0
% NewtonRaphson method
% f(x) = function
% fprim(x)= derivative of f(x)
% x    = initial value
% xprev= initialization
trace=true;
if trace
   figure;
   fplot(f, [x xprev])
   hold;
   plot([x xprev],[0 0],'--')
   disp ('*** Newton-Raphson Method');
   disp ('it---x---------x-xprev-------')
end
 %Algorithm:
k = 0;
while abs(x - xprev) > eps*abs(x)
   xprev = x;
   x = x - f(x)/fprim(x);
      if trace
         disp(sprintf('%3d %15.12f %15.12f', k, x, x-xprev))
         plot([xprev x] , [ f(xprev) 0],':r', [x x], [0 f(x)],':b')
      end
   k = k + 1;
end
if trace plot([x x], [f(x) f(x)],'or'), end
```

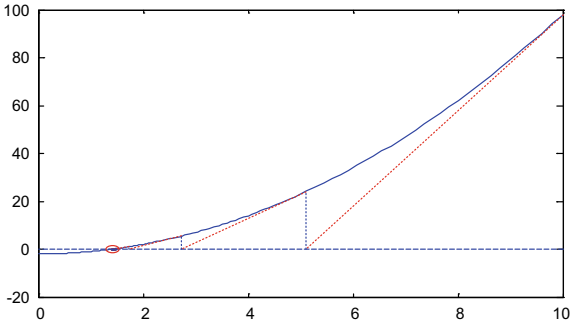


Fig. 14 Solving the equation $x^2 - 2=0$ using the Newton–Raphson method

Computation example: it is considered the numerical computation of the equation $x^2 - 2 = 0$ and the results are presented in Fig. 14.

```
>> x=zeronewton(@(x)x.^2-2,@(x)2.0*x, 10, 0)
***  Newton-Raphson Method
it---x---------------x-xprev-------
  0  5.100000000000 -4.900000000000
  1  2.746078431373 -2.353921568627
  2  1.737194874380 -1.008883556993
  3  1.444238094866 -0.292956779513
  4  1.414525655149 -0.029712439717
  5  1.414213596802 -0.000312058346
  6  1.414213562373 -0.000000034429
  7  1.414213562373 -0.000000000000
```

As a remark, Newton method has a rapid square convergence and for the first step there are necessary two values for the independent variables *x*, respectively the start value *x* and a value *xprev* (in the presented examples, this value was considered to be zero). For the graphical representation, the initial limits [*x, xprev*] were adopted.

(ii) *Gauss method implemented in Matlab*

```
function T=gauss(T)
% transforming the matrix T into an upper triangular form
% Solving the system of equation
%  Simplified algorithm – no pivoting
 [nlin, ncol]=size(T);
 for i=1:nlin-1
    % Next lines passing
    for k=i+1:nlin
       % Treating the remaining elements
       % Cancel the elements from the lines (i+1,nlin)
       %   and from the column (i+1)
       % Tii is the pivot
       cki=T(k,i)/T(i,i);
       % Line(k) = line(k)-line(i)*cki:
       for j=i:ncol
           T(k,j)=T(k,j)-T(i,j)*cki;
       end
    end
end
 % If there are extra lines (free term)
% solve the system.
% The solution will be in the last columns
 if ncol>nlin
    %Inverse passing – retro-substitution
    for j=nlin+1:ncol
       T(nlin,j)=T(nlin,j)/T(nlin,nlin);
       for i=nlin-1:-1:1
         s=0;
         for k=i+1:nlin
             s=s+T(i,k)*T(k,j);
         end
         T(i,j)=(T(i,j)-s)/T(i,i);
       end
    end
end
```

Computation example: solve the system matrix $\mathbf{T}*\mathbf{X} = \mathbf{G}$ where.

```
% The coefficients matrix (3*3:
    >> T=[1 3 5; 3 7 8; 8 1 4]
    T =
        1    3    5
        3    7    8
        8    1    4
    >> G=[ 2; 3; 8 ];
    % The extended matrix with the free terms column:
    >> B=[[T] G]
    B =
        1    3    5    2
        3    7    8    3
        8    1    4    8
    % Gauss Method:
    >> Bg=gauss(B)
    Bg =
       1.0000    3.0000    5.0000    0.7753
            0   -2.0000   -7.0000   -0.5843
            0         0   44.5000    0.5955
    % The solution is on the last column:
    >> X=Bg(:,4)
    X =
       0.7753
      -0.5843
       0.5955
    %Verification:
    >> T*X-G
    ans =
      1.0e-014 *
            0
       0.0888
       0.1776
    % Or :
    >> X=T\G
    x =
       0.7753
      -0.5843
       0.5955
```

7 Conclusion

In this chapter an overview of advanced numerical methods is presented. After errors are defined, the readers are initiated with the principles of approximating functions, numerical methods for solving equations and systems equations and optimization methods. Finally an introduction in Matlab is done to put in evidence the easy-to-use and attractiveness of this popular software. Each method is accompanied by examples to help understanding.

References

1. Stanasila O (1981) Mathematical analysis (in Romanian: Analiza matematica). Didactica si Pedagogica, Bucuresti
2. Wolfram S (1991) Mathematics—a system for doing mathematics by computers, 2nd edn. Addison Wesley
3. Zhang F, Dahlquist G, Bjorck A (2003) Numerical methods. Dover Publication
4. Chasnov JR (2012) Introduction to numerical methods. Lectures Notes for Mathematics. Creative Common
5. Andrei H, Fluerasu C, Badea AC, Caciula I, Stan F (2008) Methodes de calcul numerique en ingenerie electrique. Application en C++ et TurboPascal, Bibliotheca
6. Greenbaum A, Chartier TC (2017) Numerical methods: design, analysis, and computer implementation of algorithms
7. Goldstine H (1977) A history of numerical analysis from the 16th through the 19th century. Springer, New York
8. Quarteroni A, Sacco R, Saleri F (2000) Numerical mathematics. Springer, New York
9. Ueberhuber CW (1997) Numerical computation: Vol. 1: Methods, software, and analysis, Vol. 2: Methods, software, and analysis, Springer, New York
10. Dey S, Gupta S (2017) Numerical methods. McGrew Hill
11. Jain MK, Iyengar SRK, Jain S (2017) Numerical methods for scientific and engineering computation. New Age International Publishers
12. Dautrey R, Lions J-L (1988) Analyse Mathematique et Calcul Numerique pour les Sciences et les Techniques, Masson
13. Saad Y (1996) Iterative methods for sparse linear systems. PWS
14. Stoer J, Bulirsch R (1993) Introduction to numerical analysis. Texts in Applied Mathematics, vol 12. Springer
15. Brugnano L, Iavernaro F (2018) Line integral solution of differential problems. Axioms 7(2):36
16. Golub G, Van Loan C (1996) Matrix computations, 3rd edn. Johns Hopkins University Press
17. Simoncini V (2016) Computational methods for linear matrix equations. SIAM Rev 58:377–441
18. Atkinson K, Han W (2005) Theoretical numerical analysis: a functional analysis framework, 2nd edn. Springer, New York
19. Higham N (1996) Accuracy and stability of numerical algorithms. SIAM Pub, Philadelphia
20. Phohomsiri P (2005) On the extrema of linear least squares problems. J Optim Theory Appl 127(3):665–669
21. Glowinski R (1984) Numerical methods for nonlinear variational problems. Springer
22. Moler C (2004) Numerical computing with MATLAB. SIAM Pub, Philadelphia
23. Brugnano L, Iavernaro F (2019) Advanced numerical methods in applied sciences. Axioms 8(1):16
24. Dahlquist G, Björk Å (2008) Numerical methods in scientific computing, vol I. SIAM, Philadelphia, PA, USA
25. https://www.mathworks.com/products/matlab.html

Analysis of Partial Differential Equations in Time Dependent Problems Using Finite Difference Methods and the Applications on Electrical Engineering

Hidir Duzkaya, Suleyman Sungur Tezcan, and M. Cengiz Taplamacioglu

Abstract Finite difference methods are used as a numerical method in the time-dependent and time-independent solution of partial differential equations commonly encountered in many engineering problems. In this chapter the theoretical basics and practical applications of these methods and limits their scope in terms of electrical engineer applications are examined. This numerical method, which are commonly used to solve problems of electromagnetic fields that cannot be solved by analytical methods are described in detail. Among these methods, the Finite Difference Time Domain (FDTD) method, which is widely used in the calculation of the electric and magnetic fields in electrical engineering applications is concentrated and the results, limitations and alternatives of this method for different applications are examined. Using this numerical method, nonlinear material and structural characteristics in engineering applications can be examined depending on time. Robust and accurate analysis results can be obtained by using this method, which can also be integrated with developed models and software.

Keywords Partial differential equations (PDEs) · Finite difference method · Time dependent problems · Electrical engineering applications

H. Duzkaya · S. S. Tezcan · M. C. Taplamacioglu (✉)
Electrical and Electronics Engineering Department, Engineering Faculty, Gazi University, Ankara, Turkey
e-mail: taplam@gazi.edu.tr

H. Duzkaya
e-mail: hduzkaya@gazi.edu.tr

S. S. Tezcan
e-mail: stezcan@gazi.edu.tr

H. Duzkaya
Ampre Laboratory CNRS UMR, Ecole Centrale de Lyon, Ecully, France

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_2

Acronyms/Abbreviations

CB	Circuit Breaker
FDTD	Finite Difference Time Domain
GIS	Gas Insulated Substation
HV	High Voltage
PD	Partial Discharge
PT	Power Transformer
UHF	Ultra High Frequency

1 Introduction

The Finite Difference Time Domain (FDTD) method is one of the most widely used calculation techniques in the analysis of electromagnetic phenomena since the early 1990s. Although the first application in the use of this method for electromagnetic wave problems dates back to 1966 [1], its prevalence has increased with the development of computer technologies in the solution of numerical calculations [2]. This method used in the solution of partial differential equations is based on the discretizing of Maxwell's curl equations in time and space. At this stage, derivatives are converged using finite difference equations.

The FDTD method algorithm is very popular for its simple implementation and robust and accurate results in the analysis of electromagnetic fields [3]. Analytical methods, that can be used in simple engineering applications, cannot calculate transient and permanent state responses in complex structures with sufficient accuracy in terms of electrical and magnetic fields [4].

In the application stage of this method, first, the boundaries of the region to be analyzed are determined. These borders should be wide enough to cover all of the objects examined. In cases where suitable boundary conditions cannot be determined, calculation can be made by determining artificial boundaries where the calculation region is extended to infinity [5]. This region, whose boundaries are defined, is divided into cells according to the step intervals in the space and time. With finite difference equations solved due to these time and space variables, electric and magnetic field quantities can be calculated for a sufficient number of points in the solution region [6]. In this calculation stage, dielectric and magnetic material parameters of the design in the region where the solution set is located should be defined for each discrete region [5].

FDTD, which is a very simple and efficient alternative to solve Maxwell's equations, can produce solutions in the analysis of many electrical engineering problems [7]. Since spatially and temporally discretization is used, it allows the modeling of three-dimensional inhomogeneous materials, analysis of designs involving planar and non-planar volumes, containing multiple dielectric planes and ground layer, and examination of non-ideal conductors and insulators [3]. This method, in which

designs containing passive loads or active elements can also be modeled by adding them to the initial equation, is used in simulation of many engineering applications such as antennas, high voltage (HV) insulation systems, partial discharge (PD) imaging techniques and grounding systems [2, 7, 8].

In addition to these application areas, high frequency responses frequently encountered in high voltage and power system equipment could also be examined for some modifications made in this model [2, 8]. The propagation of electromagnetic wave in transient and non-transient modes can be examined in a wide frequency range. The FDTD method is widely used for transient analysis [9].

In engineering applications, the experimental measurement of electric and magnetic fields, which can be defined by many variables such as the size of the design, the variety of the material, and the surge caused by switching, is not always possible due to its complex nature and economic constraints [5]. In order to overcome these limitations and to understand the electromagnetic behavior of engineering designs, finite difference methods are used.

This chapter primarily examines the finite difference methods used in the solution of engineering problems defined by differential equations in a conceptual framework. Following this section, the application examples and results of the FDTD method, which stands out with its widespread use in calculating electrical and magnetic fields among these methods, are discussed. Researches on the limitations and alternative modifications of the method are also examined in this section. In the final part of the chapter, the advantages and disadvantages of FDTD, the basic modifications proposed for eliminating these disadvantages and alternative application areas are explained.

2 Finite Difference Methods for Time-Dependent Problems

2.1 Basic Concepts

A general initial problem for linear partial differential equations:

$$\begin{aligned} u_t(x,t) &= P\left(x,t,\frac{\partial}{\partial x}\right)u(x,t) \\ u_t(x,0) &= f(x) \end{aligned} \tag{1}$$

where x is a vector of s components: $x = (x_1, \ldots, x_s)$, u is a vector of p components: $u(x,t) = \big(u_1(x,t), \ldots, u_p(x,t)\big)$ and P is a polynomial $\frac{\partial}{\partial x}$.

For computational convenience, the domain of the solution.

$u(x,t)$ is restricted to a bounded region. On this bounded region, a grid of points is constructed by discretizing both space and time. Step sizes are Δt and Δx_i and the grid points are

$$t_n = n\Delta t$$
$$x_{ji} = j_i \Delta x_i \cdots i = 0, \ldots, N_i \tag{2}$$

Consider a two dimensional problem:

$$u_t = u_x + u_y \tag{3}$$

where $u(x, y, t)$ is a real valued function. Δx, Δy and Δt is positive and fixed quantities. Finite difference scheme is

$$U_{i,j}^{n+1} = 0.25\big(U_{i+1,j+1}^n + U_{i-1,j+1}^n + U_{i+1,j-1}^n + U_{i-1,j-1}^n\big) + \frac{\Delta t}{2\Delta x}\big(U_{i+1,j}^n - U_{i-1,j}^n\big) + \frac{\Delta t}{2\Delta y}\big(U_{i,j+1}^n - U_{i,j-1}^n\big) \tag{4}$$

The shift operators are E_1 and E_2 so

$$U_{i,j}^{n+1} = \left(0.25\big(E_1 + E_1^{-1}\big)\big(E_2 + E_2^{-1}\big) + \frac{\Delta t}{2\Delta x}\big(E_1 + E_1^{-1}\big) + \frac{\Delta t}{2\Delta y}\big(E_2 + E_2^{-1}\big)\right) U_{i,j}^n \tag{5}$$

2.2 Properties of Finite Difference Schemes

$u(x, t)$ is the initial value problem, $S(t, t_0)$ is the solution operator and the function u is

$$u(x, t) = S(t, t_0)u(x, t_0) \tag{6}$$

thus in particular

$$u(x, (n+1)\Delta t) = S((n+1)\Delta t, n\Delta t)u(x, n\Delta t) \tag{7}$$

If the problem is autonomous, the operator P in Eq. 1 is independent of time, S is a function of the elapsed $(t - t_0)$.

Scheme 1

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{2\Delta x}\big(U_{j+1}^n - U_{j-1}^n\big) \tag{8}$$

This scheme is useless since it will never be stable. To investigate its accuracy:

$$u_j^n = u(j\Delta x, n\Delta t) \tag{9}$$

The scheme can be rewritten in the form $u_t = u_x$

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} \tag{10}$$

Local truncation error is

$$\begin{aligned}\tau_j^n &= \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} \\ &= u_t\left(x_j, t_n\right) + O(\Delta t) - u_x\left(x_j, t_n\right) + O\left(\Delta x^2\right)\end{aligned} \tag{11}$$

This scheme is accurate of second order in space and first order in time.

Scheme 2: Lax-Friedrichs Scheme

$$U_j^{n+1} = \frac{1}{2}\left(U_{j+1}^n + U_{j-1}^n\right) + \frac{\Delta t}{2\Delta x}\left(U_{j+1}^n - U_{j-1}^n\right) \tag{12}$$

This scheme is a first order accurate scheme. This scheme describes as (FTCS) forward in time and centered in space [10].

The Lax-Friedrichs scheme has two-degree precision along space and one-degree precision along time [11].

Scheme 3: Upwind Scheme

Consider one sided difference for the spatial derivative:

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x}\left(U_{j+1}^n - U_j^n\right) \tag{13}$$

This is a first order accurate scheme. The upwind differencing scheme is conservative [12].

Scheme 4: Downwind Scheme

Consider the one-sided difference for the spatial derivative:

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x}\left(U_j^n - U_{j-1}^n\right) \tag{14}$$

This is a first order accurate scheme. However, this scheme is also useless. The domain of dependence is not included in the scheme stencil therefore such a scheme is unstable [13].

Scheme 5: Leapfrog Scheme:

If the center difference is used for both time and spatial derivatives,

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{\Delta x}\left(U_{j+1}^n - U_{j-1}^n\right) \tag{15}$$

To find its accuracy, it is rewritten as

$$\frac{U_j^{n+1} - U_j^{n-1}}{2\Delta t} = \frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} \tag{16}$$

$$\tau_j^n = \frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} - \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = u_t + O\left(\Delta t^2\right) - u_x + O\left(\Delta x^2\right) \tag{17}$$

This is a second order accurate scheme. The leapfrog method has good stability when solving partial differential equations with oscillatory solutions [14].

Scheme 6: Lax-Wendroff Scheme

It is based on the Taylor series expansion $u(x, t)$ given by

$$u(x, t + \Delta t) = u(x, t) + \Delta t\, u_t(x, t) + \frac{1}{2}\Delta t^2 u_{tt}(x, t) + O\left(\Delta t^3\right) \tag{18}$$

using $u_t = u_x$ reduces to

$$u(x, t + \Delta t) = u(x, t) + \Delta t\, u_x(x, t) + \frac{1}{2}\Delta t^2 u_{xx}(x, t) + O\left(\Delta t^3\right) \tag{19}$$

Using the centered difference, a scheme with order accuracy in both time and space is obtained by

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{2\Delta x}\left(U_{j+1}^n - U_{j-1}^n\right) + \frac{\Delta t^2}{2\Delta x^2}\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right) \tag{20}$$

The Lax-Wendroff scheme has two-degree precision along both space and time. The Lax-Wendroff scheme gives more accurate solution than that of Lax- Friedrich scheme since the Lax-Wendroff scheme has two-degree precision along time, while the Lax-Friedrichs scheme has one-degree precision along time.

The Lax-Wendroff scheme needs more computational time than that of Lax-Friedrich scheme since the Lax-Wendroff scheme need to calculate derivatives up to 4th order, while the Lax-Friedrichs scheme need to calculate derivatives up to 2nd order [11].

Scheme 7: Crank-Nicolson Scheme

This is a second order accurate implicit scheme

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{2\Delta x}\left(U_{j+1}^{n+1} - U_{j-1}^{n+1} + U_{j+1}^n - U_{j-1}^n\right) \tag{21}$$

The Crank-Nicolson method is implicit scheme with second-order accuracy in both time and space. This method is an unconditionally stable [15].

2.3 Von Neumann Stability

Stability of the scheme $V^{n+1} = C(\Delta t)V^n$ can be written in terms of the amplification matrix $G(\Delta t, k)$ as the following condition: $t > 0$

$$\left\| G(\Delta t, k)^n \right\| \leq K e^{\alpha t} \tag{22}$$

The condition must be satisfied for all multi-index k in order to establish stability of the scheme.

The Von Neumann Condition

The amplification matrix of a stable scheme satisfies the condition

$$p[G(\Delta t, k)] \leq e^{\gamma \Delta t} = 1 + O(\Delta t) \tag{23}$$

where $p[G(\Delta t, k)]$ denotes the spectral radius (largest magnitude of eigenvalues) of the matrix $G(\Delta t, k)$

The Von Neumann stability condition is necessary but not sufficient for stability. In most practical applications, turns out to be easily checked whether this condition holds or not [16].

2.4 The Leapfrog Scheme

2.4.1 The One Way Wave Equation

The one-way wave equation shows significant computational efficiency for a range of transmitted wave three-dimensional global, exploration and engineering scale applications [17]. The leapfrog scheme is

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{\Delta x}\left(U_{j+1}^n - U_{j-1}^n\right) \tag{24}$$

The periodic conditions imposed through the usual periodicity requirement, $U_{-1}^n = U_{N-1}^n$, $U_N^n = U_0^n$

The vector can be defined as

$$V_j^n = \begin{pmatrix} U_j^n \\ U_{j-1}^n \end{pmatrix} \tag{25}$$

$$V^{n+1} = C(\Delta t)V^n \tag{26}$$

and $\lambda = \frac{\Delta t}{\Delta x}$ then

$$V_j^{n+1} = \begin{pmatrix} \lambda\left(E - E^{-1}\right) & 1 \\ 1 & 0 \end{pmatrix} V_j^n \tag{27}$$

where E and E^{-1} are the shift operations. $V_j^n = \widehat{V}_k^n e^{ikj\Delta x}$ is the discrete fourier transform of V^n.

$$\widehat{V}_k^{n+1} e^{ikj\Delta x} = \begin{pmatrix} \lambda\left(E + E^{-1}\right) & 1 \\ 1 & 0 \end{pmatrix} \widehat{V}_k^n e^{ikj\Delta x} \tag{28}$$

and $x_j = j\Delta x$ so

$$E\hat{e}^{ikj\Delta x} V_k^n = e^{ik\Delta x} e^{ikj\Delta x} \widehat{V}_k^n \tag{29}$$

$$E\hat{e}^{ikj\Delta x} V_k^n = e^{-ik\Delta x} e^{ikj\Delta x} \widehat{V}_k^n \tag{30}$$

thus

$$\begin{aligned} \widehat{V}_k^{n+1} &= e^{-ikj\Delta x} \begin{pmatrix} \lambda\left(E - E^{-1}\right) & 1 \\ 1 & 0 \end{pmatrix} e^{ikj\Delta x} \widehat{V}_k^n \\ &= \begin{pmatrix} 2i\lambda \sin(k\Delta x) & 1 \\ 1 & 0 \end{pmatrix} \widehat{V}_k^n \end{aligned} \tag{31}$$

The explicit expression for the amplification matrix is

$$G(\Delta x, k) = \begin{pmatrix} 2i\lambda \sin(k\Delta x) & 1 \\ 1 & 0 \end{pmatrix} \tag{32}$$

The variable $\xi = k\Delta x$ restricted to $0 \leq \xi \leq 2\pi$. The eigenvalues of the amplification matrix $G(\Delta x, k)$ is

$$\mu_1(\xi) = i\lambda \sin(\xi) + \sqrt{1 - \lambda^2 \sin^2(\xi)} \tag{33}$$

Case 1. If $\lambda^2 > 1$, then for those values of k such that $\xi = k\Delta x = \frac{\pi}{2}$

$$\mu_1\left(\pi / 2\right) = i\left(\lambda + \sqrt{\lambda^2 - 1}\right) \tag{34}$$

so $|\mu_1(\pi/2)| > 1$ yielding that the Von Neumann stability condition is not satisfied by the amplification matrix. The leapfrog scheme is unstable when $\lambda > 1$.

Case 2. If $\lambda^2 \leq 1$, then

$$|\mu_1(\xi)|^2 = \lambda^2 \sin^2(\xi) + 1 - \lambda^2 \sin^2(\xi) = 1 \tag{35}$$

Then $p[G] = 1$ and the Von Neumann condition is satisfied. Nonetheless, this does not imply that the scheme is stable for $\lambda \leq 1$, and it is unstable for $\lambda = 1$.

The leapfrog scheme for $u_t = u_x$ is stable for $\lambda < 1$.

2.4.2 The Two Way Wave Equation

Comparison of migration results for one-way and two-way wave-equation migration shows that the two-way wave equation provides superior results [18]. The leapfrog method (second order-centered difference for time and space derivatives) for the two-way wave equation $u_{tt} = u_{xx}$ is

$$\frac{U_j^{n+1} - 2U_j^n + U_j^{n-1}}{\Delta t^2} = \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2} \tag{36}$$

The simplified 1D Maxwell's equations can be written as, $E_t = H_x$, $H_t = E_x$ which is equivalent to $E_{tt} = E_{xx}$ or $H_{tt} = H_{xx}$

The FDTD method (second order centered difference for time and space derivatives) is defined on staggered grid for H:

$$\frac{E_j^{n+1} - E_j^n}{\Delta t} = \frac{H_{j+\frac{1}{2}}^{n+\frac{1}{2}} - H_{j-\frac{1}{2}}^{n+\frac{1}{2}}}{\Delta x} \tag{37}$$

$$\frac{H_{j+\frac{1}{2}}^{n+\frac{1}{2}} - H_{j+\frac{1}{2}}^{n-\frac{1}{2}}}{\Delta t} = \frac{E_{j+1}^n - E_j^n}{\Delta x} \tag{38}$$

$\lambda = \Delta t/\Delta x$ so (36) can be written as

$$U_j^{n+1} = 2U_j^n + \lambda^2\left(E - 2 + E^{-1}\right)U_j^n - U_j^{n-1} \tag{39}$$

where E is the shift operator.

$$V_j^n = \begin{pmatrix} U_j^n \\ U_{j-1}^n \end{pmatrix} \tag{40}$$

$$V_j^{n+1} = \begin{pmatrix} 2 + \lambda^2\left(E - 2 + E^{-1}\right) & -1 \\ 1 & 0 \end{pmatrix} V_j^n \tag{41}$$

$V_j^n = \widehat{V}_k^n e^{ikj\Delta x}$ so

$$\widehat{V}_k^{n+1} = \begin{pmatrix} 2 + \lambda^2\left(e^{ikj\Delta x} - 2 + e^{-ikj\Delta x}\right) & -1 \\ 1 & 0 \end{pmatrix} \widehat{V}_k^n \tag{42}$$

Thus

$$G = \begin{pmatrix} 2 + \lambda^2(2\cos(\xi) - 2) & -1 \\ 1 & 0 \end{pmatrix} \tag{43}$$

The eigenvalues of G are $\mu_1 = a + \sqrt{a^2 - 1}$ and $\mu_2 = a - \sqrt{a^2 - 1}$ with $a = 1 + \lambda^2(\cos(\xi) - 1)$.

If $\lambda > 1$ so $\cos(\xi_0) < 1 - \frac{2}{\lambda^2}$. Then $a(\xi_0) < -1$ and $|\mu_2(\xi_0)| = \left|a - \sqrt{a^2 - 1}\right| > 1$. The Neumann stability is violated thus not stable.

If $\lambda \leq 1$, then $a^2 - 1 \leq 0$ thus $\mu_1 = a + i\sqrt{1 - a^2}$ and $\mu_2 = a - i\sqrt{1 - a^2}$. So $|\mu_i| = 1$ and the Von Neumann stability is satisfied. On the other hand, G is not a normal matrix and $\|G\| > 1$.

2.4.3 Convergence for the Two Way Wave Equation

Replace U_j^n by $u\left(x_j, t^n\right)$ in Eq. (36), the residue is the local truncation error

$$\tau^n = O\left(\Delta t^2\right) + O\left(\Delta x^2\right) \tag{44}$$

Second, replace U_j^n by $u\left(x_j, t^n\right)$ in Eq. (41), the residue is

$$\Delta t^2 \tau^n = \Delta t^2\left[O\left(\Delta t^2\right) + O\left(\Delta x^2\right)\right] \tag{45}$$

Let $V^{n+1} = C(\Delta t)V^n$ denote the leapfrog scheme. Suppose

$$V^{n+1} = \begin{pmatrix} U_0^{n+1} \\ U_0^n \\ U_1^{n+1} \\ U_1^n \\ \vdots \\ U_{N-1}^{n+1} \\ U_{N-1}^n \end{pmatrix} \tag{46}$$

$Q_{\Delta x}$ is the sampling operator at the spatial grid points and two time steps.

$$Q_{\Delta x}u(x,t) = \begin{pmatrix} u(x_0,t) \\ u(x_0,t-\Delta t) \\ u(x_1,t) \\ u(x_1,t-\Delta t) \\ \vdots \\ u(x_{N-1},t) \\ u(x_{N-1},t-\Delta t) \end{pmatrix} \tag{47}$$

2.5 Dissipative Schemes

A finite difference scheme $V^{n+1} = C(\Delta t)V^n$ is called dissipative of order 2τ if the amplification matrix satisfies

$$\rho[G(\Delta t,k)] \leq 1 - \delta|\xi|^{2\tau} \tag{48}$$

where $\xi = k\Delta x$ for all Δt, k and $\delta > 0$ is independent of k and Δt.

2.6 Difference Schemes for Hyperbolic Systems in One Dimension

$$u(x,t) = \big(u_1(x,t),\ldots,u_p(x,t)\big)^T \tag{49}$$

$$u_t(x,t) = \frac{\partial F(u(x,t))}{\partial x} \tag{50}$$

$F(u)$ is a function $F\big(u_1,\ldots,u_p\big) = \big(F_1\big(u_1,\ldots,u_p\big),\ldots,F_p\big(u_1,\ldots,u_p\big)\big)^T$

$$\frac{\partial F(u(x,t))}{\partial x} = \frac{\partial F(u)}{\partial u}\frac{\partial u(x,t)}{\partial x} \tag{51}$$

where $\frac{\partial F(u)}{\partial u}$ denotes the gradient matrix $A(u)$ with components $a_{ij}(u) = \frac{\partial F_i(u)}{\partial u_j}$ so that the nonlinear system can be written in the form

$$u_t = A(u)u_x \tag{52}$$

The above nonlinear equation is called weakly, strongly, symmetric or strictly hyperbolic if for every u_0 fixed, the corresponding linearized system:

$$u_t = A(u_0)u_x$$

is weakly, strongly, symmetric or strictly hyperbolic, respectively.

The Lax equivalence theorem states basically that an accurate scheme is stable if and only if it converges, provided that the problem is strongly well posed. Weak well posedness may give rise to instabilities.

2.6.1 First Order Schemes

Consider Friedrich's scheme:

$$U_j^{n+1} = \frac{1}{2}\left(U_{j+1}^n + U_{j-1}^n\right) + \frac{\Delta t}{2\Delta x}\left(F_{j+1}^n - F_{j-1}^n\right) \tag{53}$$

where $F_{j+1}^n = F\left(U_{j+1}^n\right)$. This scheme is based on first order approximation of the derivatives using Taylor expansion, and it can be easily shown that this scheme is first order accurate. Linearizing the function $F(u)$ around some arbitrary value to u_0, $A(u)$ is replaced by a constant matrix A, so that the linearized problem is equivalent to the original problem with $F(u) = Au$. Substituting in the Freidrich's scheme, the linearized form is obtained.

$$U_j^{n+1} = \frac{1}{2}\left(U_{j+1}^n + U_{j-1}^n\right) + \frac{\Delta t}{2\Delta x}A\left(U_{j+1}^n + U_{j-1}^n\right) \tag{54}$$

The corresponding amplification matrix is given by

$$G(\xi) = I\cos(\xi) + i\lambda\sin(\xi) \tag{55}$$

where $\xi = k\Delta x$, and I is the $p \times p$ identity matrix. If the original problem is strongly or strictly hyperbolic, then it follows that the matrix $A = A(u_0)$ is diagonalizable, there exist a matrix T

$$T^{-1}AT = \begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_p \end{pmatrix} \tag{56}$$

where $a_1, \ldots, a_p$ are the real eigenvalues of A. Therefore:

$$T^{-1}G(\xi)T = I\cos(\xi) + i\lambda\begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_p \end{pmatrix}\sin(\xi) \tag{57}$$

and the eigenvalues are

$$\mu_k(\xi) = \cos(\xi) + i\lambda a_k \sin(\xi) \tag{58}$$

which implies that

$$\begin{aligned} |u_k(\xi)|^2 &= \cos^2(\xi) + i\lambda^2 a_k^2 \sin^2(\xi) \\ &= 1 - \left(1 - \lambda^2 a_k^2\right) \sin^2(\xi) \end{aligned} \tag{59}$$

Therefore, if $\rho(A) = \max_k |a_k|$ satisfies the inequality $\frac{\Delta t}{\Delta x}\rho(A) \leq 1$ then Von Neumann stability condition will hold and $|u_k(\xi)| \leq 1$ for k and ξ. It is an exercise to prove under strict inequality of Von Neumann condition, the scheme is dissipative of order 2.

Upwind schemes are motivated by the scalar equation $u_t = au_x$ when $p = 1$. If $a > 0$ the characteristics are straight lines moving to the left, and the scheme constructed in order to "follow" the physical characteristics is:

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} a\left(U_{j+1}^n - U_j^n\right), \quad a > 0 \tag{60}$$

And the scheme is accurate and stable for $0 < a\lambda \leq 1$ for $\lambda = \frac{\Delta t}{\Delta x}$. On the other hand, if $a < 0$, then the characteristics point to the right and

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} a\left(U_j^n - U_{j-1}^n\right), \quad a < 0 \tag{61}$$

In this case, stability follows from the condition $-1 \leq \lambda a < 0$.

2.6.2 Second Order Schemes

A scheme for approximating the solution of $u_t = A(u)u_x$ is called a Lax-Wendroff scheme if under the assumption $A(u) = A$ (or $F(u) = Au$ is linear) the scheme reduces to

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{2\Delta x} A\left(U_{j+1}^n - U_{j-1}^n\right) + \frac{1}{2}\left(\frac{\Delta t}{\Delta x} A\right)^2 \left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right) \tag{62}$$

The above scheme is actually the only second order scheme for the linear problem.

Lax-Wendroff schemes arise from the idea of replacing time derivatives by space derivatives, using the equation $u_t = F(u)$ and approximating the later by finite differences. Using a Taylor expansion for u

$$u(x, t + \Delta t) = u(x, t) + \Delta t u_t(x, t) + \frac{\Delta^2}{2} u_{tt}(x, t) + O\left(\Delta t^3\right) \tag{63}$$

$u_t(x, t) = F(u(x, t))$ in the linear case where $F(u) = Au$

$$u_t(x,t) = Au_x(x,t) \tag{64}$$

$$u_{tt}(x,t) = A^2 u_{xx}(x,t) \tag{65}$$

The amplification matrix of the linear form of the Lax-Wendroff scheme is

$$G(\xi) = I + i\lambda A\sin(\xi) + \lambda^2 A^2(\cos(\xi) - 1) \tag{66}$$

$\xi = k\Delta t$, $\lambda = \frac{\Delta t}{\Delta x}$ and $\eta = \sin\left(\frac{\xi}{2}\right)$,

$$G(\xi) = I + 2i\lambda A\eta\sqrt{1-\eta^2} - 2\lambda^2 A^2\eta^2 \tag{67}$$

Any eigenvalue $\mu(\eta)$ of the amplification matrix

$$\mu(\eta) = 1 + 2i\lambda\mu(A)\eta\sqrt{1-\eta^2} - 2\lambda^2 A^2\eta^2 \tag{68}$$

The eigenvalues $\mu(\eta)$ of the amplification matrix

$$|\mu(\eta)|^2 = 1 - \lambda^2\mu(A)^2\eta^4\left(1 - \lambda^2\mu(A)^2\right) \tag{69}$$

which holds for every eigenvalue of $G(\xi)$. The spectral radius of $G(\xi)$ is defined as the maximum value of $\mu(\eta)$. μ_* is the eigenvalue of A which maximizes the above expression $|\mu(\eta)|$

$$|\rho(G)|^2 = 1 - \lambda^2\mu_*^2\eta^4\left(1 - \lambda^2\mu_*^2\right) \tag{70}$$

Von Neumann condition will be satisfied if

$$\lambda_p(A) \leq 1 \tag{71}$$

which implies $\lambda_\mu(A) \leq 1$ for all eigenvalues of A. Furthermore, if $\lambda\mu_* < 1$, then the scheme given by (Eq. 62) is dissipative of order 4.

For the nonlinear case,

$$u_{tt} = [F(u)]_{xt} = [F(u)_t]_x = [A(u)u_t]_x = [A(u)F(u)_x]_x \tag{72}$$

Substituting $u_t = F(u)_x$ and using Taylor expansion,

$$u(x,t+\Delta t) = u(x,t) + \Delta t F(u)_x + \frac{\Delta t^2}{2}[A(u)F(u)_x]_x + O\left(\Delta t^3\right) \tag{73}$$

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{2\Delta x}\left(F_{j+1}^n - F_{j-1}^n\right)$$

$$+\frac{1}{2}\left(\frac{\Delta t}{\Delta x}\right)^2\left(A^n_{j+\frac{1}{2}}\left(F^n_{j+1}-F^n_j\right)-A^n_{j-\frac{1}{2}}\left(F^n_j-F^n_{j-1}\right)\right) \tag{74}$$

$F^n_j = F\left(U^n_j\right)$ so

$$A^n_{j+\frac{1}{2}} = A\left(\frac{U^n_{j+1}+U^n_j}{2}\right) \tag{75}$$

Scheme Eq. 74 becomes rather inefficient in practical applications due to the many computations involved at each time step iteration in order to evaluate A and F. A modification of this scheme which is very popular considers approximating derivatives at "half stages" of the iteration

$$u(x,t+\Delta t) = u(x,t)+\Delta t u_t\left(x,t+\frac{1}{2}\Delta t\right)+O\left(\Delta t^2\right) \tag{76}$$

and it is known as the MacCormack scheme. Each iteration has two steps corresponding to first order approximations of the solution at half steps.

The scheme is given by:

$$U^*_j = U^n_j+\frac{\Delta t}{\Delta x}\left(F^n_{j+1}-F^n_j\right) \tag{77}$$

$$U^{n+1}_j = \frac{1}{2}\left(U^n_j+U^*_j+\frac{\Delta t}{\Delta x}\left(F^*_j-F^*_{j-1}\right)\right) \tag{78}$$

where $F^n_j = F\left(U^n_j\right)$, $F^*_j = F\left(U^*_j\right)$.

This scheme is a two-stage which evaluates a "predictor" U^*_j and a "corrector" $U^{**}_j = U^*_j+\frac{\Delta t}{\Delta x}\left(F^*_j-F^*_{j-1}\right)$ and then forms U^{n+1}_j as the average $\left(U^{**}_j+U^*_j\right)/2$.

It is clear that in order to evaluate U^{n+1}_j the scheme uses the same points in the grid at time in as Lax-Wendroff scheme. The "efficiency" of a scheme is often related to the cost in computer time of each iteration. In these terms, one can compere different schemes. For the Lax-Wendroff scheme, F^n_{j+1}, F^n_j, F^n_{j-1}, $A^n_{j+\frac{1}{2}}$ and $A^n_{j-\frac{1}{2}}$ need to evaluate and perform matrix multiplications in each iteration, whereas MacCormack Scheme requires only the evaluation of F^n_{j+1}, F^n_j, F^*_j and F^*_{j-1}.

It only remains to prove the order of accuracy of MacCormack scheme. The local truncation error of the MacCormack scheme is $O\left(\Delta t^2\right)+O\left(\Delta x^2\right)+O(\Delta t\Delta x)$ in which $\Delta t = O(\Delta x)$. Thus it has a second order accuracy in space and time.

The MacCormack scheme uses forward difference for the predictor and backward difference for the corrector steps. It has second order accuracy as the Lax-Wendroff method. It is much easier to apply, since it is no need to evaluate the second time derivatives [19].

Among the class of second order non-dissipative schemes is the leapfrog scheme. For the general non-linear equation, the scheme is given by:

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{\Delta x}\left(F_{j+1}^n - F_{j-1}^n\right) \tag{79}$$

This scheme is analyzed in detail for the linear case, found out that it is not dissipative but stable, provided that $\frac{\Delta t}{\Delta x}\rho(A) < 1$. The fact that Eq. 79 is accurate of second order follows a straightforward calculation. This scheme is generally more efficient than Lax-Wendroff schemes, although it needs roughly twice as much memory due to the dependence on two previous time stages to evaluate U^{n+1}, therefore in practice, must face the trade-off between efficiency and storage requirements. Since this is a non-dissipative scheme, it will not give good approximations for nonlinear equations. A dissipative term is introduced to Eq. 79 to deal with problem. When adding a dissipative term in the form of a small perturbation, care must be taken so that the resulting linear scheme retains stability. Recall that in the linear case $F(u) = Au$, the amplification matrix $G(\xi)$ is a $2p \times 2p$ matrix (A itself is a $p \times p$ matrix)

$$G(\xi) = \begin{pmatrix} 2i\lambda A \sin(\xi)\ I \\ I \qquad\quad 0 \end{pmatrix} \tag{80}$$

where now each of the entries is itself a $p \times p$ matrix. In order to express the eigenvalues $\mu(\xi)$ of G in terms of those of A, if A is diagonalizable by a matrix T, then G possesses the same eigenvalues of G.

$$\begin{aligned} \widehat{G}(\xi) &= \begin{pmatrix} T^{-1} & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} 2i\lambda A \sin(\xi)\ I \\ I \qquad\quad 0 \end{pmatrix}\begin{pmatrix} T & 0 \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} 2i\,\lambda\, T^{-1} A\, T \sin(\xi)\ I \\ I \qquad\qquad\quad 0 \end{pmatrix} \end{aligned} \tag{81}$$

Recall that $T^{-1}AT$ is a diagonal matrix with diagonal entries $a_1, \ldots, a_p$. From this expression, it follows that any eigenvalue $\mu(\xi)$ of the amplification matrix satisfies:

$$\mu^2(\xi) = 1 + 2i\lambda a_j \sin(\xi)\mu(\xi), \quad j = 1, 2, \ldots, p \tag{82}$$

If a dissipative term is added to the leapfrog scheme at time level n, this will cause rise to instabilities.

$$\varepsilon\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right) \tag{83}$$

added to the scheme (Eq. 79) where ε is a small perturbation. Notice that any modification at time level n will affect the first block in the amplification matrix. The modified amplification matrix will be of the form:

$$G(\xi) = \begin{pmatrix} 2i\lambda A \sin(\xi) + \varepsilon \sin^2(\xi/2) I & I \\ I & 0 \end{pmatrix} \tag{84}$$

and therefore the eigenvalues will now satisfy:

$$\mu^2(\xi) = 1 + \big(2i\lambda a_j \sin(\xi) + \varepsilon \sin^2(\xi/2)\big)\mu(\xi) \tag{85}$$

E denotes the shift operator $E\,U_j^n = U_{j+1}^n$, adding a dissipative term at time level n amounts to modifying Eq. 74 yielding the scheme:

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{\Delta x} A\big(U_{j+1}^n - U_{j-1}^n\big) + \varepsilon P(E) U_j^n \tag{86}$$

where $P(E)$ is a function of the shift operator $\big(P(E) = E - 2I + E^{-1}\big)$. Since $P(E)$ approximates a second order derivative, its Fourier transform $\widehat{P}(\xi)$ will be a real function of ξ and thus the modified eigenvalues will in general satisfy:

$$\mu^2(\xi) = 1 + \Big(2i\lambda a_j \sin(\xi) + \varepsilon \widehat{P}(\xi)\Big)\mu(\xi) \tag{87}$$

for some eigenvalue at of A.

Let x_1 and x_2 are the solutions of the equation $x^2 - ax - 1 = 0$. If both $|x_1| \leq 1$ and $|x_2| \leq 1$, then necessarily the coefficient α is purely imaginary.

Using exactly the same analysis, the leapfrog scheme gives rise to instabilities when it is used to approximate parabolic equations. For the heat equation, this can also be explained by the stability region of the leapfrog method, which is only on the imaginary axis, while the centered finite difference used in approximating the second order derivatives will give real eigenvalues.

In order to introduce the correct amount of dissipation, the dissipation term at time level $n - 1$ should be added. The operator $E^{1/2}U_j^n = U_{j+1/2}^n$ so the leapfrog scheme Eq. 79 can be rewritten in the form:

$$U_j^{n+1} = U_j^{n-1} + \frac{\Delta t}{\Delta x}\big(E^{1/2} - E^{-1/2}\big)\big(E^{1/2} + E^{-1/2}\big)F_j^n \tag{88}$$

in general form:

$$\begin{aligned} U_j^{n+1} = U_j^{n-1} &+ \frac{\Delta t}{\Delta x}\big(E^{1/2} - E^{-1/2}\big)\big(E^{1/2} + E^{-1/2}\big)F_j^n \\ &- \frac{\varepsilon}{16}\big(E^{1/2} - E^{-1/2}\big)^4 U_j^{n-1} \end{aligned} \tag{89}$$

$\eta = \sin(\xi/2)$ the amplification matrix of the linearized scheme

$$G(\xi) = \begin{pmatrix} 2i\lambda A \sin(\xi) & \big(1 - \varepsilon\eta^4\big) I \\ I & 0 \end{pmatrix} \tag{90}$$

and the eigenvalues hold the relations:

$$\mu^2(\xi) = 1 - \eta^4 + 2i\lambda\mu(A)\sin(\xi)\sin(\xi)\mu(\xi) \tag{91}$$

for some eigenvalue $\mu(A)$ of A. Therefore:

$$\mu(\xi) = i\lambda\mu(A)\sin(\xi) \pm \sqrt{1 - |\mu(A)|\sin^2(\xi) - \varepsilon\eta^4} \tag{92}$$

And $|\mu(\xi)|^2 = 1 - \varepsilon\eta^4$ provided that

$$1 - |\lambda\mu(A)|^2 \sin^2(\xi) - \varepsilon\eta^4 > 0 \tag{93}$$

for all eigenvalues of A and ξ. Under this condition, the modified scheme Eq. 89 is stable and dissipative. Remark, though, that in order for Eq. 93 to hold, whenever add dissipation ($\varepsilon > 0$), and also must decrease the value of $\lambda = \frac{\Delta t}{\Delta x}$. This means that for a fixed space grid, a large number of time steps must be evaluated to get an approximate solution at some given time t.

3 Finite Difference Time Domain Applications in Electrical Engineering

The FDTD method used in the solution of Maxwell's equations allows to analyze the electric and magnetic fields and interactions with medium. Maxwell's equations, which have differentials in time and space dimensions, are solved by using the future and past values of time and space [20]. In the solution of this discrete set of space and time, the electric and magnetic fields are resolved interconnectedly and the value obtained in each step becomes the first value for the next step [21]. The relationship between these two parameters is described in Fig. 1.

This method is used in many different applications since it is a very powerful tool for solving partial differential equations. These applications include many engineering problems such as the percussion instrument model, where different sampling frequencies are used to reduce the simulation time, the grounding characteristics of wind turbines in low resistive soil, the propagation of partial discharge signals in HV current transformers, and the analysis of electromagnetic interaction currents flowing in the power cables of DC-DC converters [20, 22, 23]. Since the examination of all these application areas is beyond the scope of this chapter, the applications of the FDTD method in the power system and high voltage industry are discussed.

Electromagnetic transient and non-transient simulations become an important tool for planning, operation and fault analysis in electrical power systems [8]. These simulations concentrate on transient state analyzes that occur in power system equipment such as circuit breakers (CBs), lightning arrestor, overhead and underground cables, ultra-high frequency (UHF) sensors and power transformers (PTs) [8, 9, 24, 25].

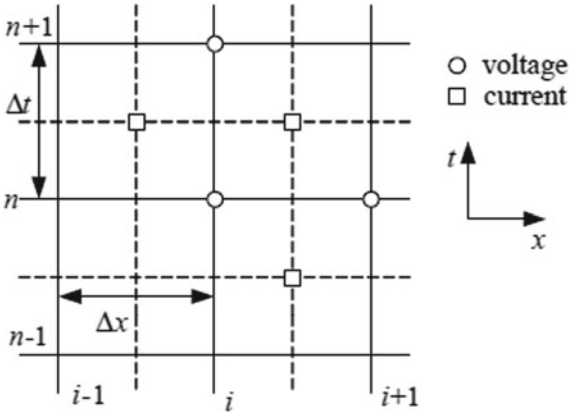


Fig. 1 Time and special discretization in FDTD method [21]

Determination of lightning induced voltage and current caused by lightning discharges, one of the major sources of fault in power systems, is critical for the protection of power system equipment [26]. Aodsup and Kulworawanichpong [24] examined the propagation and reflection of the lightning strike in the lightning arrester with silicon carbide (SiC) and metal oxid varistor (MOV) by adapting the Telegraphist equations to the FDTD method. According to the simulation results, MOV arrester reflects and transmits the impulse surge smooter than SiC arrester [24]. Nagarjuna and Chandrasekaran [21] adapted the transmission line approach to FDTD equations to examine the current and voltage characteristics of the horizontal ground electrode at high impulse currents. Izadi et al. [26] calculated the electric and magnetic fields in different time and space using Maxwell's equations and 2nd order FDTD while advancing of the lightning channel in the power system, and the proposed algorithm showed a good agreement with the measurement results.

Analyzing the electromagnetic behavior of overhead and underground cables, one of the important parts in power system transmission and distribution, improves system design. These cables, consisting of multiple layers with different characteristics between the cable core and the shell, can be successfully modeled by the FDTD method with a high spatial discretization [27]. In addition, frequency-dependent FDT models are numerically unstable or computational time is excessive. Additionally, underground cable applications of these models are very limited. To overcome these limitations, the FDTD method can be developed by taking into account distributed fixed parameters such as skin effect and imperfect earth in overhead lines [27]. It is an important problem to analyze the transient state responses of electromagnetic fields in underground cables, which are used more and more for environmental, political and technical reasons in high voltage applications. Barakou et al. [28] used the universal line model (ULM) and FDTD method to model these lines. While the FDTD method provides very high accuracy for slow front surge, these results are distorted by temporary fluctuations for fast front surges. Either way, simulation times are almost six times the ULM and are quite slow [28].

As a result of the operation of power system equipment such as disconnectors or CBs in gas-insulated substations (GISs), switching pulses called very fast transient (VFT) may occur in the frequency range from several MHz to more than 100 MHz [9]. Calculated the transient electromagnetic disturbances caused by these frequencies using FDTD and EMTP and found that the results obtained by the FDTD are less oscillating. Shakeri et al. [25] examined the effect of VFTs on power transformers, one of the most important equipment in power systems, by adapting the multi-conductor transmission line theory to the FDTD method and observed the effect of electromagnetic waves. In order to increase the accuracy of this model, the winding capacity matrix is calculated by FEM analysis, and the simulation results are obtained with a certainty to confirm the experimental results [25].

FDTD method is also used in the electromagnetic modeling stage to understand the behavior and improve the performance of UHF sensors used to detect partial discharges that can be dangerous for power systems and transformers [6]. Ishak et al. [22], using the FDTD integrated UHF sensor developed for this purpose, achieved agreement results with experimental results in a wide frequency range of 500–1500 MHz. Another proposed approach to determine the behavior of the UHF based test system used as a PD sensor in high voltage cables and to investigate the PD coupling process is the combination of the FDTD method and the transfer function theory [2]. This proposed approach has been applied to 11 kV XLPE cable by Hu et al. [2]. In another application where the amplitude and charge of the partial discharge current in gas insulated switches are examined, the data obtained with a voltage probe placed on the outer surface of a three-phase gas-insulated switch in the 84 kV-class are verified by simulation results obtained by the three-dimensional FDTD method [29].

Busbar structures commonly used in high voltage transmission are also affected by electromagnetic fields radiated from switching operations. These analyzes become even more important for high voltage equipment located close to switching equipment and electronic circuits of these [4]. The FDTD method is successfully used in the modeling of busbar structures in high voltage air-insulated substations [5]. Musa et al. [4] modeled transient electromagnetic fields as a result of switching operations in a 400 kV air insulated substation using simply specifying their constitutive parameters with this method.

Grounding behavior, which is one of the important parameters to ensure system reliability in transmission and distribution systems, can also be examined with the FDTD method. In this context, the soil ionization phenomenon, which reduces the ground electrode resistance, has been investigated by the FDTD method and applied to a typical high voltage substation of 500/220 kV [7].

Finite element and difference methods are also widely used in high voltage technique to calculate the breakdown characteristics of gas dielectrics [30]. The electron drift velocity, mean energy, ionization and attachment coefficients of dielectric gases such as SF_6, CF_4, CHF_3, and argon, which are frequently used in the insulating gas industry, can be calculated by using the finite difference method for solving Boltzmann equations [31–33].

Despite this widespread use of the FDTD method, there are also limitations such as defining the uncertainties they have due to the nature of electromagnetic fields, and the excessive computation time needed in the analysis of large objects [9, 34].

In the calculation of electromagnetic fields, the properties of the object, such as geometric properties, electrical parameters, material characteristics and input sources, can increase randomness and thus uncertainty [35]. The uncertainty in these input parameters is reflected in the electromagnetic fields, which are the output parameters, and a parametric uncertainty appears in the resulting components [36]. Identifying these uncertainties, which are very important in some engineering problems, is also a major research topic in the analysis of electromagnetic fields [35]. As an alternative to the Monte Carlo method used to identify parametric uncertainties, there are many methods and approaches combined with FDTD. In this context, methods such as stochastic, polynomial chaos, control variations, and the method of moments are combined with the FDTD method and defined the uncertainties in the calculation of electromagnetic fields [35–38].

Chen [34] used the hybrid implicit explicit approach in combination with the FDTD method to overcome the problem of electromagnetic modeling in very fine structures. This proposed method is applicable for many boundary conditions including connect boundary, absorbing boundary and periodic boundary. In order to overcome the computation time problems encountered in the electromagnetic modeling of electrically large objects, Shi et al. [39] proposed FDTD method combined with Internet of Things, in which multiple processors are connected in parallel.

Another important problem of the FDTD method is the increased response time at high frequencies and the decrease in the accuracy of the analysis results [27]. In order to overcome this disadvantage, alternative models and software combined with FDTD method are used in the analysis of high frequency transient situations frequently encountered due to switching and lightning in high voltage equipment, especially cables [2, 27].

4 Conclusions

In this chapter, the time-dependent finite difference method, which is widely used in the solution of engineering problems defined by differential equations, is examined in the theoretical framework and application examples. In order to limit the examination in terms of engineering applications, the use of the finite difference method in the analysis of electric and magnetic fields in power systems and high voltage equipment is concentrated. The limitations of the finite difference method are defined and how these limitations can be overcome by combining them with different methods and approaches are discussed. Finite difference method, which is an important tool in robust and accurate calculation of electromagnetic fields, has been used more widely in transient analysis as well as in steady state analysis.

References

1. Yee K (1966) Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. IEEE Trans Antennas Propag 14(3):302–307. https://doi.org/10.1109/TAP.1966.1138693
2. Hu X, Siew WH, Judd MD, Reid AJ, Sheng B (2019) Modeling of high-frequency current transformer based partial discharge detection in high-voltage cables. IEEE Trans Power Delivery 34(4):1549–1556. https://doi.org/10.1109/TPWRD.2019.2910076
3. Gedney SD (1996) The application of the finite-difference time-domain method to EMC analysis. In: Proceedings of symposium on electromagnetic compatibility, Santa Clara, CA, USA, pp 117–121. https://doi.org/10.1109/isemc.1996.561212
4. Musa BU, Siew WH, Judd MD, Wang T, Li QM (2013) Application of finite difference time domain method to high voltage substations: switching transient fields. Int J Eng Sci Innovat Technol (IJESIT) 2(5):20–29
5. Musa BU, Siew WH, Judd MD (2010) Computation of transient electromagnetic fields due to switching in high-voltage substations. IEEE Trans Power Delivery 25(2):1154–1161. https://doi.org/10.1109/TPWRD.2009.2034008
6. Li T, Wang X, Zheng C, Liu D, Rong M (2014) Investigation on the placement effect of UHF sensor and propagation characteristics of PD-induced electromagnetic wave in GIS based on FDTD method. IEEE Trans Dielectr Electr Insul 21(3):1015–1025. https://doi.org/10.1109/TDEI.2014.6832244
7. Ghania SM (2019) Grounding systems under lightning surges with soil ionization for high voltage substations by using two layer capacitors (TLC) model. Electric Power Syst Res 174(105871). https://doi.org/10.1016/j.epsr.2019.105871
8. Noda T (2018) A study of an FDTD-based frequency-dependent line model for electromagnetic transient simulations. Electr Eng Jpn 202(2):33–42. https://doi.org/10.1002/eej.23038
9. Ametani A, Xue H, Natsui M, Mahseredjian J (2018) Electromagnetic disturbances in gas-insulated substations and VFT calculations. Electr Power Syst Res 160:191–198. https://doi.org/10.1016/j.epsr.2018.02.014
10. DuChateau P, Zachmann DW (2002) Applied partial differential equations. Dover Publications, New York. ISBN: 978-0-486-41976-3
11. Sheikh MAA, Andallah LS, Kowser MA (2014) A comparative study of finite difference scheme for Burger's equation. Gazi Univ J Sci GU J Sci 27(4):1045–1052
12. Versteeg HK, Malalasekera W (2007) An introduction to computational fluid dynamics, 2nd edn. Pearson Education Limited. ISBN: 978-0-13-127498-3
13. Zhang X, MA615 numerical methods for PDEs, spring 2020 lecture notes. Math Dept, Purdue University
14. Shampine L (2009) Stability of the leapfrog/midpoint method. Appl Math Comput 208(1):293–298. https://doi.org/10.1016/j.amc.2008.11.029
15. Sun C, Trueman CW (2003) Unconditionally stable Crank-Nicolson scheme for solving two-dimensional Maxwell's equations. Electron Lett 39(7):595–597. https://doi.org/10.1049/el:20030416
16. Käppeli R, Balsara DS, Chandrashekar P, Hazra A (2020) Optimal, globally constraint-preserving, DG(TD)2 schemes for computational electrodynamics based on two-derivative Runge-Kutta time stepping and multidimensional generalized Riemann problem solvers—a von Neumann stability analysis. J Comput Phys 408:109238. https://doi.org/10.1016/j.jcp.2020.109238
17. Angus DA (2014) The one-way wave equation: a full-waveform tool for modeling seismic body wave phenomena. Surv Geophys 35:359–393. https://doi.org/10.1007/s10712-013-9250-2
18. Mulder WA, Plessix RE (2004) A comparison between one-way and two-way wave-equation migration. Geophysics 69(6):1491–1504. https://doi.org/10.1190/1.1836822
19. Ching MH, Deiwert GS, Inouye M (2002) The MacCormack method—historical perspective. Front Computat Fluid Dyn 45–59. https://doi.org/10.1142/9789812810793_0003

20. Fontana F, Bozzo E, Novello M (2015) Decimation in time and space of finite-difference time-domain schemes: standard isotropic lossless model. IEEE Trans Signal Process 63(20):5331–5341. https://doi.org/10.1109/TSP.2015.2453139
21. Nagarjuna K, Chandrasekaran K (2019) Analysis of horizontal grounding electrode in transmission line approach. In: 2019 international conference on communication and electronics systems (ICCES). Coimbatore, India, pp 267–272
22. Ishak AM, Ishak MT, Jusoh MT, Syed Dardin SF, Judd MD (2017) Design and optimization of UHF partial discharge sensors using FDTD modeling. IEEE Sens J 17(1):127–133. https://doi.org/10.1109/jsen.2016.2628035
23. Laour M, Tahmi R, Vollaire C (2017) Experimental evaluation and FDTD method for predicting electromagnetic fields in the near zone radiated by power converter systems. Turk J Elec Eng Comp Sci 25:1460–1471. https://doi.org/10.3906/elk-1506-278
24. Aodsup K, Kulworawanichpong T (2017) Analysis of surge propagation with lightning arrester using FDTD for 25 kV-AC transmission line. Int J Eng Appl Sci 4(4):66–70
25. Shakeri J, Abbasi AH, Shayegani AA, Mohseni H (2010) FDTD simulation of voltage distribution in transformer winding under VFTO Phenomena. Int Rev Electr Eng 5(1):130–137
26. Izadi M, Ab Kadir MZA, Gomes C, Wan Ahmad WF (2010) An analytical second-FDTD method for evaluation of electric and magnetic fields at intermediate distances from lightning channel. Prog Electromag Res 110:329–352. https://doi.org/10.2528/pier10080801
27. Kaloudas CG, Chrysochos AI, Papagiannis GK (2014) FDTD analysis of multiphase power cable systems using distributed constant parameters. In: MedPower 2014, Athens, pp 1–8. https://doi.org/10.1049/cp.2014.1711
28. Barakou F, De Silva HMJ, Wouters PAAF, Steennis EF (2018) Evaluation of FDTD model for transient studies with complicated cable configurations. In: 2018 power systems computation conference (PSCC), Dublin, pp 1–7. https://doi.org/10.23919/pscc.2018.8442741
29. Tanaka H, Tanahashi D, Baba Y, Nagaoka N, Okada N, Ohki H, Takeuchi M (2016) Finite-difference time-domain simulation of partial discharges in a gas insulated switchgear. High Volt 1(1):52–56. https://doi.org/10.1049/hve.2016.0006
30. Duzkaya H, Dincer MS, Hiziroglu HR (2009) Calculation of partial discharge inception voltages in ultradilute $SF_6 + N_2$ gas mixtures. In: Conference on electrical insulation and dielectric phenomena, Virginia Beach, VA, USA, pp 531–534. https://doi.org/10.1109/ceidp.2009.5377903
31. Tezcan SS, Akcayol M, Ozerdem OC, Dincer MS (2010) Calculation of electron energy distribution functions from electron swarm parameters using artificial neural network in SF_6 and argon. IEEE Trans Plasma Sci 38(9):2332–2339. https://doi.org/10.1109/TPS.2010.2049588
32. Tezcan SS, Duzkaya H, Dincer MS, Hiziroglu HR (2016) Assessment of electron swarm parameters and limiting electric fields in $SF_6 + CF_4 +$ Ar gas mixtures. IEEE Trans Dielectr Electr Insul 23(4):1996–2005. https://doi.org/10.1109/TDEI.2016.005435
33. Duzkaya H, Tezcan SS (2019) Boltzmann analysis of electron swarm parameters in $CHF_3 + CF_4$ mixtures. Turk J Electr Eng Comp Sci 27:615–622. https://doi.org/10.3906/elk-1804-187
34. Chen J (2018) A review of hybrid implicit explicit finite difference time domain method. J Comput Phys 363:256–267. https://doi.org/10.1016/j.jcp.2018.02.053
35. Zygiridis TT (2017) A short review of FDTD-based methods for uncertainty quantification in computational electromagnetics. Mathemat Prob Eng (9247978). https://doi.org/10.1155/2017/9247978
36. Edwards RS, Marvin AC, Porter SJ (2010) Uncertainty analyses in the finite-difference time-domain method. IEEE Trans Electromagn Compat 52(1):155–163. https://doi.org/10.1109/TEMC.2009.2034645
37. Gu Z, Zhang X, Sood N, Sarris CD (5) Efficient multi-parametric uncertainty quantification methods for EMC/EMI applications. In: 2015 IEEE symposium on electromagnetic compatibility and signal integrity, Santa Clara, CA, pp 361–364. https://doi.org/10.1109/emcsi.2015.7107715
38. Ganta SS, Van Veen BD, Hagness SC (2017) On the accuracy of polynomial models in stochastic computational electromagnetics simulations involving dielectric uncertainties. IEEE Antennas Wirel Propag Lett 16:2594–2597. https://doi.org/10.1109/LAWP.2017.2733543

39. Shi Q, Zou B, Zhang L, Liu D (2019) Hybrid parallel FDTD calculation method based on MPI for electrically large objects. Wireless Commun Mobile Comput (7309431). https://doi.org/10.1155/2019/7309431

Theoretical Approaches of Finite Elements Method (FEM)

Marilena Stanculescu, Sorin Deleanu, Paul Cristian Andrei, Horia Andrei, Lavinia Bobaru, and Mihai Iordache

Abstract In engineering practices, electromagnetic field problems are of two main types: analysis and synthesis problems. In principle, a synthesis problem can be solved by an iterative technique, repeating the solving of some analysis problems, for configurations adapted because of some preliminary results of analysis simulations. Therefore, knowing how to compute the electromagnetic field precisely is a crucial issue in most of electrical engineering problems. This chapter is dedicated to Theoretical Approaches of Finite Elements Method (FEM).

Keywords Finite element method (FEM) · Electromagnetic field computation · Numerical methods

M. Stanculescu (✉) · P. C. Andrei · L. Bobaru · M. Iordache
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania
e-mail: marilena.stanculescu@upb.ro

P. C. Andrei
e-mail: paul.andrei@upb.ro

L. Bobaru
e-mail: lavinia.bobaru@upb.ro

M. Iordache
e-mail: mihai.iordache@upb.ro

S. Deleanu
Northern Alberta Institute of Technology, Edmonton, Canada
e-mail: sorind@nait.ca

H. Andrei
Department of Doctoral School, University Valahia, Targoviste, Romania
e-mail: hr_andrei@yahoo.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_3

Nomenclatures

A. Acronyms

FEA	Finite element analysis
FEM	Finite element method (FEM)
FDM	Finite difference method (FDM)
FE	Finite elements
BVP	Boundary value problems
BC	Boundary conditions

B. Symbols/parameters

$\boldsymbol{M}_p$	Permanent magnetization
I	Magnetic polarization
B	Magnetic flux density
H	Magnetic field strength
μ	Absolute magnetic permeability
μ	Magnetic permeability of the vacuum
V	Scalar magnetic potential
J	Electric conduction current density vector
A	Magnetic potential vector
φ_i	Form functions
ψ_i	Trial functions
D_Σ	Computation domain
$\Sigma = S_D \cup S_N$	Surface which bounds the computation domain

1 Introduction

Finite element analysis (FEA) applies to a wide variety of electrical engineering applications, including solid mechanics, fluid mechanics, heat transfer, acoustics. The description of the modalities of applying the FEA within the different disciplines is quite complicated, because it involves the use of an abstract mathematical approach, while the particularization of a domain or discipline limits the mathematical abstraction [1].

The finite element method (FEM) is a technique utilized in the process of obtaining approximate solutions of the boundary values for mathematical physics problems. This method was first proposed in the 1940s and began its journey in the early 1950s for aircraft design. Later, this method was developed and has been widely used in structural analysis problems and on a smaller scale in other areas. At present, this method is recognized to be one of the general methods with applicability in various fields, from those related to engineering and to mathematical problems [2].

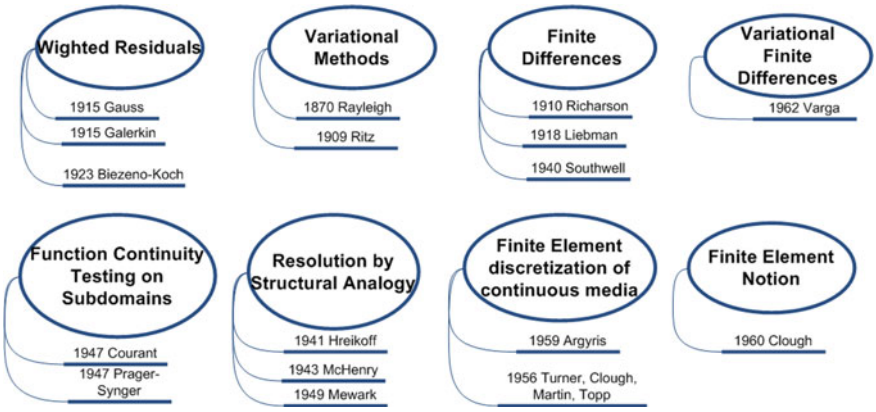


Fig. 1 Methods and theories on which FEM is based

There are many papers written in this field, yet Zienkiewicz's book is one of the most well-known texts of this genre. Zienkiewicz and Cheung have given a wide interpretation of FEM and practically underlines FEM's applicability to any engineering problem. With this general interpretation, FEM equations can also be obtained using the weighted residuals method, such as the Galerkin method or the least-squares approach. All of these have led to an increasing interest among specialists in applied mathematics, applying FEM for solving linear and nonlinear problems. A plethora of literature on the topic, from conferences and books, confirms the continuous scientific interest in FEM. The application of FEM in the current engineering applications has its fundamentals in the theories and methods synthesized in Fig. 1 [3–12].

Computers provide fast means of performing a large volume of calculations involved in FEA and made the method practically applicable. FEM requires the utilization of computationally powerful computers. FEM application has progressed at an impressively high speed due to the development of high-speed digital computers.

The necessity of improving the performances of electrotechnical and electronic components leads to the development of the electromagnetic field computation methods. These problems imply solving some complicated equation systems containing partial derivatives. If the laws referring to the materials (called "laws of material") are linear (linear media), then the corresponding systems are linear. Most of the time, the laws of material are not linear; therefore, the corresponding systems of equations are not linear. The computation of such systems involves many iterations (i.e., "iterative manner"), at each iteration being computed a linear field problem, using the known procedures from the linear media.

In engineering practices, electromagnetic field problems are of two main types: analysis and synthesis problems. Analyses problems consist of determining the corresponding magnetic field to some given uniqueness conditions problems. Synthesis problems consist in determining the configurations capable of producing an electromagnetic field holding some characteristics.

In principle, a synthesis problem can be solved by an iterative technique, repeating the solving of some analysis problems, for configurations adapted because of some preliminary results of analysis simulations.

Whereas using a mathematical approach as a classification criterion, there are analytical methods and numerical methods. Analytical methods utilize a set of equations (in general, with partial derivatives) of the electric and or magnetic field directly. Such methods of computation require simple enough applications (e.g., configurations with some symmetries). However, they may have their niche, as auxiliary means in more complex studies.

The numerical methods seek to determine the solution in a discrete set of points in the studied domain. Although apparently, the numerical methods can approach the most real problems of determining some electromagnetic fields, the reality is much more "modest." Even if some calculus systems get more and more capable and more and more approachable, numerous limitations persist in this domain. Despite the abundance of performant (and generally expensive!) software packages commercially available, addressing the problem of numerical computation of electromagnetic field, there is a strong demand for proper knowledge of electromagnetism when it comes to modeling. Nowadays, having access on a large scale to performant computers allowed numerical methods to gain more steam.

Firstly, scientists developed and applied the Finite Difference Method (FDM), which originated from the expansion of Taylor series. FDM has its well-known drawback when it comes to express the Neumann boundary conditions at interfaces. The answer to such a problem was the development of the Finite Element Method (FEM), which conveniently addressed some limitations of the FDM, particularly the "stiffness" in the construction of the meshing scheme.

The main goal of this chapter is to give an insight into the theoretical approaches of the FEM. The construction of the finite element approximation relies on partial differential equations as expressions of the solutions defined by a partition of the field study in disjoint elements, called "finite elements" (FE), giving the name of the method. Although such an approach looks similar to the FDM, the procedure used in FEM is different.

This chapter starts with a presentation of the importance of the principle of the FEM. The description of the electromagnetic field (and a field in general) appeals to several formulations for its equations, including the integral form of laws and the differential forms of laws.

In terms of a concrete approach, FEM has two main variants: the Galerkin method and the Ritz method. The distinction between these approaches is not always obvious. For this reason, the terminology may show some differences (i.e., the Ritz method is sometimes called the Ritz-Galerkin). The sections of this chapter contain a detailed presentation of these methods.

The so-called "weak formulation" of the electromagnetic field equations characterized the Galerkin method; therefore, it is presented the weak form using the scalar electric potential and weak form using magnetic vector potential, followed by the principle of the Galerkin method.

For a better understanding, each section comes with clear examples. In such examples, the FEM turns to be extremely flexible and easy to customize, pending upon the type of problem in place for analysis (e.g. electrostatics, magnetostatics, AC analysis, transient analysis and the geometry, 2D or 3D).

The chapter ends with the conclusions section, and a comprehensive set of bibliographic references accompanies it.

2 The Principle of Finite Elements Method (FEM)

In many situations, existing analytical equipment is not enough for finding the exact solution (or, in some cases, the approximate one) for problems with immediate practical application. The basic idea in FEM is to get the solution to a complex problem by replacing it with a less complicated one.

The basic idea is that if the structure suffers a division into several parts called "finite elements". For each of them, it follows the application of the computation theories corresponding to the adopted schematization. The division of the whole into smaller parts, an operation that is called "discretization," has the effect of obtaining simple forms for the finite elements of the structure. The calculation model used in FEA is approximate, obtained by assembling finite element components, considering the geometry of the structure. The connection of the finite elements is made only at specific points called nodes. The nodes represent the intersection points of the straight lines or curves of the finite elements. Function of the structure's modeled geometry, the finite elements can be 1D, 2D, or 3D kind. Usually, the nodes are points placed on the contours of the element, being points of connection with the adjacent elements of the mesh. A simple function can approximate the variation of the field inside the domain of an FE. These approximation functions (interpolation models) are extracted from the nodes' values of the field variables [13–16].

The approximate character of FEM results from the replacement of the real geometry with a finite element network that follows the original form but can be accurately reproduced only for specific geometries. The finite number of elements imposes the calculation of unknowns only in the nodes of the structure. Consequently, the computation precision increases as the number of finite elements increases.

FEM relies on the concept of constructing complicated objects from simpler objects or dividing complicated objects into simpler objects, suitable for the application of known computation schemes.

Steps of applying FEM

1. Study of the structure

This step is essential for choosing a computation model and the types of finite elements suitable to reproduce as accurately as possible the real structure. The choice of the types of finite elements follows the demands regarding the precision and quality of the results we want to obtain [17–19, 21].

2. The discretization of the structure (mesh building)

In this step, the objective of the discretization mechanism, in areas of interest, must be to obtain as small as possible finite elements in the area of interest. The transition from areas with small finite elements to large finite elements must follow a progressive pass. When choosing the discretization mode, one must not distort the finite elements. Most of the professional programs dedicated to FEM analysis provide modules of data preprocessing and make automatic discretization.

Moreover, in these cases, verifying the configuration of the used finite elements is an essential step in solving the problem with minimal errors in the proposed analysis [20].

The electromagnetic field equations require adequate formulation for applications in the energy problems domain.

The laws, expressed in two kinds of forms, integral, respectively differential, address the FEM problem formulation Furthermore, Galerkin' s Method applies the so-called weak form of laws. In contrast, Ritz's Method considers the variational form based on minimizing a function that derives from the field equations.

FEM or FEA is a computation technique used to obtain approximate boundary value problems (BVP) in engineering. BVPs are also known as field problems.

The field represents the area of interest, and most of the time, represents a physical structure. Domain variables are those dependent interest variables governed by differential equations. Boundary conditions (BC) are those specific values of domain variables (or related variables such as their derivatives) attached to the domain boundaries. Certain differential equations, followed by a set of constraints, called boundary conditions, are assembling the BVP problems. Finding the solutions for such problems, one applies either the Galerkin Method or Ritz Method mentioned before. The distinction between these approaches is less clear than it seems. For this reason, the terminology may also have some differences (for example, the Ritz method is sometimes called the Ritz-Galerkin) [21].

3 Galerkin Method

3.1 Weak Form of Equations of the Steady-State Magnetic Field

Galerkin Method requires the so-called weak formulation of the electromagnetic field equations [22].

3.1.1 Weak Form Formulation Using the Scalar Magnetic Potential

Let us consider a magnetostatics field problem, in a domain in which there can exist bodies with permanent magnetization. For simplicity, we assume there are no discontinuity surfaces. The corresponding field equations are [22, 23]:

$$\oint_{\Gamma} \boldsymbol{H} dl = 0 \rightarrow \boldsymbol{H} = -\text{grad}V \,(\text{Magnetic circuit Law}) \tag{1}$$

$$\boldsymbol{B} = \mu \boldsymbol{H} + \mu_0 \boldsymbol{M}_p = \mu \boldsymbol{H} + \boldsymbol{I} \,(\text{Material Law}) \tag{2}$$

where the vectors: $\boldsymbol{M}_p$—the "permanent magnetization", $\boldsymbol{I}$—"magnetic polarization" (with its components permanent and temporary), $\boldsymbol{B}$—"magnetic flux density", $\boldsymbol{H}$—"magnetic field strength", respectively the scalars: μ—"absolute magnetic permeability" of the medium, and μ_0—"magnetic permeability of the vacuum".

$$\oint_{\Sigma} \text{B} ds = 0 \rightarrow \text{divB} = 0 (\text{Magnetic Flux Law}) \tag{3}$$

From Eqs. 3.1, 3.2, 3.3, it results:

$$\text{divB} = 0 = -\text{div}(\mu \, \text{grad}V) + \text{divI} \tag{4}$$

where V is the scalar magnetic potential. If we define the "magnetic charge", having the density $\rho = -div\boldsymbol{I}$, the following equation becomes:

$$-\text{div}(\mu \, \text{grad}V) = \rho \tag{5}$$

Regarding the computation domain from Fig. 2, one uses the notations by D_Σ—the computation domain, respectively $\Sigma = S_D \cup S_N$, $(S_D \cap S_N = \emptyset)$, the surface which bounds the computation domain (Fig. 2).

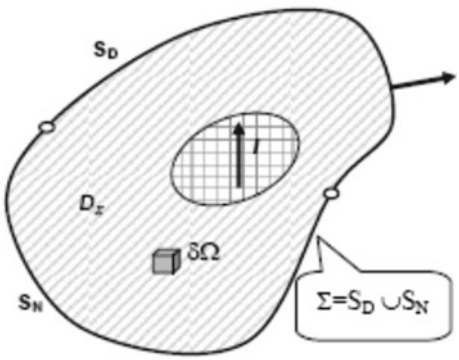


Fig. 2 Computation domain

On this surface Σ, the boundary conditions can be:
Dirichlet conditions

$$V(P) = f(P),\ P \in S_D \tag{6}$$

Neumann conditions:

$$\frac{dV(P)}{d\mathrm{n}} = g(p),\ P \in S_N \tag{7}$$

We are about to consider the functions Y = {Ul: with the following properties:

– its square and derivatives are integrable
– is equal to zero on S_D:

$$U(P) = 0,\ P \in S_D \tag{8}$$

The domain of definition for function U(r) is D_Σ.
By multiplying the Eq. (3.5) with this function, followed by its integration on the domain, it results:

$$\int_{D_\Sigma} U\,\mathrm{div}(\mu\ \mathrm{grad}V)\delta\Omega = -\int_{D_\Sigma} \rho U \delta\Omega \tag{9}$$

At this point there is a computational procedure that involves calculations in both differential and integral form, as follows:

$$\mathrm{div}(U(\mu\ \mathrm{grad}V)) = U div(\mu\ \mathrm{grad}V) + \mu\ \mathrm{grad}U\ \mathrm{grad}V \tag{10}$$

or, by using Gauss-Ostrogradski relation:

$$\int_{D_\Sigma} \mu\ \mathrm{grad}U\ \mathrm{grad}V\delta\Omega = \int_{D_\Sigma} \rho U\delta\Omega + \oint_{\Sigma} \mu U \frac{dV}{dn}\mathrm{dS} \tag{11}$$

From (8), in (11) the surface integral refers only to the Neumann conditions surface S_N:

$$\int_{D_\Sigma} \mu\ \mathrm{grad}U\ \mathrm{grad}V\delta\Omega = \int_{D_\Sigma} \rho U\delta\Omega + \int_{S_N} \mu \mathrm{U} \frac{dV}{dn}\mathrm{dS} \tag{12}$$

respectively

$$\int_{D_\Sigma} \mu \, \text{grad}U \, \text{grad}V \delta\Omega = \int_{D_\Sigma} \rho U \delta\Omega + \int_{S_N} \mu g \text{UdS} \tag{13}$$

The above equation is the "weak form" of the magnetic field equation using scalar potential, and therefore the solution is called the problem's weak solution. It has this name because to obtain the solution, the imposed conditions (derivability, continuity) are less restrictive than those imposed in the "strong form."

3.1.2 Weak Form Using the Magnetic Potential Vector

The considered problem refers to the steady magnetic field, where there are permanent magnetization bodies and conductors circulated by currents. Again, for simplicity, there are no discontinuity surfaces. The corresponding field equations are (14) the magnetic field law (Ampere's Theorem), (15) the magnetic flux law (Gauss's Theorem) respectively (16) the law of material [23, 24]:

$$\oint_\Gamma \boldsymbol{H} dl = \oint_{S_\Gamma} \boldsymbol{J} dS \rightarrow \boldsymbol{rot}\ \boldsymbol{H} = \boldsymbol{J} \tag{14}$$

$$\oint_\Sigma \boldsymbol{B} dS = 0 \rightarrow \boldsymbol{B} = \boldsymbol{rot}\ \boldsymbol{A} \tag{15}$$

$$\boldsymbol{B} = \mu \boldsymbol{H} + \mu_0 \boldsymbol{M}_p \tag{16}$$

where $\boldsymbol{J}$ represents the vector of electric conduction current density.

From Eqs. (14), (15) and (16) results:

$$\text{H} = \frac{1}{\mu}\text{rot A} - \frac{\mu_0 \text{rot M}}{\mu} \rightarrow \text{rot}\left(\frac{1}{\mu}\text{rot A} - \frac{\mu_0 \text{rot M}}{\mu}\right) = \text{J} \tag{17}$$

where $\boldsymbol{A}$ represents the magnetic potential vector. If one defines $\boldsymbol{J}_M = \frac{\mu_0 rot \boldsymbol{M}}{\mu}$, it results:

$$\text{rot}\left(\frac{1}{\mu}\text{rotA}\right) = \text{J} + \text{J}_M = \text{J}_T;\ \boldsymbol{J}_T = \boldsymbol{J} + \boldsymbol{J}_M \tag{18}$$

D_Σ is the computation domain and is bounded by the closed surface $\Sigma = S_D \cup S_N ((S_D \cap S_N = \emptyset)$ (Fig. 3). The boundary conditions on this surface are:

– Homogenous *Dirichlet conditions*, (which comes from the limit condition expressed using the normal component of the vector representing the magnetic flux density or induction B)

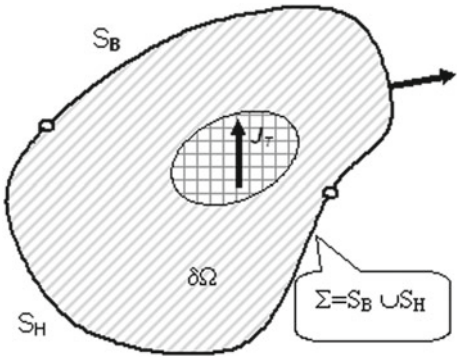


Fig. 3 Computation domain

$$n \times A\big|_{S_n} = b, \mathrm{P} \in \mathrm{S_H} \tag{19}$$

– *Neumann conditions* (from the limit conditions expressed for the tangential component of the magnetic field strength $\boldsymbol{H}$):

$$n \times \left(\frac{1}{\mu} rot \boldsymbol{A}\right)\big|_{S_n} = h(\boldsymbol{P}), \mathrm{P} \in \mathrm{S_H} \tag{20}$$

Let us consider the vector function: $\boldsymbol{U}(r)$ defined in D_Σ.
Let $\mathrm{Y} = \{\mathrm{Ul}$:, be the set of vector functions, with the following properties:

– its square and derivatives are integrable
– there is the following fulfilled condition on S_B:

$$n \times U\big|_{S_n} = 0, \mathrm{P} \in \mathrm{S_B} \tag{21}$$

Then, the Eq. (18) multiplied by the function U, and integrated on the given domain leads to:

$$\int_{\mathrm{D}_\Sigma} \mathrm{Urot}\left(\frac{1}{\mu}\mathrm{rotA}\right)\delta\Omega = \int_{\mathrm{D}_\Sigma} \mathrm{UJ}_T \delta\Omega \tag{22}$$

Using the identity:

$$\mathrm{div}\left(\mathrm{U} \times \frac{\mathrm{rot}\boldsymbol{A}}{\mu}\right) = \mathrm{rotU}\frac{\mathrm{rot}\boldsymbol{A}}{\mu} - \mathrm{Urot}\left(\frac{\mathrm{rot}\boldsymbol{A}}{\mu}\right) \tag{23}$$

results:

$$\int_{D_\Sigma} \frac{1}{\mu} \text{rot} \boldsymbol{U} \text{rot} \boldsymbol{A} \delta\Omega = \int_{D_\Sigma} \boldsymbol{U} \boldsymbol{J}_T \delta\Omega + \int_{D_\Sigma} \text{div}\left(\boldsymbol{U} \times \frac{\text{rot}\boldsymbol{A}}{\mu}\right)\delta\Omega \tag{24}$$

or, by applying Gauss-Ostrogradski:

$$\int_{D_\Sigma} \frac{1}{\mu} \text{rot} \boldsymbol{U} \text{rot} \boldsymbol{A} \delta\Omega = \int_{D_\Sigma} \boldsymbol{U} \boldsymbol{J}_T \delta\Omega + \oint_{\Sigma} \left(\boldsymbol{U} \times \frac{\text{rot}\boldsymbol{A}}{\mu}\right) n \text{dS} \tag{25}$$

The application of the identities:

$$\left(\boldsymbol{U} \times \frac{\text{rot}\boldsymbol{A}}{\mu}\right) n = (n \times \boldsymbol{U}) \frac{\text{rot}\boldsymbol{A}}{\mu} = -\left(n \times \frac{\text{rot}\boldsymbol{A}}{\mu}\right)\boldsymbol{U} \tag{26}$$

gives the possibility for decomposition of the surface integral into the sum of two surface integrals S_B and S_H

$$\int_{D_\Sigma} \frac{1}{\mu} \text{rot} \boldsymbol{U} \text{rot} \boldsymbol{A} \delta\Omega = \int_{D_\Sigma} \boldsymbol{U} \boldsymbol{J}_T \delta\Omega + \int_{S_n} (n \times \boldsymbol{U}) \frac{\text{rot}\boldsymbol{A}}{\mu} \text{dS} - \int_{S_n} \boldsymbol{U}\left(n \times \frac{\text{rot}\boldsymbol{A}}{\mu}\right)\text{dS} \tag{27}$$

In the Eq. (37), whereas considering the condition on the surface S_H, we've obtained:

$$\int_{D_\Sigma} \frac{1}{\mu} \text{rot} \boldsymbol{U} \text{rot} \boldsymbol{A} \delta\Omega = \int_{D_\Sigma} \boldsymbol{U} \boldsymbol{J}_T \delta\Omega - \int_{S_n} \boldsymbol{h}\boldsymbol{U} \text{dS} \tag{28}$$

Equation (28) is the "weak form" using the magnetic potential vector of the magnetic field equation. Therefore, the solution is the "so-called" problem's weak solution.

3.2 The Principle of the Galerkin Method

The Galerkin method requires the application of the following procedure [22, 23]:

a. *The domain is divided (discretized) in disjoint finite elements*

Depending on the function of the nature of the unknown (scalar or vector) and the configuration (2D or 3D problems), one can use various types of discretization.

There are the following notations in place:

N the number of the nodes corresponding to the discretization mesh
Ne the number of elements of the discretized domain

I the number of nodes of the network
I_D the set of nodes with Dirichlet conditions.

b. *Seek solution of the following form*:

$$V(r) = \sum_{i \in I} V_i \varphi_i(r) \tag{29}$$

for the case when the unknown is either a scalar field (e.g. scalar magnetic potential), or even the magnetic potential vector. The quantities V_i, $i \in I$ are the *unknowns* in the network's nodes. One can impose some of the unknowns following the using of Dirichlet type conditions.

The functions φ_i are called form functions and have the following property:

$$\varphi_i(r) = \begin{cases} 1 \; r = r_i \\ 0 \; r \neq r_i \end{cases}; i \in I \tag{30}$$

The function φ_i attached to the node $i \in N$ is equal to 1 in this node and 0 in the remaining nodes. The case of vector fields is treated using a similar procedure.

c. Use the weak form of the equations for a complete set of trial functions $\psi_i \in \mathrm{Y}$, $\mathrm{i} \in \mathrm{N}$. The computation of the integrals present in the weak form, utilizes the analytical expressions of the form functions φ_i and of the trial functions ψ_i, resulting a system of N algebraic equations with N unknowns.

The same function can have both either form and test function, such that, which is a characteristic of Galerkin Method.

d. Solve the equations system obtained at point c.

Remarks

(i) The Dirichlet conditions are active from the beginning of the approximation expression of the solution. That's why they are called essential conditions.
(ii) Neumann conditions don't act directly on the set of functions in which the solution is searched for. They implicitly appear in the weak form of the solution. For this reason, the Neumann condition are also called "natural conditions".

3.3 Approximation Using Finite Elements

FEM, Galerkin and Ritz use the approximation with finite elements.

3.3.1 2D Finite Elements

a. *Local approximation*

The 2D finite elements used to discretize a given domain are disjoint polygonal domains, whose reunion aims to cover as well as possible the shape (exact or approximated) of the domain on which the two variables function under study is defined [23].

The most common case is that corresponding to the linear triangular finite elements, presented as follows: (Fig. 4).

Let's consider one of the triangular elements (e), with the nodes i, j, and $k \in N$, labeled using the direct trigonometric direction (counterclockwise) (Fig. 5).

For linear elements, the form functions are linear and inside each element, they can have value one in one of the nodes, being zero in the remaining ones:

$$\varphi_m^e = \begin{cases} 1 \; \boldsymbol{r} = \boldsymbol{r}_m \\ 0 \; \boldsymbol{r} \neq \boldsymbol{r}_m \end{cases}; m = i, j, k \tag{31}$$

Let be one form function, corresponding to node i. It can have a polynomial form (Fig. 6):

$$\varphi_i^e(x, y) = a_i x + b_i y + c_i \tag{32}$$

From the definition condition, results:

$$\varphi_i^e(P_i) = a_i x_i + b_i y_i + c_i = 1$$
$$\varphi_i^e(P_j) = a_i x_j + b_i y_j + c_i = 0$$

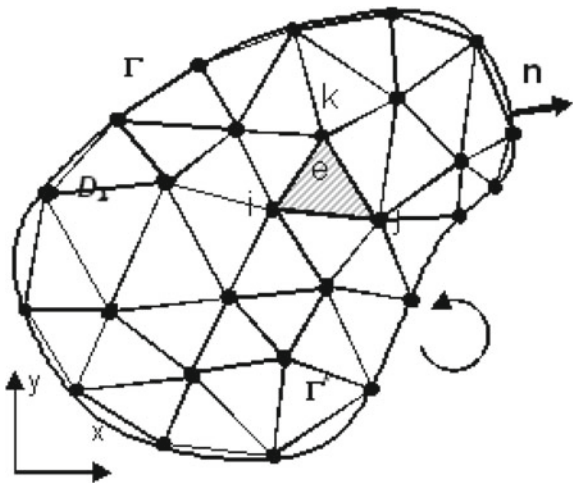


Fig. 4 Creating the mesh in a plane domain using triangular finite elements

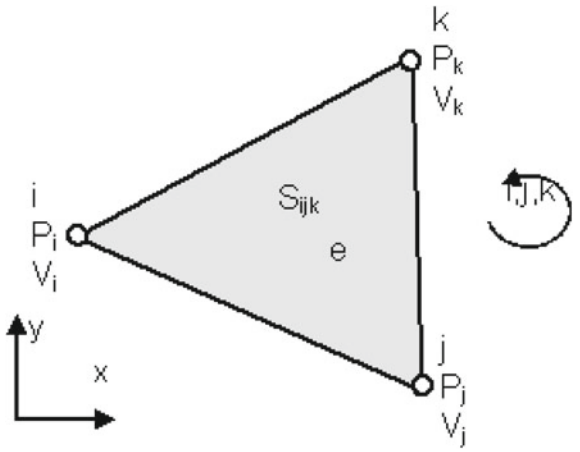


Fig. 5 Triangular finite element

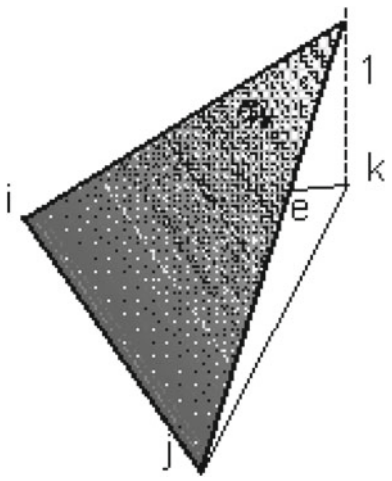


Fig. 6 Linear form function

$$\varphi_i^e(P_k) = a_i x_k + b_i y_k + c_i = 0 \tag{33}$$

The unknowns a_i, b_i, c_i are the solution of the following system of equations:

$$\begin{bmatrix} x_i & y_i & 1 \\ x_j & y_j & 1 \\ x_k & y_k & 1 \end{bmatrix} \begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{34}$$

Using Cramer rule, it results:

$$\begin{cases} a_i = \dfrac{\begin{bmatrix} y_i & 1 \\ y_k & 1 \end{bmatrix}}{\Delta_{ijk}} = \dfrac{y_i - y_k}{\Delta_{ijk}} = \dfrac{y_i - y_k}{2S_{ijk}} \\ b_i = \dfrac{\begin{bmatrix} 1 & x_j \\ 1 & x_k \end{bmatrix}}{\Delta_{ijk}} = \dfrac{x_k - x_j}{\Delta_{ijk}} = \dfrac{x_k - x_j}{2S_{ijk}} \\ c_i = \dfrac{\begin{bmatrix} x_i & y_j \\ x_k & y_k \end{bmatrix}}{\Delta_{ijk}} = \dfrac{x_j y_k - x_k y_j}{\Delta_{ijk}} = \dfrac{x_j y_k - x_k y_j}{2S_{ijk}} \end{cases} \tag{35}$$

where

$$\Delta_{ijk} = \begin{bmatrix} x_i & y_i & 1 \\ x_j & y_j & 1 \\ x_k & y_k & 1 \end{bmatrix} = 2S_{ijk} \tag{36}$$

and S_{ijk} is the area of the triangle *ijk*.

Replacing the expression of the function φ_i for the element (e), results:

$$\varphi_i^e(\text{x}, \text{y}) = \frac{1}{2\text{S}_{\text{ijk}}}\left(\begin{bmatrix} \text{y}_\text{j} & 1 \\ \text{y}_\text{k} & 1 \end{bmatrix}\text{x} + \begin{bmatrix} 1 & \text{x}_\text{j} \\ 1 & x_k \end{bmatrix}\text{y} + \begin{bmatrix} \text{x}_\text{j} & \text{y}_\text{j} \\ \text{x}_\text{k} & y_k \end{bmatrix}\text{x}\right) \tag{37}$$

$$\varphi_i^e(\text{x}, \text{y}) = \frac{1}{2\text{S}_{\text{ijk}}}\big((y_j - y_k)x + (x_k - x_j)y + (x_j y_k - x_k y_j)\big) \tag{38}$$

The above relation has the following replacement:

$$\varphi_i^e(x, y) = \varphi_i^e(P) = \frac{\begin{bmatrix} x & y & 1 \\ x_j & y_j & 1 \\ x_k & y_k & 1 \end{bmatrix}}{2S_{ijk}} = \frac{2S_{Pjk}}{2S_{ijk}} = \frac{S_{Pjk}}{S_{ijk}} \tag{39}$$

where $P(x, y)$ is a point inside the triangular element.

One can express the barycentric coordinates of a point $P \in e$, defined as the ratios between the areas of the triangles consisting of this specific point and each branch and the area of the triangle (Fig. 7).

The three values are not independent, being related by the completeness relation:

$$\lambda_i + \lambda_j + \lambda_k = 1 \tag{40}$$

Each barycentric coordinate associated with a branch has an association with the opposite vertex. It is equal to zero when the point belongs to the branch and is equal to 1 when the point is in the associated vertex. The interior of the triangular element can be graded using parallel lines associated with the three barycentric triangles (Fig. 8).

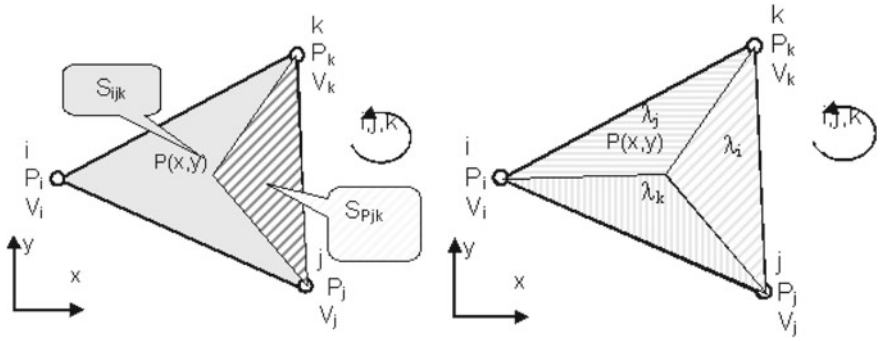


Fig. 7 Barycentric coordinates

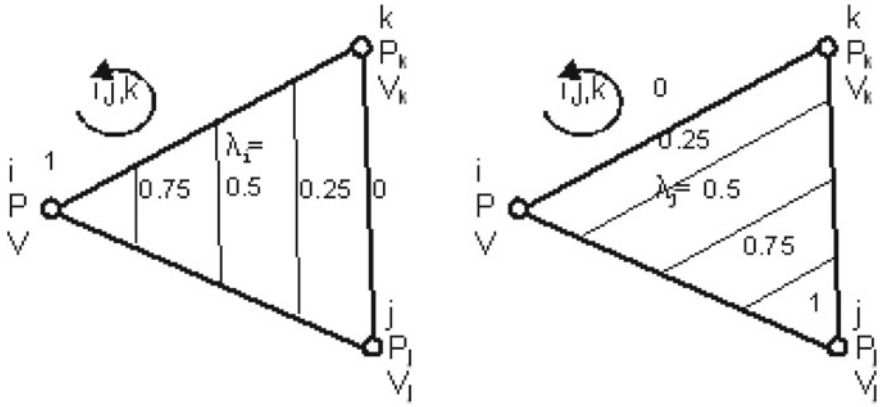


Fig. 8 Barycentric coordinates associated with the elements of the branches

The cases presented above correspond to the triangles with sharp angles, because they are appropriate. The presence of obtuse angles leads to numerical difficulties. The scalar function defined inside of the element has an approximation determined by its values in the element's nodes, using the relation:

$$V^e(r) = \sum\nolimits_{m=i,j,k\in e} V_m \phi_m^e(r);\ r \in e; \tag{41}$$

Another approximation of the function inside the element uses the barycentric coordinates:

$$V^e(r) = \sum\nolimits_{m=i,j,k\in e} V_m \lambda_m^e(x, y);\ r \in e \tag{42}$$

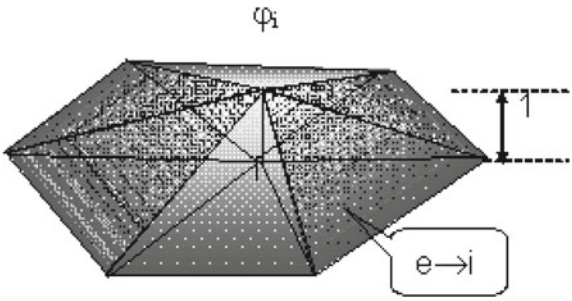


Fig. 9 Linear global approximation function attached to a node

b. *Global approximation*

The assembling of local approximations (for each finite element) represents a preliminary condition for obtaining a global approximation, for the entire domain. The global approximation function attached to one node is the reunion of the local interpolation functions for all the elements which has in common the respective node [23]. Consequently, the function is equal to 1 in the node to which is attached and equal to zero in the remaining nodes. For linear interpolation functions, the global interpolation function attached to a node is:

$$\varphi_i(x, y) = \sum_{e \to i} \varphi_i^e(x, y) = \sum_{e \to i} \lambda_i^e(x, y); i \in I \tag{43}$$

where $\sum_{e \to i} \ldots$ represents the sum for all elements containing the node i; I—the mesh's set of the nodes. Such a function has a pyramidal shape, having as base the reunion of the triangles of the component functions (Fig. 9):

$$\text{supp}\, \varphi_i = \bigcup_{e \to i} \text{supp}\, \varphi_i^e; i \in I \tag{44}$$

The global approximation of a function $V(x, y)$ defined on the given domain and which has the values V_i in the network's nodes will be:

$$V(x, y) \approx \sum V_i \varphi_i(x, y); i \in I \tag{45}$$

3.3.2 3D Finite Elements

The presented principles apply for the 3D case, as well. The simple solution consists of discretizing the domain (building the mesh) in elements having the shape of a tetrahedron. The neighboring tetrahedrons have a common face, and inside each

tetrahedron, the interpolation function is linear. The presentation of the case is the following section.

3.4 Galerkin Method Using the Scalar Magnetic Potential

3.4.1 Parallel Plane Field Computation

For simplicity, we consider only the case of parallel plane field in which the unknown—the scalar magnetic potential V depends only on the x and y coordinates [21, 23]. Consequently, the computation domain is represented by the transversal section S_Γ bounded by the curve Γ (Fig. 10).

The boundary conditions:

– Dirichlet conditions:

$$V(P) = f(P);\ P \in C_D \tag{46}$$

– Neumann conditions:

$$\frac{dV(P)}{dn} = g(p),\ P \in C_N \tag{47}$$

In this case, the weak form of the equations expressed using the scalar potential, becomes:

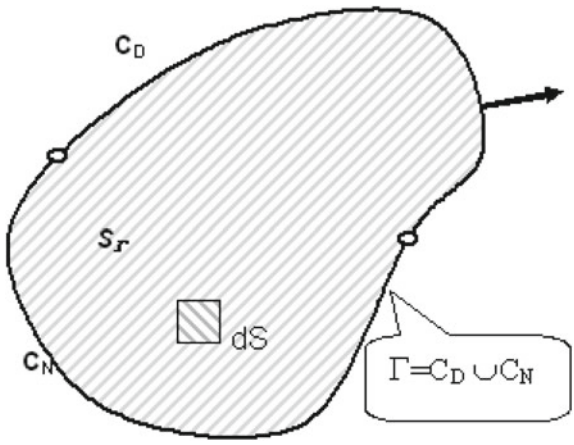


Fig. 10 The computation domain for the 2D problem

$$\int_{S_\Gamma} \mu \, \text{gradU} \, \text{grad} \, \text{VdS} = \int_{S_\Gamma} \rho \text{UdS} + \int_{C_N} \mu g \text{Udl} \tag{48}$$

Approximation using finite elements

We adopt a discretization using triangular elements and linear form functions. So, the approximation of the vector potential V is the relation (3.54). At the limit, the condition becomes:

$$V(P)|_i - V_i - f(P), P \in C_D \Rightarrow i \in I_D \tag{49}$$

$$\left.\frac{dV(P)}{dn}\right|_i = g(P), P \in C_N \Rightarrow i \in I \backslash I_D \tag{50}$$

where I_D are the nodes fulfilling Dirichlet conditions. According to the principle of Galerkin method, the trail (form) functions belong to the same class.

It results that:

$$\int_{S_\Gamma} \mu \, \text{grad} \left(\sum_{i \in I} V_i \varphi_i \right) \text{grad} \varphi_j \text{dS} = \int_{S_\Gamma} \rho \varphi_\text{j} \text{dS} + \int_{C_N} \mu g \varphi_j \text{dl}; \forall j \in I \tag{51}$$

or, condensed:

$$\sum_{i \in I} a_{ji} V_i = b_j + c_j; \forall j \in I; [a][V] = [d]; [d] = [b] + [c] \tag{52}$$

where

$$\begin{cases} a_{ij} = \int_{S_\Gamma} \mu \, \text{grad} \varphi_i \text{grad} \varphi_j \text{dS}; i, j \in I \\ b_j = \int_{S_\Gamma} \rho \varphi_j \text{dS}; j \in I \\ c_j = \int_{C_N} \mu g \varphi_j \text{dl}; j \in I_N \end{cases} \tag{53}$$

Because the nodes $i \in I_D$ in which there are given the Dirichlet conditions and the values of V are the known, the corresponding terms from the equations system appears put on the right side:

$$\sum_{i \in I \backslash I_D} a_{ji} V_i = b_j + c_j - \sum_{i \in I_D} a_{ji} V_i; \forall j \in I \backslash I_D \tag{54}$$

$$[a][V] = [d]; d_j = b_j + c_j - \sum_{i \in I_D} a_{ij} f_i; j \in I \backslash I_D \tag{55}$$

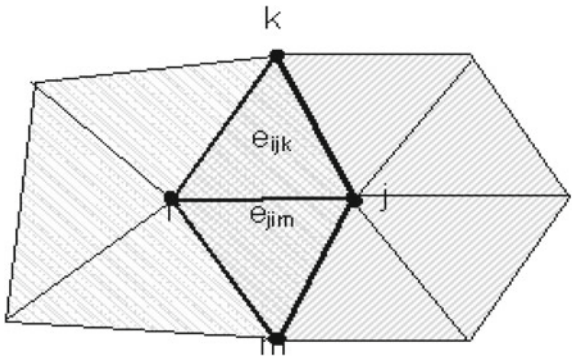


Fig. 11 The integration domain for computing the coefficients

The coefficients of the equations system

– The a_{ji} coefficients.

The interpolation functions φ_i and φ_j have as a support the reunion of the finite elements which has in common the nodes i, respectively j, according to relation (3.53). If we denote the support of the interpolation function φ_i with S_i, the integral from the expression of the coefficient a_{ji}, must be computed only on the intersection of the support functions φ_i and φ_j, $S_i \cup S_j$:

$$a_{ji} = \int\limits_{S_i \cup S_j} \mu \, \mathrm{grad}\varphi_i \mathrm{grad}\varphi_j \mathrm{dS}; \tag{56}$$

These coefficients display the property of symmetry: $a_{ji} = a_{ij}$ $(j \neq i)$

The only elements which intervene are those attached to the nodes i and j, connected by a segment (Fig. 11).

The form functions corresponding to i and j from the first of the two elements ($e_{\text{-ijk}}$) which has in common the branch ij are:

$$\begin{cases} \varphi_i^{e_{ijk}}(x, y) = \frac{1}{\Delta_{ijk}}\big((y_j - y_k)x + (x_k - x_j)y + (x_j y_k - x_k y_j)\big) \\ \varphi_j^{e_{ijk}}(x, y) = \frac{1}{\Delta_{ijk}}((y_k - y_i)x + (x_i - x_k)y + (x_k y_i - x_i y_k)) \end{cases} \tag{57}$$

It results:

$$\begin{cases} \mathrm{grad}\varphi_i^{e_{ijk}}(x, y) = \frac{1}{\Delta_{ijk}}\big((y_j - y_k)\boldsymbol{i} + (x_k - x_j)\boldsymbol{j}\big) \\ \mathrm{grad}\varphi_j^{e_{ijk}}(x, y) = \frac{1}{\Delta_{ijk}}((y_k - y_i)\boldsymbol{i} + (x_i - x_k)\boldsymbol{j}) \end{cases} \tag{58}$$

Similarly, for the second element bounded by the ij branch (e_{jim} element), the system of equations becomes valid:

$$\begin{cases} \text{grad}\varphi_i^{e_{jim}}(x, y) = \frac{1}{\Delta_{jim}}\left(\left(y_m - y_j\right)\boldsymbol{i} + \left(x_j - x_m\right)\boldsymbol{j}\right) \\ \text{grad}\varphi_j^{e_{jim}}(x, y) = \frac{1}{\Delta_{jim}}\left(\left(y_i - y_m\right)\boldsymbol{i} + \left(x_m - x_i\right)\boldsymbol{j}\right) \end{cases} \tag{59}$$

Replacing these expressions, it results:

$$a_{ji} = \int\limits_{e_{ijk}} \mu\, \text{grad}\varphi_i^{e_{ijk}} \text{grad}\varphi_j^{e_{ijk}} \text{dS} + \int\limits_{e_{jim}} \mu\, \text{grad}\varphi_i^{e_{jim}} \text{grad}\varphi_j^{e_{jim}} \text{dS}; i, j \in I, j \neq i \tag{60}$$

respectively:

$$a_{ji} = \int\limits_{e_{ijk}} \left[\left(y_j - y_k\right)\boldsymbol{i} + \left(x_k - x_j\right)j\right]\left[\left(y_k - y_i\right)\boldsymbol{i} + \left(x_i - x_k\right)j\right]\frac{\mu dS}{\left(\Delta_{ijk}\right)^2} + \int\limits_{e_{ijk}} \left[\left(y_m - y_j\right)\boldsymbol{i} + \left(x_j - x_m\right)j\right]\left[\left(y_i - y_m\right)\boldsymbol{i} + \left(x_m - x_i\right)j\right]\frac{\mu dS}{\left(\Delta_{jim}\right)^2}; j \neq i \tag{61}$$

$$a_{ji} = \frac{\left(y_j - y_k\right)\left(y_k - y_i\right) + \left(x_k - x_j\right)\left(x_i - x_k\right)}{\left(\Delta_{ijk}\right)^2} \int\limits_{e_{ijk}} \mu dS + \frac{\left(y_m - y_j\right)\left(y_i - y_m\right) + \left(x_i - x_m\right)\left(x_m - x_i\right)}{\left(\Delta_{jim}\right)^2} \int\limits_{e_{jim}} \mu dS; j \neq i \tag{62}$$

The computation of the integral must target the form function φ_{jj} attached to node j, represented by the reunion of all finite elements which have in common the node j and the interpolation functions φ_i and φ_j.

$$a_{jj} = \int\limits_{S_j} \mu\, \text{grad}^2 \varphi_j dS; j \in I \tag{63}$$

It results:

$$a_{jj} = \sum_{e \to j} \frac{\left(y_j - y_k\right)^2 + \left(x_k - x_j\right)^2}{\left(\Delta_e\right)^2} \int\limits_e \mu dS \tag{64}$$

To compute the integrals, one adopted the hypothesis referring to the magnetic permeability repartition in the mesh elements. The simplest approximation consists of considering a constant permeability μ_e for the mesh elements (e).

For this approximation, we've obtained the following coefficients of the system:

$$a_{ji} = \frac{(y_j - y_k)(y_k - y_i) + (x_k - x_j)(x_i - x_k)}{2(\Delta_{ijk})^2}\mu_{e_{ijk}} + \frac{(y_m - y_j)(y_i - y_m) + (x_j - x_m)(x_m - x_i)}{2(\Delta_{jim})^2}\mu_{e_{jim}};\ j \neq i \tag{65}$$

$$a_{jj} = \sum_{e \to j} \frac{(y_j - y_k)^2 + (x_k - x_j)^2}{2(\Delta_e)^2}\mu_e \tag{66}$$

– The b_j coefficients.

In relation (62) the integral computation involves the support of the form function φ_j, which represents the reunion of the finite elements that has in common the node j.

$$b_j = \int_{S_j} \rho\varphi_j dS = \sum_{e \to j} \int_{S_j} \rho\varphi_j^e dS;\ j \in I \tag{67}$$

For simplicity, we've assumed that the function ρ is constant inside the finite elements. It results:

$$b_j = \sum_{e \to j} \int_{S_j} \rho\varphi_j^e dS = \sum_{e \to j} \rho_e \int_{S_j} \varphi_j^e dS = \sum_{e \to j} \rho_e \int_{S_j} \lambda_j^e dS;\ j \in I \tag{68}$$

The computation of integrals requires the use of the Holland-Bell formulae [14]:

$$\int_{S_{ijk}} \lambda_i^\alpha \lambda_j^\beta \lambda_k^\gamma dS = 2S_{ijk}\frac{\alpha!\beta!\gamma!}{(\alpha+\beta+\gamma+2)!} = \Delta_{ijk}\frac{\alpha!\beta!\gamma!}{(\alpha+\beta+\gamma+2)!} \tag{69}$$

This case corresponds to $\alpha = \chi = 0$, $\beta = 1$, and the result is:

$$\int_{S_{ijk}} \lambda_j dS = \frac{\Delta_{ijk}}{6} \tag{70}$$

Furthermore:

$$b_j = \sum_{e \to j} \frac{\rho_e \Delta_e}{6};\ j \in I \tag{71}$$

– The c_j coefficients.

In relation (62), the integral computation involves the support of the form function φ_j corresponding to a point placed on the boundary with Neumann conditions, which

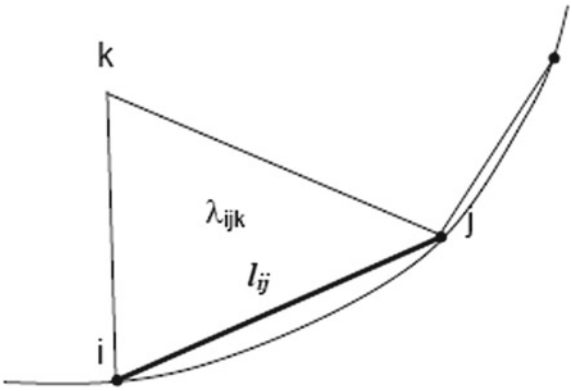


Fig. 12 The integration contour for computing c_j coefficients

represents the reunion of the finite elements which have in common the node j and which have one of the branches on the C_N boundary.

$$c_i = \int_{C_N} \mu g \varphi_j dl = \sum_{e \to j \in I_N} \int_{C_N} \mu g \varphi_j dl; \; j \in I_N \tag{72}$$

If, for simplicity, one admits that the magnetic permeability μ and the Neumann boundary condition (g) have constant values inside the involved finite elements, it results:

$$c_j = \sum_{e \to j \in I_N} \mu g \int_{ij} \varphi_j dl = \sum_{e \to j \in I_N} \mu_e g_e \int_{ij} \lambda_j dl = \sum_{e \to j \in I_N} \mu_e g_e \int_{\substack{ij \\ i,j \in I_N}} \lambda_j dl; \; i, j \in I_N \tag{73}$$

Using one of the Holland-Bell formulae [14], results (Fig. 12):

$$\int_{ij} \lambda_i^{\alpha} \lambda_j^{\beta} dS = L_{ij} \frac{\alpha! \beta!}{(\alpha + \beta + 1)!} \tag{74}$$

for $\alpha = 0, \beta = 1$:

$$\int_{ij} \lambda_j dl = \frac{L_{ij}}{2} = \frac{\sqrt{(x_i + x_j)^2 + (y_i + y_j)^2}}{2} \tag{75}$$

Consequently:

$$c_j = \sum_{\substack{e \to j \in I_N \\ i \in i_N}} \mu_e g_e \frac{L_{ij}}{2} = \sum_{\substack{e \to j \in I_N \\ i \in i_N}} \mu_e g_e \frac{\sqrt{(x_i + x_j)^2 + (y_i + y_j)^2}}{2}; i, j \in I_N \quad (76)$$

Building the equations system

The matrix of the equations systems has the following properties:

- It is symmetrical $a_{ji} = a_{ij}$.
- It is sparse (the number of non-zero elements is relatively small).
- It is positively-defined.
- Using a convenient labeling of the nodes, the matrix can have a band structure; the wider the band is, the more advantageous is.

These particularities are exploited by using adequate equation system computation techniques. This is a common subject for all numerical methods.

3.5 Galerkin Method Using the Magnetic Potential Vector

3.5.1 Parallel-Plane Field Computation

Let us consider the case of parallel-plane fields, in which the vector potential has an orientation along the *oz* axis (Fig. 13).

The boundary conditions are, for 3D problems, those given in paragraph 3.2.1.2:

For parallel-plane fields, these conditions become:

$$A(x, y) = f(P), P \in C_D \quad (77)$$

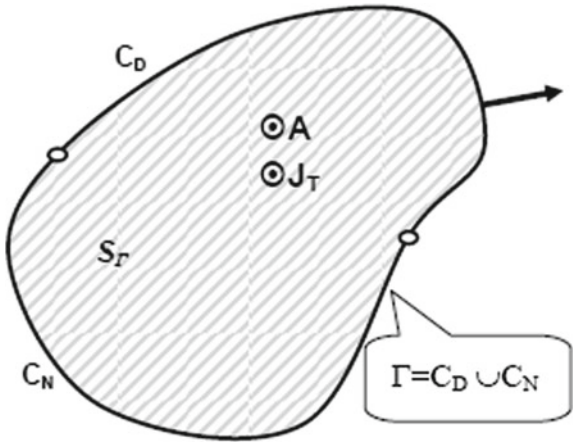


Fig. 13 The computation domain for 3D problem

$$-\frac{1}{\mu}\frac{dA}{dn}\bigg|_{C_N} = g(P),\ P \in C_N \tag{78}$$

The weak forms of the field equations are given by the relation (79):

$$\int_D \frac{1}{\mu} rot\boldsymbol{U} rot\mathrm{A}\delta\Omega = \int_D \boldsymbol{U}\mathrm{J}_T\delta\Omega - \int_D h\mathrm{U}dS \tag{79}$$

For parallel-plane magnetic fields:

$$\boldsymbol{A} = \boldsymbol{k}A(x, y) \rightarrow \mathrm{rot}\boldsymbol{A} = -\boldsymbol{k} \times \nabla A \tag{80}$$

$$\boldsymbol{U} = \boldsymbol{k}\mathrm{U}(x, y) \rightarrow \mathrm{rotU} = -\boldsymbol{k} \times \nabla\mathrm{U} \tag{81}$$

$$\mathrm{J}_T = k\mathrm{J}_T(x, y) \tag{82}$$

From these relations results:

$$\mathrm{rot}\boldsymbol{U}\mathrm{rot}A = (-\mathrm{k} \times \nabla\mathrm{A})(-\mathrm{k} \times \nabla\mathrm{U}) = \nabla\mathrm{A}\nabla\mathrm{U} \tag{83}$$

$$\boldsymbol{hU} = \left(n \times \frac{rot\mathrm{A}}{\mu}\right)\boldsymbol{U} = \left(n \times \frac{-\boldsymbol{k} \times \mathrm{A}}{\mu}\right)\boldsymbol{k}\mathrm{U} = -\frac{1}{\mu}\frac{d\mathrm{A}}{dn}\mathrm{U} \tag{84}$$

The weak form becomes:

$$\int_{S_C} \frac{1}{\mu}\,\mathrm{grad}\,\mathrm{U}\,\mathrm{grad}\mathrm{A}dS = \int_{S_C} \mathrm{U}\mathrm{J}_T dS + \int_{C_N} \frac{1}{\mu}\frac{d\mathrm{A}}{dn}\mathrm{U}dl \tag{85}$$

The resulting expression is similar to the one used for solving the problem using the scalar potential. The necessary transformations are obvious. Therefore, all the above-presented considerations are valid for the present case.

4 Numerical Example

The subject of the numerical modeling is a three-phase squirrel-cage induction motor with the following ratings: nominal power $P_n = 7.5$ kW synchronous speed $n_1 = 3000$ rot/min (one pole pair), nominal voltage $V_n = 380$ V, nominal frequency $f_1 = 50$ Hz, Y-connected. The copper made distributed stator winding has one layer, whereas, the squirrel cage is molded from aluminum [27, 28].

Starting with Maxwell's equations, one can conclude the mathematical model of the induction motor in the quasi-stationary harmonic magnetic regime as [25–27]:

(a) In the stator winding area:

$$\mathrm{rot}\left[\upsilon(\boldsymbol{B})\cdot\mathrm{rot}\underline{\boldsymbol{A}}\right]=\underline{\boldsymbol{J}} \tag{86}$$

(b) In the areas of the rotor bars:

$$\mathrm{rot}\left[\upsilon(\boldsymbol{B})\cdot\mathrm{rot}\underline{\boldsymbol{A}}\right]=-\sigma\left(j\omega\underline{A}+\mathrm{grad}\underline{V}\right) \tag{87}$$

(c) In the air gap and core ferromagnetic areas:

$$\mathrm{rot}\left[\upsilon(\boldsymbol{B})\cdot\mathrm{rot}\underline{\boldsymbol{A}}\right]=0 \tag{88}$$

One can express the relationship between the vector potential and the scalar potential as (89):

$$\mathrm{div}\left[-\sigma\left(j\omega\underline{\boldsymbol{A}}+\mathrm{grad}\underline{\boldsymbol{V}}\right)\right]=0 \tag{89}$$

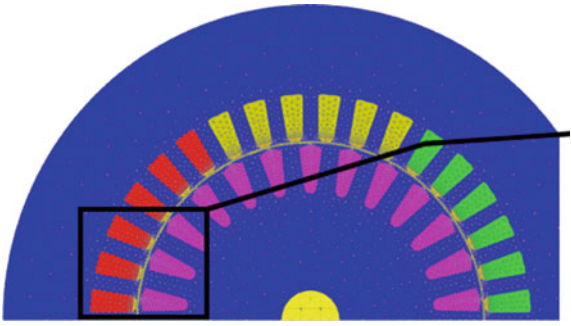
The application of Coulomb calibration condition turns the system of Eqs. (86)–(89) the "AV" formulation (90), (91):

$$\mathrm{rot}\left[\upsilon\cdot\mathrm{rot}\underline{\boldsymbol{A}}\right]-\mathrm{grad}\left(\upsilon\mathrm{div}\underline{\boldsymbol{A}}_p\right)+\sigma\left(j\omega\underline{\boldsymbol{A}}+\mathrm{grad}\underline{V}\right)=0 \tag{90}$$

$$\mathrm{div}\left[\sigma\left(j\omega\underline{\boldsymbol{A}}+\mathrm{grad}\underline{\boldsymbol{V}}\right)\right]=0 \tag{91}$$

The axial symmetry of the induction motor cross-sectional area makes the 2D FEM appropriate. Circular symmetry gives the possibility to reduce the area of the cross-section to an area covered by a pole-pitch, as in Figs. 14 and 15. The borders of the domain are the axis of the cross-sectional area, the stator's external surface, respectively, the internal surface of the rotor's magnetic circuit [28].

Fig. 14 Calculation domain



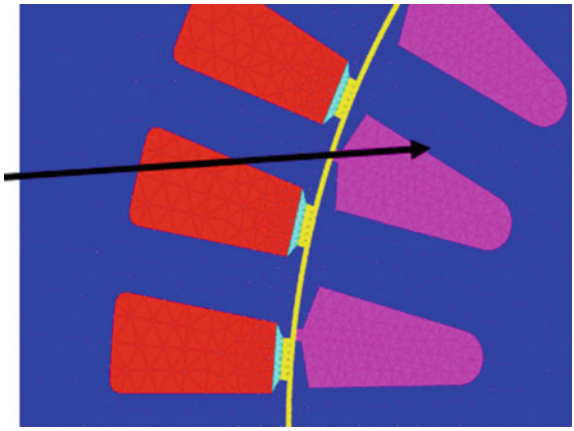


Fig. 15 Discretization detail

There are subdomains displaying distinct physical properties as following: following

- STATOR (blue), nonconductive magnetic region.
- STATOR TOOTH (blue), nonconductive magnetic region.
- ROTOR (blue), nonconductive magnetic region.
- ROTOR TOOTH (blue), nonconductive magnetic region.
- AIR GAP(yellow), nonconductive nonmagnetic region.
- INSULATION (cyan), non-conductive nonmagnetic region.
- PHASE1 (in red), PHASE2 (in green), PHASE3 (in yellow), nonmagnetic nonconductive regions made of fascicular conductors placed into the stator's slots: in this way they are not subjected to eddy currents.
- BAR1, BAR2, …, BAR14 (magenta), regions of massive conductor type nonmagnetic placed inside of the 14 slots of the rotor included in the calculation domain.

Triangular finite elements participated in the domain meshing process. The air gap requires a discretization of higher elements density due to a higher magnetic field variation: discretization starts from the air gap towards exterior.

In the quasi-stationary regime of the magnetic type, the rotor has a fixed position relative to the stator. The sources for the magnetic field are in such a case, either the densities of the currents circulating in the stator winding. If the motor operates energized from a three-phase voltage supply, one must solve a circuit-field coupled problem. Following the definition and discretization of the calculation domain (mesh build-up), the allocation of the physical properties to the regions is the next step. For example, the magnetic core is characterized by the magnetization curve $B = f(H)$ as in Fig. 16. Moreover, the boundary conditions impose a vector potential null at the exterior of the stator magnetic core and anti-symmetry properties for the boundary sections defined by the two axes of symmetry.

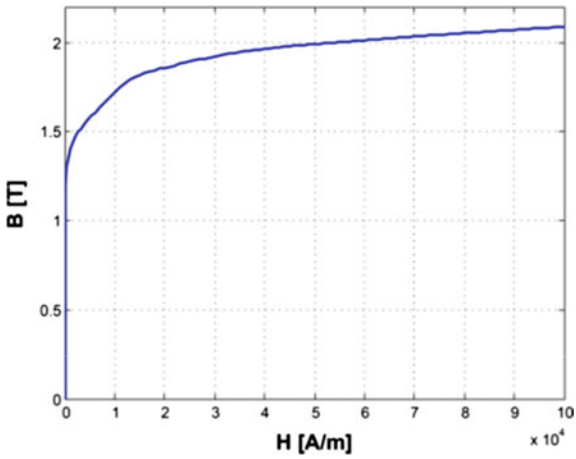


Fig. 16 Magnetization curve of the stator and rotor magnetic core

The magneto-harmonic model of the induction machine runs for constant slip, respectively, constant rotor speed. For this type of simulations, there is no relative movement between the stator and the rotor. Due to the presence of the stator and rotor slots, simulation results depend on their positions. There is a requirement to determine that particular position of the rotor for which the electromagnetic torque is equal to average electromagnetic torque for a cycle of variation, whereas the rotor relative position to the stator changes. There, such a position requires a rotor rotation with two degrees. The simulations of interest point out to the full load and to the starting as well (Figs. 17, 18 and 19). From Fig. 16, one can visualize the map of the

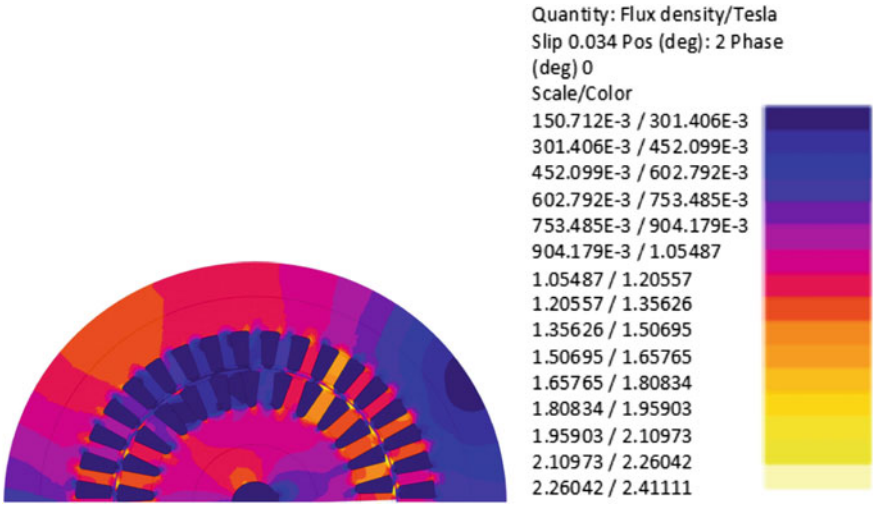


Fig. 17 The magnetic field induction distribution at full-load

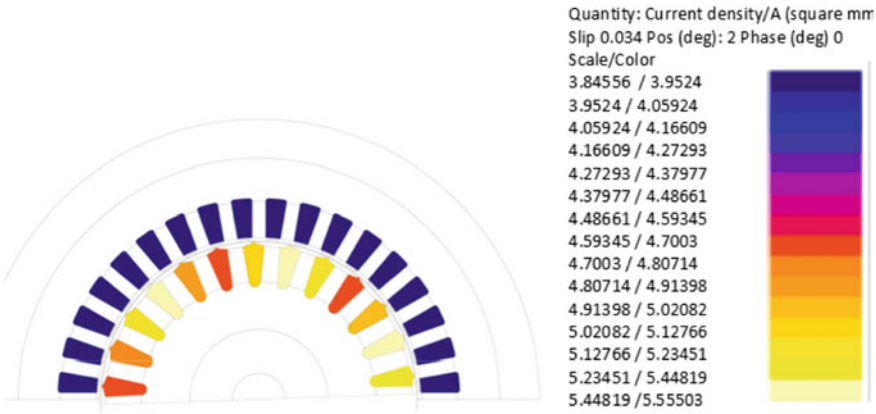


Fig. 18 The current density distribution in the stator and rotor windings

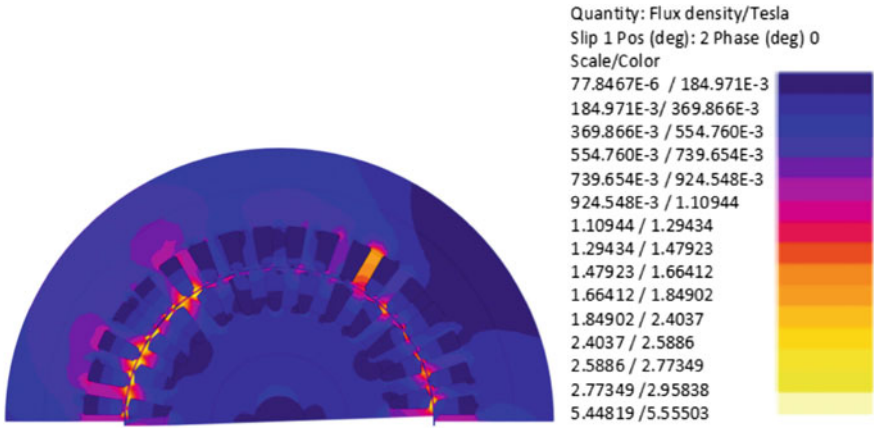


Fig. 19 The magnetic field induction distribution at starting

magnetic induction in the cross-sectional area at full load. The stator teeth adjacent to the slots containing phase coils carrying the highest instantaneous current, suffer the highest magnetic stress with the magnetic induction, $B = 1.8$ T. Similarly, in the rotor region teeth adjacent to bars carrying the highest instantaneous current carry a magnetic field having the induction $B = 2.4$ T. The current density in the rotor bars varies between 4.7 and 5.5 A/mm^2. At starting, the magnetic induction is $B = 2.6$ T in the stator's teeth, 1 T in the stator's yoke, whereas the current density in the rotor bars varies 4.7 and 10 A/mm^2.

5 Conclusions

The necessity of improving the performances of electrotechnical and electronic components leads to the development of the electromagnetic field computation methods. These problems imply solving some complicated equation systems containing partial derivatives. The manner of computation of such systems involves many iterations.

In engineering practices, electromagnetic field problems are of two main types: analysis problems, which consist of determining the corresponding magnetic field to some given uniqueness conditions and synthesis problems, which consist in determining the configurations capable of producing an electromagnetic field holding some characteristics.

The numerical methods seek to determine the solution in a discrete set of points in the studied domain. Although apparently, the numerical methods can approach the most real problems of determining some electromagnetic fields, the reality is much more "modest." Even if some calculus systems get more and more capable and more and more approachable, numerous limitations persist in this domain. Despite the abundance of performant (and generally expensive!) software packages commercially available, addressing the problem of numerical computation of electromagnetic field, there is a strong demand for proper knowledge of electromagnetism when it comes to modeling. Nowadays, having access on a large scale to performant computers allowed numerical methods to gain more steam.

Firstly, scientists developed and applied the Finite Difference Method (FDM), which originated from the expansion of the Taylor series. FDM has its well-known drawback when it comes to express the Neumann boundary conditions at interfaces. The answer to such a problem was the development of the Finite Element Method (FEM), which conveniently addressed some limitations of the FDM, particularly the "stiffness" in the construction of the meshing scheme.

This chapter provided an insight into the theoretical approaches of the FEM. The construction of the finite element approximation relies on partial differential equations as expressions of the solutions defined by a partition of the field study in disjoint elements, called "finite elements" giving the name of the method. Although such an approach looks like the FDM, the procedure used in FEM is different.

In terms of a concrete approach, FEM has two main variants: the Galerkin method and the Ritz method. The distinction between these approaches is not always obvious. For this reason, the terminology may show some differences (i.e., the Ritz method is sometimes called the Ritz-Galerkin). The sections of this chapter contain a detailed presentation of these methods.

The so-called weak formulation of the electromagnetic field equations characterized the Galerkin method; therefore, it is presented the weak form using the scalar electric potential and weak form using magnetic vector potential, followed by the principle of the Galerkin method.

For a better understanding, each section comes with clear examples. In such examples, the FEM turns to be extremely flexible and easy to customize, pending

upon the type of problem in place for analysis, like electrostatics, magnetostatics, AC analysis, transient analysis, and geometry, 2D or 3D.

References

1. Virjoghie EO, Enescu D, Ionel M, Stan MF (2010) 3D finite element analysis for arcing chamber optimization of the current-limiting circuit breaker. WSEAS Trans Power Syst 1(5):1–10
2. Ramuhalli P, Udpa L, Udpa SS (2005) Finite-element neural networks for solving differential equations. IEEE Trans Neural Netw 6(6):1381–1392
3. Zhan Y, Knight A (2008) Parallel time-stepped analysis of induction machines with Newton-Raphson iteration and domain decomposition. IEEE Trans Magn 44(6):1546–1549
4. Yoon H, Kim I, Shin PS, Koh CS (2011) Finite element implementation of a generalized chua-type vector hysteresis model and application to loss analysis of three-phase transformer. IEEE Trans Magn 47(5):1122–1125
5. Niu S, Ho L, Fu WN, Zhu J (2102) A convenient mesh rotation method of finite element analysis using sub-matrix transformation approach. IEEE Trans Magn 48(2):303–306
6. Sibué JR, Meunier G, Ferrieux JP, Roudet J, Périot R (2013) Modeling and computation of losses in conductors and magnetic cores of a large air. IEEE Trans Magn 49(1):586–589
7. Bastos JPA, Sadowski N (2003) Electromagnetic modeling by finite element methods. Marcel Dekker Inc
8. Radulet R (1975) Bazele teoretice ale electrotehnicii (in English: Fundamentals of Electrical Engineering) Litografia Învăţământului. Bucharest, Romania
9. Timotin A, Hortopan V, Ifrim A, Preda M (1970) Lecţii de bazele Electrotehnicii (Lectures of Fundamentals in Electrical Engineering) Editura Didactică şi Pedagogică. Bucharest, Romania
10. Mocanu CI (1981) Teoria câmpului electromagnetic (In English: Electromagnetic Field theory) Editura Didactică şi Pedagogică. Bucharest, Romania
11. Moraru A (2002) Bazele electrotehnicii - Teoria câmpului Electromagnetic (In English: Fundamentals of Electrical Engineering – Electromagnetic field Theory). MatrixRom, Bucharest, Romania
12. Fluerasu C, Fluerasu C (2007) Electromagnetism. Printech, Bucharest, Romania
13. Hantila IF (2004) Electrotehnica teoretică (In English-Theoretical Electrical Engineering). Editura Electra, Bucharest, Romania
14. Hantila IF, Vasiliu M (2005) Câmpul electromagnetic variabil în Timp (In English -Time variable electromagnetic field). Editura Electra, Bucharest, Romania
15. Vasiliu M, Hantila IF (2006) Electromagnetics. Editura Electra, Bucharest, Romania
16. Hantila IF, Vasile N, Demeter E, Marinescu S, Covrig M (1997) Câmpul electromagnetic staţionar în medii nelineare (In English: Nonlinear media stationary electromagnetic field). Editura ICPE, Bucharest, Romania
17. Hantila IF (1975) A method for solving stationary magnetic field in nonlinear media. Revue Roum Sci et Tech – Elth et Energ (3):397–407
18. Hantila IF (1974) Mathematical models of the relation between B and H, for nonlinear media. Revue Roum Sci et Tech Elth et Energ (3):429–448
19. Hantila IF, Preda G, Vasiliu M (2000) Polarization method for static fields. IEEE Trans Magn 36(4):672–675
20. Silvester P, Chari MV (1970) Finite-element solutions of saturable magnetic field problems. IEEE Trans Power App Syst 89(7):1642–1651
21. Moraru A (1968) Finite difference equations for magnetic field in nonlinear, media. Revue Roum Sci et Tech Elth et Energ (1) 3–13
22. Fluerasu C, Fluerasu C (2009) Electromagnetism-applications. Printech, Bucharest, Romania
23. Andrei H et al (2010) Metode numerice, modelări şi simulări cu aplicaţii în ingineria electrică (in English - Numerical Methods, Modeling and Simulation Applied in Electrical Engineering). Editura Electra, Bucharest, Romania

24. Onuki T, Wakao S, Saito H (1994) Improvement in the calculation of electromagnetic force by the FEM. IEEE Trans Magn 30(4):1863–1866
25. Vaananen J (1996) Circuit theoretical approach to couple two-dimensional finite element models with external circuit equations. IEEE Trans Magn 32(2):400–410
26. Chevalier T, Kedous-Lebouc A, Cornut B, Cester C (1999) Estimation of magnetic loss in an induction motor fed with sinusoidal supply using a finite element software and a new approach to dynamic hysteresis. IEEE Trans Magn 35(5):3400–3402
27. Knight AM (2003) Time-stepped eddy-current analysis of induction machines with transmission line modeling and domain decomposition. IEEE Trans Magn 39(4):2030–2035
28. Florin Rezmerita, Contributii privind imbunatatirea performantelor motoarelor asincrone (Contributions to improve performance of induction motors). PhD Thesis, December 2013, PhD supervisor prof. Mihai Iordache

Advanced Numerical Methods Based on Artificial Intelligence

Levente Czumbil, Dan D. Micu, and Andrei Ceclan

Abstract The chapter presents some advanced numerical methods based on Artificial Intelligence (AI) techniques applied to specific electrical engineering problems. A theoretical description is done regarding the basic aspects of the nowadays most commonly used AI techniques: Neural Networks, Fuzzy Logic, and Genetic Algorithms respectively. The presented AI techniques are exemplified through two specific electrical engineering application implemented by the authors in their previous research projects. The first application consist in the identification of the optimal equivalent horizontally layered earth structure by means of a Genetic Algorithm according in site soil resistivity measurements. The second application provides a Neural Network alternative two evaluate the impedance matrix that describes the electromagnetic coupling between overhead powerlines and nearby underground pipelines for different separation distances and various vertically layered soil structures.

Keyword Artificial intelligence · Neural networks · Genetic algorithms · Electrical engineering application

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network

L. Czumbil · D. D. Micu (✉) · A. Ceclan
Numerical Methods Research Center, Department of Electrical Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, Romania
e-mail: dan.micu@ethm.utcluj.ro

L. Czumbil
e-mail: levente.czumbil@ethm.utcluj.ro

A. Ceclan
e-mail: andrei.ceclan@ethm.utcluj.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_4

EA Evolutionary Algorithms
FEM Finite Element Method
FL Fuzzy Logic
FLS Fuzzy Logic Systems
GA Genetic Algorithms
HVPL High Voltage Power Line
MGP Metallic Gas Pipeline
ML Measurement Location
NN Neural Networks
SM Surrogate Models

1 Introduction

The first definition for AI, which is still one of most accepted ones, was given in 1955 by McCarthy: "making a machine behave in ways that would be called intelligent if a human were so behaving" [1].

The AI technics represents a class of heuristic methods for solving the last decade's issues, that were born from the desire of implementing a system with the capacity to mimic the human mind. One of the most fundamental methods is the capacity of learning with or without external help and even with the purpose of permeant improvement. This method is usually used as a quick alternative for the old methods that requires a high effort and a long time of calculus compilation.

The main AI techniques that are used nowadays are Fuzzy Logic (FL), Neural Networks (NN), Genetic Algorithms (GA), Surrogate Models (SM) and Evolutionary Algorithms (EA) [2].

The main purpose of this chapter is to highlight the basic theoretical aspects of the most commonly used AI techniques (Genetic Algorithms, Fuzzy Logic and Neural Networks) and to exemplify how they could be implemented in case of specific electrical engineering applications.

Section 2 makes a brief introduction to the theoretical aspects regarding Genetic Algorithms based optimization techniques, presenting their structure, different chromosomal coding techniques and the basic GA operators. The next section describes how Fuzzy Logic Systems (FLS) work and main implementation steps (fuzzification, FL rule base interface and defuzzification respectively). Section 4 presents the basic theoretical aspects regarding Neural Networks: the structure of an artificial neuron, the main activation functions, the most commonly used NN architectures and training techniques.

The first demonstrative application (Sect. 5) shows how a genetic algorithm could be implemented to determine the optimal equivalent horizontal soil structure based on soil resistivity measurements. The second application (Sect. 6) exemplifies the implementation of Neural Networks in order to determine the inductive coupling

matrix in case of electromagnetic interference problems between overhead power line and nearby metallic pipelines.

2 Genetic Algorithms

Genetic Algorithms are part of the evolutionary computing strategies and represent a series of adaptive heuristic techniques based on the principle of natural selection: "The one who is best suited survives". The idea of evolutionary calculus was introduced in 1960 when several biologists began to use computers to simulate biological systems [3, 4].

Usually, genetic algorithms are used to solve multi-criteria optimization, planning or nonlinear search problems. They constitute a set of adaptive procedures that could find the solutions of a problem through a mechanism of natural selection and genetic recombination/evolution. The mechanism was introduced and analysed by J. Holland [5], being characterized by the fact that only the species (the solutions) that are better adapted to the environment (to the investigated problem) are able to survive and evolve over generations, while the less adapted ones will disappear. The likelihood of a species to survive and evolve over generations becomes greater as the degree of adaptation grows, which in terms of optimization it means that the solution is getting closer to an optimum.

2.1 Structure of a Genetic Algorithm

Genetic Algorithms start from an initial set of solutions, randomly generated or based on prior knowledge, referred as "population" in the literature. In this population, each individual represents a possible solution of the investigated problem and is defined by its "chromosome" structure (its coding). Within the GA the starting population is subjected to an iterative process, exemplified in Fig. 1, through which an optimal

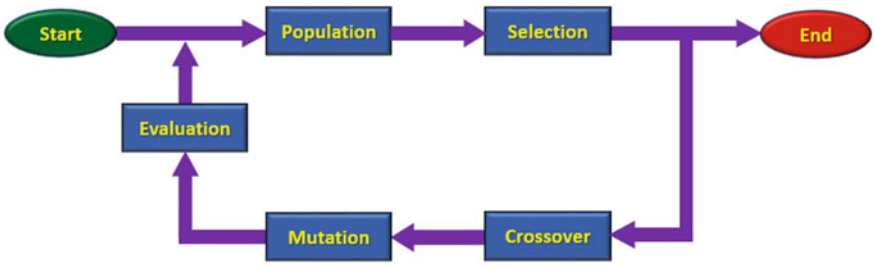


Fig. 1 The structure of a genetic algorithm

solution of the studied problem is determined. An iteration of this optimization/search process is known in literature as "a generation" of the genetic algorithm.

The iterative process that underlies any genetic algorithm can be defined by the following steps:

Step 1: Creation of a set of initial possible solutions ("individuals") that will form the starting population of the investigated problem;
Step 2: Selection based on an objective ("fitness") function of the individuals from the current generation population, that have adapted best to the needs of the problem that has to be solved;
Step 3: The selected individuals are subjected to genetic operators (such as "crossover" and "mutation") to form the population of the next generation;
Step 4: Evaluate the degree to which the members of the new generation correspond more adequately to the requirements of the studied problem;
Step 5: The population of old generation is abandoned, and the iterative process is resumed from Step 2.

Such a cycle is repeated until the best solution of the problem is identified or a predetermined number of generations/iterations has been reached [6, 7].

2.2 *Chromosome Structure of an Individual*

The chromosome structure of an individual defines how a candidate solution of the investigated problem is represented within a genetic algorithm. This constitutes the whole set of "genes", the parameters of an individual that must be determined/optimized for the studied problem. The genes of an individual can be represented either in binary form (Fig. 2a), through a finite string of 0 and 1 values, or in natural form (Fig. 2b) by a real value, generally in the range of 0 to 1.

In order to evaluate how each parameter ("gene"), that has to be optimized, correspond to the requirements of the investigated problem, a cost function, f_c, has to be defined, for each gene, g. The overall performance of an individual (possible solution) regarding the problem in question is determined by the GA objective (or "fitness") function, that is given by the weighted sum of these cost functions, see Eq. (1):

$$F = \frac{1}{n} \cdot \sum_{i=1}^{n} (p_i \cdot f_{Ci}(g_i)) \tag{1}$$

Fig. 2 Chromosome structure of an individual: **a** binary form, **b** natural form

where: n is the total number of parameters, p_i indicates the importance of the gene g_i, $i = 1..n$.

Within the iterative process of optimizing the solution, a minimization or maximization of the fitness function must to be achieved according to the investigated problem [8].

2.3 Selection Operator

The selection operator plays an important role in a genetic algorithm. This operator decides which of the individuals of a population will be able to participate in the formation of the next generation population. The purpose of the selection is to ensure more chances of "survival" / "reproduction" for the best performing individuals in a given population. The selection aims to maximize the performance of individuals (possible solutions to the problem in question).

2.4 Crossover Operator

The crossover operator is the most important operator in the optimization process. This operator applies to individuals in the current population for the purpose of generating individuals for the next generation, and thus ensuring the convergence of the problem. The mechanism of the crossover operator is highly dependent on the gene coding mode of the chromosome structure. Usually, the crossover operator applies to two parents (possible solutions) from the current population and provides two descendants (new solution configurations) for the next generation population. Descendants obtained through the crossover operation, will have characteristics from both parents. Due to its major importance, several crossing methods have been proposed in the literature [9, 10].

2.5 Mutation Operator

The mutation operator has the role of maintaining the diversity of the search space population by introducing individuals that could not be obtained through other mechanisms. This operator consists in randomly changing the value of a single gene/position in the chromosome structure of an individual. In the case of a binary gene coding, the process of mutation implies the negation of a bit in a gene, while in the case of natural form coding it implies a small variation of the value of a gene, see Eq. (2):

$$genA = genA + \xi \tag{2}$$

where: *genA* is the value of the parameter represented by gene *A*, and ξ is the applied perturbation [11].

Mutation is a probabilistic operator. Considering a population of N individuals, each one having n genes, the probability of a gene to undergo a mutation will be p_m, $m = 1..N\,n$. According to GA implementation these probability values could be equal or not equal for each gene.

There are several ways to apply a mutation operator. One of them would be the *change in formatting*. In this case for each position in the chromosome structure of an individual selected for mutation, a random number, q, is generated in the interval [0,1]. If $q > p_m$, then the mutation operator is executed, for that chromosome position, otherwise the position remains unchanged [10].

3 Fuzzy Logic

Fuzzy Logic (FL) offers an alternative, to classical linear equation-based methods, for dealing with problems that describe system operations. It is used especially when the connections between the input and output data of a system are too complex and cannot be defined exactly, due to a significant level of uncertainty in the analysed problem. In case of Fuzzy Logic Systems (FLS), conventional algorithms are replaced by a set of rules of the form *IF (premise) THEN (conclusion)*. This results in a heuristic algorithm that takes into account operator's experience in describing the investigated system.

The basis of the fuzzy set theory was laid by L.A. Zahed in 1965 [12]. From a mathematical point of view, the object of FL is to make a connection (application) between the set of input data of a system and its output values. This connection is made based on a set of *IF—THEN* type laws or reasoning. All the laws are evaluated in parallel and their order do not affect the outcome values.

Fuzzy Logic Systems work only with linguistic/fuzzy values. Therefore, all the input data must undergo a "fuzzification" process that transforms the actual values into fuzzy sets, and the obtained results has to be subjected to a "defuzzification" process for later use [13, 14], as in Fig. 3 can be seen.

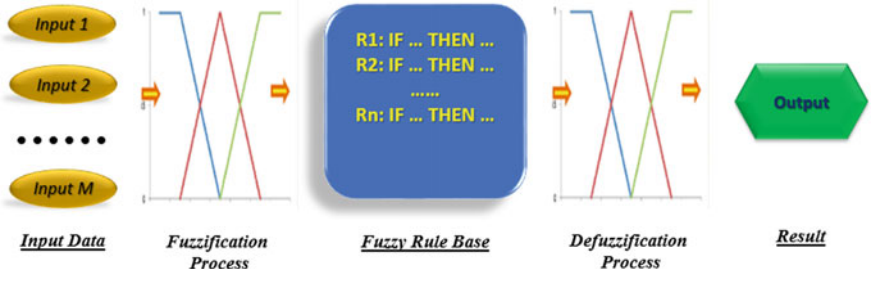


Fig. 3 Operation of a fuzzy logic system

3.1 Fuzzification

Fuzzification is the process by which the actual data provided at the input of a FLS block is transformed into linguistic variables defined by fuzzy sets. The notion of a fuzzy set has been introduced as a generalization of the concept of binary membership of an element to a set. Each fuzzy set is associated with a characteristic/membership function that provides a value in the [0,1] range. This value describes the degree of the belonging of an element to that fuzzy set [15], as in (3) is presented.

$$\mu_A : X \rightarrow [0, 1] \tag{3}$$

A fuzzy set is completely defined by its membership function. Most of the fuzzy sets used practical applications have a membership function defined over the set of real numbers. Therefore, is the most convenient way to express these membership functions as analytical equations [15].

3.2 Inference

The most important component in describing a fuzzy logic system is set of rules (laws) that are applied. The mathematical interpretation of these *IF – THEN* sentences is done through the *inference* process, which has several distinct parts. First, the premises are evaluated, which involves providing the input data and applying the fuzzy membership functions. Then the proper consequence of a fuzzy law is applied to the output values, this operation is known as an *implication.* The premise of a fuzzy rule can have several parts joined by fuzzy operators of "AND" or "OR" type [16], like in (4):

$$\begin{aligned} &\text{IF } x_1 \text{ is } A \text{ AND } x_2 \text{ is } B \text{ THEN } y \text{ is } V \\ &\text{IF } x_1 \text{ is } C \text{ AND } x_2 \text{ is } D \text{ THEN } y \text{ is } W \end{aligned} \tag{4}$$

where: x_1, x_2 are input values; A, B and C, D are fuzzy sets for input data x_1 and x_2 respectively; y is the FLS output and V, W are fuzzy sets corresponding to y.

Each part of the premise is evaluated separately, assigning a specific value to the fuzzy operators. The way in which these "AND" / "OR" operators are mathematically interpreted depends on the inference method adopted. The most commonly used inference methods in the literature are the Max–Min, the Max-Product and the Sum-Product respectively [15].

3.3 *Defuzzification*

The result of the inference are fuzzy sets attached to the FLS output values. In order to turn these fuzzy sets into real values, they must undergo a defuzzification process. The task of this operation is to determine that unique value from a given range belonging to each output data that best fits the resulting fuzzy sets.

Among the most common methods of defuzzification, in the literature, there are the centre of gravity method, the centre of sums method and the height method respectively [17].

4 Neural Networks

The most complex neural network in nature is the human brain, this inspired scientist to try to mimic it by designing Artificial Neural Networks (ANN). As in nature, ANN are constructed from smaller building blocks called neurons. The first attempt to schematically represent the mathematic model of an artificial neuron was made in the early 1940s by McCulloch and Pits [18].

As Fig. 4 shows, the architecture of an artificial neuron follows the structures of the biological neuron, being a system with variable number of m input data x_k, $k = 1..m$, and a single output value y. The m input values of an artificial neuron are multiplied by coefficients w_k, called the *weights*, and then summed together. The value thus obtained is added to a parameter b called *bias* value. The final sum, denoted by h, is applied as argument to the transfer function of the artificial neuron. This function is also known as activation function, f_a, in the literature and can have various mathematical implementations [19–21].

Thus, the output of an artificial neuron is generally described by Eq. (5):

$$y = f_a(h) \tag{5}$$

where:

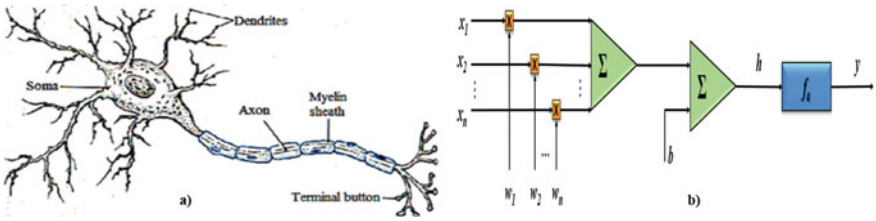


Fig. 4 Structure of a biological (**a**) and artificial (**b**) neurons [22]

$$h = \sum_{k=1}^{m} (x_k \cdot w_k) + b \tag{6}$$

The weights and the bias of an artificial neuron are adjustable parameters, and their values are determined during the neural network training process, in order to obtain the desired network behaviour. Therefore, when using a neuron, the output depends only on the set of input data and the used activation function.

4.1 Activation Functions

The activation function of a neuron is generally a bounded and monotony increasing function, as in Eq. (7), with values between 0 and 1 or between −1 and 1:

$$|f_a(h)| \leq M, \; M \in (0, +\infty), f_a'(h) > 0 \tag{7}$$

Each neuron of an artificial neural network can have its own activation function, however, usually, the same activation function is used for all neurons that form a layer. If back-propagation error technique is used to train the neural network, then it is necessary to know the first derivative of applied activation/transfer functions. In most application, for the output layer of neurons a linear transfer function is used while for the intermediate (hidden) layer neurons sigmoid type transfer functions are implemented.

4.2 Neuronal Networks Architecture

The output of a neural network is highly influenced by its architecture, how the neurons that form it are interconnected. As a NN architecture we understand the structure, more precisely the number of layers, the activation functions and the number of neurons used on each layer. A layer of neurons is formed by all the neurons that work in parallel with the same input data and have the same destination for their output data. Figure 5 shows the working configuration of a layer of neurons:

The weights of the neurons forming a layer can be grouped and placed in a matrix of weights W, while the bias values could be collected in a vector B [20], see Eq. (8):

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{1,n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{m,1} & w_{m,2} & \cdots & w_{m,n} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{bmatrix} \tag{8}$$

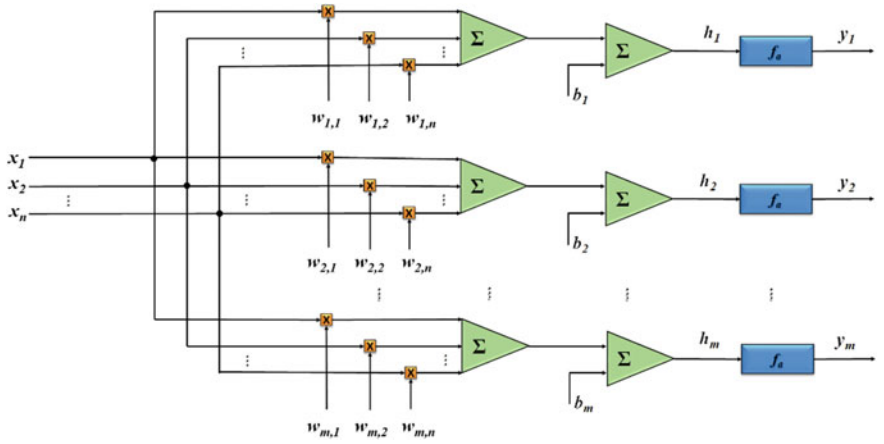


Fig. 5 Structure of a layer of neurons

Usually, the same activation function f_a is used for all the neurons from a specific layer. Therefore, the output values of a layer of neurons could be expressed as:

$$y_i = f_a(h_i),\ i = 1..n \tag{9}$$

with:

$$h_i = \sum_{k=1}^{m} \left(x_k \cdot w_{k,i}\right) + b_i,\ i = 1..n \tag{10}$$

The matrix form of the above equation in given by:

$$[H] = [W]^T \cdot [x] + [B] \tag{11}$$

The architecture of a specific NN could contain one or several layers of neurons. *Output layer* is called the layer of neurons that provides the final output data a neural network. This layer cannot be missing from the structure of any neural network. The neuron layers that interpose between the NN input data and the input values of the output layer are called *hidden layers*. In some literature references the first layer of neuros is called also as the *input layer* [21].

Figure 6 shows a simplified representation a *feed-forward neural network* with one hidden/input layer and one output layer. The following notations are used in Fig. 6: x_k, $k = 1..m$ for the m NN input data values; v_j, $j = 1..r$ for the r output values of the hidden/input layer neurons; y_i, $i = 1..n$ for the n output values of the NN; and $w^1{}_{k,j}$, respectively $w^2{}_{k,j}$ for the weights of the neurons in the two layers of the presented network configuration.

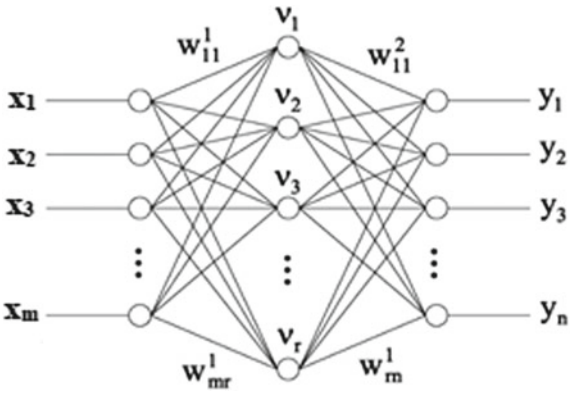


Fig. 6 Simplified diagram of a multi-layer feed-forward neural network [22]

Based on this simplified representation the mathematical form of the output data of a feed-forward neural network, with one hidden layer and one output layer, can be easily deduced according to equations (9–11). Thus, the arguments of the hidden layer neurons activation function are given by:

$$h_j^1 = \sum_{k=1}^{m} \left(x_k \cdot w_{k,j}^1 \right) + b_j^1, \;\; j = 1..r \tag{12}$$

while the output values of the hidden layer neurons are obtained through:

$$v_j = f_a^1 \left(h_j \right) = f_a^1 \left(\sum_{k=1}^{m} \left(x_k \cdot w_{k,j}^1 \right) + b_j^1 \right), \;\; j = 1..r \tag{13}$$

Therefore, the arguments of the output layer activation functions will be given by:

$$h_i^2 = \sum_{j=1}^{r} \left(v_j \cdot w_{j,i}^2 \right) + b_i^2 = \sum_{j=1}^{r} \left(f_a^1 \left(\sum_{k=1}^{m} \left(x_k \cdot w_{k,j}^1 \right) + b_j^1 \right) \cdot w_{j,i}^2 \right) + b_i^2, \;\; i = 1..n \tag{14}$$

Finally, the general NN output data could be evaluated with equation (15):

$$y_i = f_a^2 \left(h_i^2 \right) = f_a^2 \left(\sum_{j=1}^{r} \left(f_a^1 \left(\sum_{k=1}^{m} \left(x_k \cdot w_{k,j}^1 \right) + b_j^1 \right) \cdot w_{j,i}^2 \right) + b_i^2 \right), \;\; i = 1..n \tag{15}$$

Due to the above presented mathematical form, feed-forward neural networks can approximate/replace any kind of function. By using multiple hidden layers of neurons with sigmoid activation functions, a very good approximation could be obtained even for nonlinear relations between the NN input and output data. Linear

transfer functions applied on output layer neurons allows the network to provide any kind of output values. On the other hand, if it is desired to limit the output values, a sigmoid transfer function is advised to be used on the output layer [23].

A particular type of NN is the so-called *radial basis neural network*. This network contains a single hidden layer of neurons that uses the exponential function as transfer function. On the output layer, the linear activation function is used, similar to most feed-forward neural networks.

Recurrent neural networks have been also developed. In this case, the neurons from the hidden layers of the network, receive as input data and their own output value or the output data of nearby neuron layers. Recurrent neural networks are usually used in for the implementation of dynamic systems. For this reason, these networks are also called sequential networks [24].

However, the most common network architecture remains the feed-forward one due to the ease of implementation and training.

4.3 Training of Neural Networks

The process through which a neural network is taught to provide at its output the values of a specific desired function is called *training*. During the training process, the weights and bias values of the neurons are established so that the y outputs generated by network for a set of x input data, would be as close as possible to the target y^* values. Figure 7 graphically presents the error backpropagation principle in the NN training process.

Based to this principle the weights and bias values are continuously adjusted through an iterative process according to the error between the actual NN output values and the network desired ones. Several training algorithms were developed from this basic error backpropagation principle like: gradient descent algorithms; conjugate gradient algorithms; quasi-Newton algorithms; and Jacobian based Levenberg–Marquardt and Bayesian Regularization algorithms [24].

Usually, a large number of input/target output pairs *(x, y*)* are used to train a neural network. The values of the weights and biases depend on the applied training algorithm and error evaluation technique. The evaluation of the error between the provided NN output data and the target values is done through a *cost* ("fitness") *function*. Since

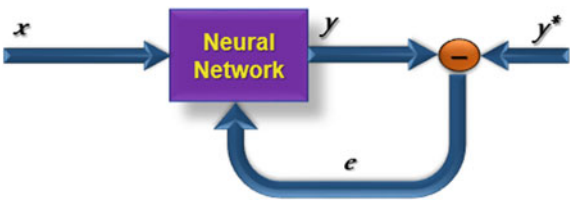


Fig. 7 Error backpropagation principle used to train neural networks

the cost functions express the deviation from the desired NN behaviour, these functions are also called as *quality indicators* of the networks. The most commonly used quality indicator is the mean square error [24], see Eq. (16).

$$MSE = \frac{1}{n} \cdot \left(\sum_{i=1}^{n} \left(y_i^* - y_i \right)^2 \right) \tag{16}$$

5 Identification of the Proper Equivalent Multi-layer Earth Structure Through a Genetic Algorithm Based AI Technique

5.1 Description of the Presented Application

Several electrical engineering applications like grounding grid design for power substations or strategic buildings, cathodic protection design of underground metallic gas or oil pipelines, design of lightning protection system require as a first step a proper knowledge of the earth structure from an electrical (soil resistivity) point of view. The following application meant to exemplify how genetic algorithm based optimization techniques could be applied in electrical engineering.

In order to determine the equivalent multi-layer earth structure corresponding to on site soil resistivity measurements the authors have developed an AI based optimization technique [25, 26]. The implemented genetic algorithm identifies the optimum value of the resistivity and the width of each soil layer considering horizontal multi-layer earth model, in order to reconstruct the measured apparent soil resistivity data.

The developed GA optimization will be applied to determine the proper multi-layer soil structure for two different locations (ML1 and ML2) where on-site Wenner type soil resistivity measurements were carried out. Obtained earth structure data are compared to soil configurations provided by dedicated software applications (the RESAP tool from the CDEGS software package [27]).

5.2 Implemented Genetic Algorithm

The implemented GA starts from a population of 30 individuals, randomly generated, each of them representing a possible configuration of the multi-layered soil model that has to be determined. The chromosome structure of each possible solution contains the resistivity, respectively the thickness of the equivalent soil layers.

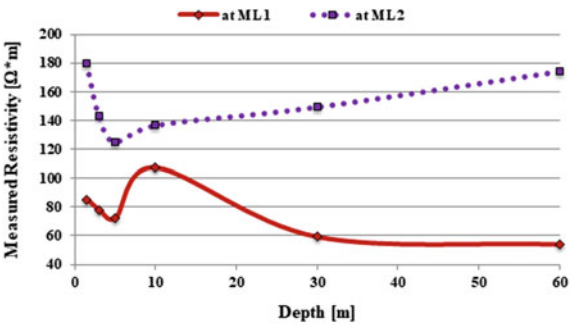


Fig. 8 Measured Wenner apparent soil resistivity curve at location MP1 and MP2

To determine the optimal soil structure, the implemented GA uses a cost function, Eq. (17), that evaluates the mean square error between the Wenner apparent resistivity curve, obtained through soil resistivity measurements (see Fig. 8), and the apparent resistivity curve, related to a possible multi-layer earth configuration:

$$MSQ_{Err} = \frac{1}{n}\sum_{i=1}^{n}[\rho_a(d_i) - \rho_{Ea}(d_i)]^2 \tag{17}$$

where: MSQ_{Err} is the mean square error; n is the number of measurement points; $\rho_a(d_i)$ is the apparent soil resistivity value measured through the Wenner method; and $\rho_{Ea}(d_i)$ is the apparent soil resistivity value corresponding to a horizontal equivalent multi-layered earth model, numerically computed for a depth d_i, $i = 1..N$.

For the numerical evaluation of the apparent soil resistivity, related to a possible soil configuration, the following equation was adopted [28]:

$$\rho_{Ea}(d) = \rho_1 \cdot [1 + 2 \cdot F_L(d) - F_L(2 \cdot d)] \tag{18}$$

where the value of the $F_L(d)$ function is given by the semi-infinite integral:

$$F_L(d) = 2 \cdot d \cdot \int_0^{\infty} \frac{K_{L1} \cdot e^{-2\cdot\lambda\cdot h_1}}{1 - K_{L1} \cdot e^{-2\cdot\lambda\cdot h_1}} \cdot J_o(\lambda \cdot d) \cdot d\lambda \tag{19}$$

with $J_0(\lambda \bullet d)$ the first kind, zero order, Bessel function and K_{L1} a coefficient determined by:

$$K_{L1} = \frac{k_1 + K_{L2} \cdot e^{-2\cdot\lambda\cdot h_2}}{1 - k_1 \cdot K_{L2} \cdot e^{-2\cdot\lambda\cdot h_2}} \tag{20}$$

and

$$K_{Lj} = \frac{k_j + K_{Lj+1} \cdot e^{-2\cdot\lambda\cdot h_{j+1}}}{1 - k_j K_{Lj+1} \cdot e^{-2\cdot\lambda\cdot h_{j+1}}}, \quad j = 1..L-2 \tag{21}$$

where $K_{LL\text{-}1} = k_{L\text{-}1}$, L being the number of horizontal layers, h_j the thickness of the jth layer and k_j the reflection coefficient between layers j and $j + 1$ with soil resistivity ρ_j and ρ_{j+1} respectively:

$$k_j = \frac{\rho_{j+1} - \rho_j}{\rho_{j+1} + \rho_j} \tag{22}$$

5.2.1 The Iterative Optimization Process

To obtain the optimal equivalent earth horizontal model, the set of possible solutions, from the initial GA population, is involved in an iterative process defined by the following steps [25]:

Step 1: The cost function is evaluated for all the possible soil configuration from the current GA population and the best suited ones are directly transferred to the next GA generation;
Step 2: Two soil configuration are randomly selected and subjected to the crossover operator to obtain two new equivalent soil configurations with lower cost function values for the GA next generation;
Step 3: The previous step is repeated until the next GA generation will have the same number of individuals (possible solutions) as the current one;
Step 4: In order to maintain solution diversity four soil configurations are randomly selected and the mutation operator is applied on them;
Step 5: The iterative GA optimization process restarted form *Step 1*.

The maximum number of GA iterations was set to $N = 2000$, a value identified by the authors to be high enough to obtain accurate soil configurations. This way, the implemented GA identifies the optimal parameters of a specific multi-layer earth structure, according to on site Wenner apparent soil resistivity measurements.

5.2.2 Chromosome Structure

Each possible soil configuration solution is represented in the GA optimization process by its chromosome structure formed by the resistivity and thickness of each soil layer scaled to [0,1] range, as in Eq. (23).

$$C = \{\rho_1, h_1, \rho_2, h_2, \ldots, \rho_L\} \tag{23}$$

5.2.3 Crossover Operator

During the crossover process, six new soil configurations are obtained from the initial two solutions selected for crossover recombination, applying three different crossover operators. The first two configuration (GA children) are obtained through an arithmetic crossover operator, Eq. (24):

$$\begin{aligned} C_1^t &= \alpha \cdot P_1^t + (1-\alpha) \cdot P_2^t \\ C_2^t &= \alpha \cdot P_2^t + (1-\alpha) \cdot P_1^t \end{aligned} \tag{24}$$

where: α is a randomly determined scaling factor, t denotes the tth parameter of an equivalent soil model, C_i and P_j represent the ith GA child configuration and jth GA parent soil configuration.

Another two new soil configurations are obtained using a Max–Min type crossover operator, Eq. (25):

$$\begin{aligned} C_3^t &= \min(P_1^t, P_2^t) \\ C_4^t &= \max(P_1^t, P_2^t) \end{aligned} \tag{25}$$

The last two GA child soil configurations are generated applying the classical cut-point crossover operator [8], Eq. (26):

$$\begin{aligned} C_5 &= \left(P_1^1 \cdots P_1^k \; P_2^{k+1} \cdots P_2^r\right) \\ C_6 &= \left(P_2^1 \cdots P_2^k \; P_1^{k+1} \cdots P_1^r\right) \end{aligned} \tag{26}$$

where k is a randomly selected cut point and r is the total number of chromosome structure parameters.

From these six GA child configurations the best two ones with lower cost function values are transferred the next GA generation population.

5.2.4 Mutation Operator

Within the mutation process, each parameter that has to be optimized from a possible multi-layer earth configuration is subjected to a probabilistic test. If the test is passed, then the value of the selected parameter is slightly changed through the following arithmetic mutation operator:

$$C^t = C^t + (0.5 - \alpha) \cdot M \tag{27}$$

with α a random value from the [0,1] range and M a predefined mutation coefficient.

5.3 Computed Equivalent Soil Models

The above presented GA optimization process was applied by the authors to determine the equivalent soil structure based on the on-site Wenner soil resistivity measured at location ML1 and ML2 (see measured apparent soil resistivity curves from Fig. 8) considering a three horizontal layer earth structure. To validate the obtained multi-layer earth configurations a comparison has been done with the RESAP module of CDEGS software package (see Table 1).

Based on the layer resistivity and thickness values obtained through the implemented GA optimization process and the RESAP module respectively (see Table 1, Fig. 9a and Fig. 10a) the apparent soil resistivity curves were generated according

Table. 1 Obtained equivalent three horizontal layer soil models

		$\rho_1[\Omega/m]$	$h_1[m]$	$\rho_2[\Omega/m]$	$h_2[m]$	$\rho_3[\Omega/m]$	$h_3[m]$
ML1	CDGES	80.77	0.99	82.52	13.61	49.31	Inf
	GA	83.84	5.25	101.18	4.63	51.78	Inf
ML2	CDGES	210.69	1.29	109.02	5.16	168.55	Inf
	GA	178.678	1.29	112.34	6.67	174.99	Inf

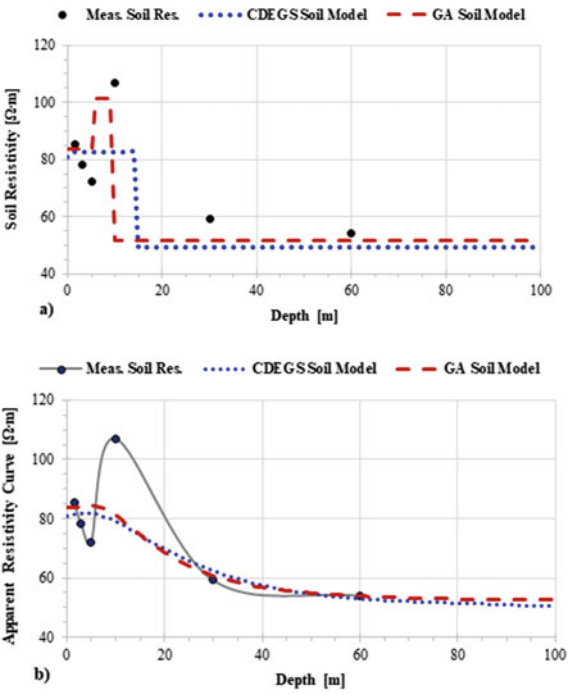


Fig. 9 Obtained three-layer equivalent earth structure (**a**) and Apparent soil resistivity curves (**b**) with the CDEGS software and the implemented GA for measurement location ML1

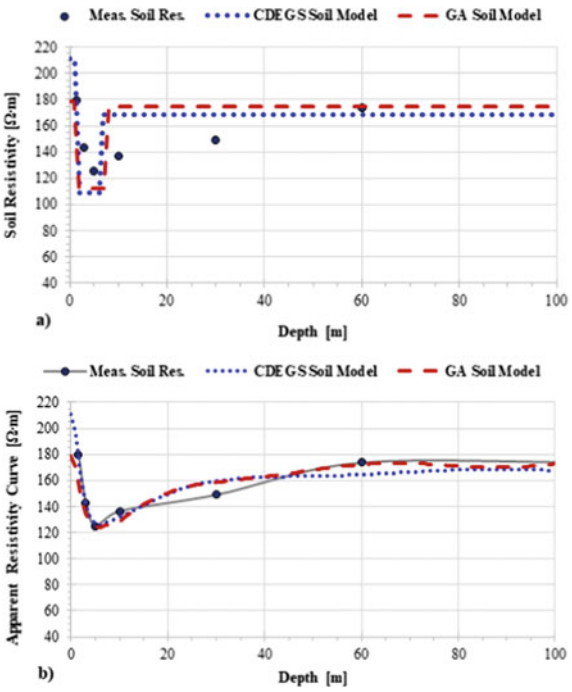


Fig. 10 Obtained three-layer equivalent earth structure (**a**) and Apparent soil resistivity curves (**b**) with the CDEGS software and the implemented GA for measurement location ML2

to equations (18–22) and compared to the on-site Wenner apparent soil resistivity measurements (see Fig. 9b and Fig. 10b).

For measurement location ML1, the average deviation from the measured apparent soil resistivity curve is 7.66% for the equivalent three-layer soil structure provided by the RESAP module of the CDEGS software package, while the average deviation for the soil structure provided by the implemented genetic algorithm is 7.16% (see Fig. 9b).

In case of measurement location ML2 the average deviation from the measured apparent soil resistivity curve are 3.83% for RESAP and 3.30% with the implemented GA optimization process (see Fig. 10b).

Similar comparisons have been carried out by the authors for uniform and two-layer horizontal earth structures in [25] and [26]. Based on the obtained results it can be concluded that the implemented GA provides an accurate alternative to evaluate the equivalent multi-layer earth structure using to on-site apparent soil resistivity measurements.

The above presented multi-layer soil structure GA optimization technique was also applied by the authors at archaeological sites in order to identify and establish the trajectory of buried walls, according to the obtained equivalent earth configurations [29].

6 Neural Network Implementation to Evaluate the Inductive Coupling Matrix in Case of a HVPL – MGP Electromagnetic Interference Problem

6.1 Description of the Studied Problem

Due to economic policies meat to limit construction costs and to environmental regulations meant to protect wildlife and nature, the access of utility systems to new right-of-ways is highly limited. Therefore, in many situations gas, oil or water transportation metallic pipelines are forced to share the same distribution corridor with high voltage power lines and/or AC electrical railway systems (see Fig. 11) and to be exposed to induced AC currents and voltages [30, 31].

In case of underground or above ground metallic pipelines, the induced electromagnetic interferences produced by nearby high voltage power lines could be dangerous on both the operating personnel (that may be exposed to electric shocks), and to the structural integrity of the pipeline, due to corrosion phenomena [31].

Induced AC currents and voltages may appear as effect of inductive, conductive or capacitive coupling mechanisms. However, during power line normal operating conditions, only the inductive coupling, described by the self and mutual inductance matric, has to be considered for underground pipelines. Conductive and capacitive coupling may be, also, neglected when a phase to ground fault happens on the power line far away from the common distribution corridor [30, 32].

To evaluate the self and mutual inductance between all the present conductors in the analysed problem geometry (phase wires, sky wires and pipelines) the magnetic vector potential must be evaluated on the cross section of these conductors as presented in [33, 34]. The longitudinal z-direction component of the magnetic

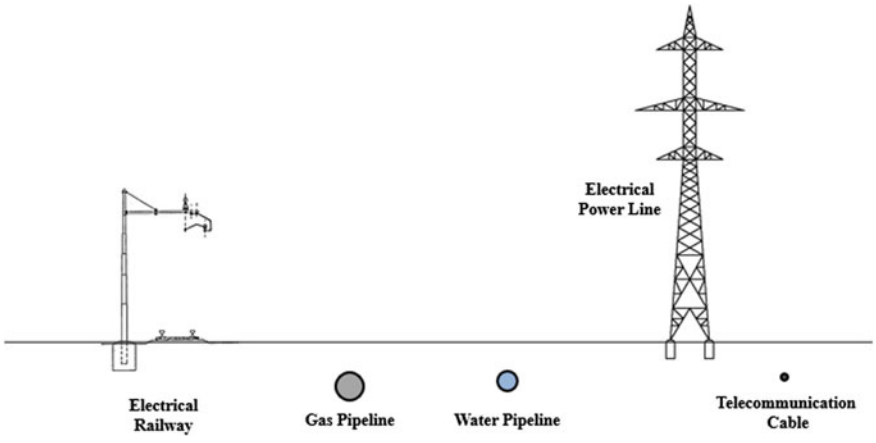


Fig. 11 Common distribution corridor of multiple utilities

vector potential A_z and the total current density J_z are described by the following system of differential equations:

$$\begin{cases} \dfrac{1}{\mu_0 \mu_r} \cdot \left[\dfrac{\partial^2 A_z}{\partial x^2} + \dfrac{\partial^2 A_z}{\partial y^2} \right] - j\omega\sigma A_z + J_{sz} = 0 \\ -j\omega\sigma A_z + J_{sz} = J_z \\ \displaystyle\iint_{S_i} J_z ds = I_i \end{cases} \tag{28}$$

where σ is the conductivity, ω is the angular frequency, μ_0 is the magnetic permeability of free space ($\mu_0 = 4 \cdot \pi \cdot 10^{-7}$ H/m), μ_r is the relative permeability of the environment, J_{sz} is the source current density in the z-direction and I_i is the imposed current on conductor i of S_i cross section.

To solve this differential equation system, the finite element method (FEM) is recommended to be used. Although the calculation process based on FEM, used in the hybrid method presented in [34], provides accurate solution for the magnetic vector potential, regardless of the complexity of the problem, the computation time of the method increases with the complexity of the geometry, the size of the discretization network, the characteristics of the material and the number of parameters being evaluated.

Therefore, the authors have implemented a neural network solution to evaluate the inductive coupling matrix for a specific electromagnetic interference problem between a 220 kV/50 Hz overhead High Voltage Power Lines (HVPL) and underground Metallic Gas Pipeline (MGP) [35], considering a stratified soil structure for the common distribution corridor with three vertical layers (see Fig. 12).

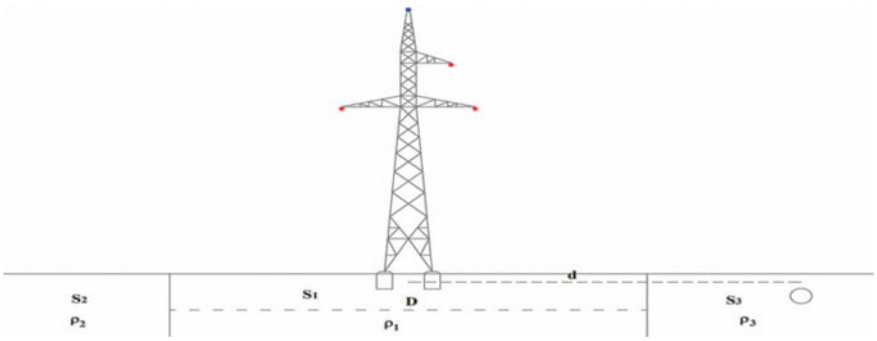


Fig. 12 Interference problem HVPL-MGP with vertically layered earth

6.2 Proposed Neural Network Solution

Once the proposed neural network solution will be trained it will have to be able to instantly evaluate the self and mutual inductance matrix for any possible geometric configuration of the investigated electromagnetic interference problem. Therefore, to implement the proposed NN, the input and desired output data values must be analysed. The following geometrical and electrical parameters of the studied problem were chosen as input values:

- d—HVPL-MGP separation distance (with variation in the 0–1000 m range);
- ρ_1—middle layer resistivity S1 (with variation in the 10–5000 Ω m range);
- D—middle layer width S1 (with variation in the 20–1200 m range);
- ρ_2, ρ_3 (considering $\rho_2 = \rho_3$) —sideways layers resistivity S2 and S3 (with variation in the 10–5000 Ω m range);

Tacking into the account that the inductance matrix is a symmetrical one, the proposed NN should provide only the elements above the main diagonal. For the investigated HVPL-MGP interference problem (three phase wires, one sky wire and one underground pipeline) these inductance elements are: $L_{11}, L_{12}, L_{13}, L_{14}, L_{15}, L_{22}, L_{23}, L_{24}, L_{25}, L_{33}, L_{34}, L_{35}, L_{44}, L_{45}, L_{55}$, with L_{ii} representing the self-inductance of conductor i ($i = 1.0.3$ for phase wires, $i = 4$ for the sky wire and $i = 5$ for the underground pipeline) and L_{ij} representing the mutual inductance between conductor i and j.

Due to the large variation range of the inductance matrix elements value (the self-inductance values are much higher than the mutual inductance values), it was concluded to implement three different neural networks: NN1 for the self-inductance values ($L_{11}, L_{22}, L_{33}, L_{44}, L_{55}$), NN2 for the MGP mutual inductances ($L_{15}, L_{25}, L_{35}, L_{45}$) and NN3 for the remaining mutual inductances between HVPL conductors. This way the complexity of the implemented NN will be reduced, so that the required training time will also be reduced, and the obtained results accuracy will be increased.

6.3 Matlab Implementation of Proposed Neural Network

The Neural Networks toolbox from the MATLAB software package [24] was used to implement, test and validate the proposed NN solution. A feed-forward architecture with two hidden layers and an output layer was chosen (as in Fig. 13).

To identify the optimal configuration of the chosen NN architecture different transfer functions and various number of neurons on the NN hidden layers were tested. The number of neurons on each hidden layer was varied between 5 and 30 with a step of 5. The “tansig” (sigmoid tangent) and “logsig” (logarithmic sigmoid) transfer functions were tested for the NN hidden layer neurons while the “purelin” (linear) transfer function was used for the output layer neurons. To automatically generate and test all these different possible NN configurations a Matlab code “.m”

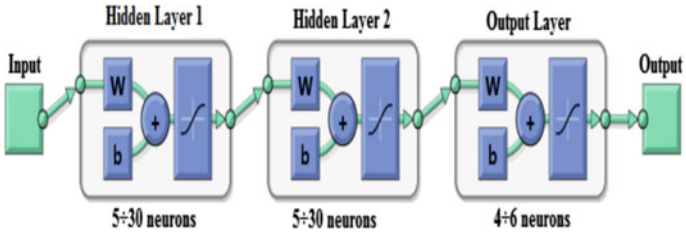


Fig. 13 Implemented feed-forward architecture with two hidden layers and an output layer

file was implemented, using the *feedforwardnet* Matlab function [24]:

$$\text{net} = \textit{feedforwardnet}(\text{hiddenSizes}, \text{trainFcn}) \tag{29}$$

where: *net* is the created feed-forward neural network, *hddinSizes* is a vector of values specifying the number of neurons used on each hidden layer and *trainFcn* is a vector of strings defining the transfer function used on each NN layer.

To create a useful training database for the investigated HVPL-MGP electromagnetic interference problem approximately 4000 inductance matrixes were determined through FEM analysis for various problem geometries. The HVPL-MGP separation distance was varied between 0 and 1000 m, the resistivities of the vertical soil layers were varied between 10 Ω m and 5000 Ω m while the width of the middle soil layer was varied between 20 and 1200 m. Table 2 shows some of the HVPL-MGP problem geometries used to train the proposed neural networks. different configurations used to stimulate the NN.

The NN training process took between 1 and 25 min depending on the NN configuration complexity. The Levenberg–Marquardt training method ("trainlm") was used with a mean square error ("mse") cost function on a i7-3632QM 2.2 GHz Intel Core

Table. 2 Different problem geometry configurations used for NN training

Case no	d $[m]$	D $[m]$	ρ_1 $[\Omega \cdot m]$	ρ_2 $[\Omega \cdot m]$	ρ_3 $[\Omega \cdot m]$	**Case No**	d $[m]$	D $[m]$	ρ_1 $[\Omega \cdot m]$	ρ_2 $[\Omega \cdot m]$	ρ_3 $[\Omega \cdot m]$
8	5	60	500	50	500	2134	0	550	50	250	50
104	100	60	150	250	150	2301	20	550	30	250	30
206	20	60	50	500	50	2532	100	550	100	500	100
373	100	60	500	750	500	2751	500	550	30	100	30
481	150	60	500	250	500	2914	5	1050	10	250	10
692	1000	60	750	50	750	3096	20	1050	100	250	100
875	20	120	750	100	750	3274	100	1050	500	1000	500
1064	50	120	750	1000	750	3545	750	1050	30	750	30
1231	500	120	100	30	100	3754	5	1500	50	30	50

Table. 3 HVTL-MGP problem geometries used for the NN testing procedure

Case no	d $[m]$	D $[m]$	ρ_1 $[\Omega \cdot m]$	ρ_2 $[\Omega \cdot m]$	ρ_3 $[\Omega \cdot m]$	Case No	d $[m]$	D $[m]$	ρ_1 $[\Omega \cdot m]$	ρ_2 $[\Omega \cdot m]$	ρ_3 $[\Omega \cdot m]$
1	310	800	900	850	900	85	310	800	900	850	900
13	105	1100	550	550	550	97	170	700	300	350	300
25	250	800	150	150	150	109	240	500	80	750	80
37	340	400	600	150	600	121	420	100	550	20	550
49	170	800	650	750	650	135	105	1200	250	950	250
54	55	1000	900	400	900	148	85	400	140	160	140
61	40	200	600	800	600	176	15	300	140	700	140
73	120	900	750	350	750	198	10	1000	200	750	200

PC, with a 64-bit operating system and 8 GB RAM memory. To train the implemented NN configurations the *train* Matlab function was applied [24].

6.4 Obtained NN Results

In order to determine the accuracy of the generated NN architectures and to identify the optimal NN configuration for each of the three implemented NN solutions (NN1, NN2 and NN3 respectively) an addition set of approximatively 200 randomly generated, testing HVPL-MGP problem geometries were used. These testing HVPL-MGP problem geometries were not supplied to the implemented NN configuration during the NN training process. Table 3 shows some of the testing HVPL-MGP problem geometries.

To identify the optimal NN configurations the evaluation error of the provided NN output data was analysed for both the training and testing data sets [35, 36]. To obtain NN provided output data for the training and testing HVPL-MGP problem geometries the *sim* Matlab function was applied.

For the neural network meat to evaluate the self-inductance values of the conductors (NN1 network) the best identified NN configuration was a feed-forward architecture with 15 neurons on the first hidden layer and 25 neurons on the second hidden layers, with "tansig" transfer function on both hidden layers. The obtained average evaluation errors are 0.064% for the testing geometries and 0.043% for the training geometries. The maximum recorded evaluation error was 0.77%. Figure 14 presents the evaluation error distribution on different error classes for both training and testing HVPL-MGP problem geometries.

In case of the neural network implemented to compute the mutual inductance values that define the electromagnetic coupling between MGP and the nearby HVPL (NN2 network) the best NN configuration has 30 neurons on the first hidden layer and 20 neurons on the second layer with "logsig" transfer function. The average

Fig. 14 Evaluation error distribution for the optimal NN1 architecture

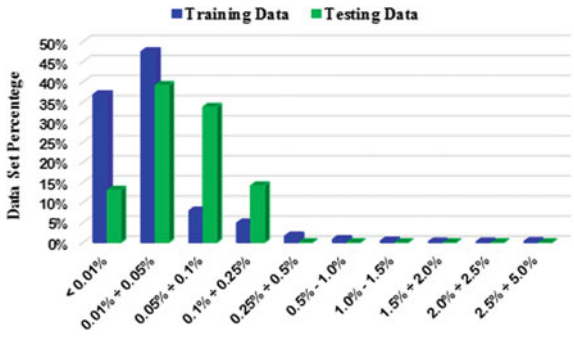


Fig. 15 Evaluation error distribution for the optimal NN2 architecture

evaluation error was around 0.060% for both testing and training data sets, while the maximum recorded evaluation error was 2.67%. The evaluation error distribution over the analysed error classes for the testing and training HVPL-MGP problem geometries is presented in Fig. 15.

For the neural network used to compute the mutual inductance values between HVPL conductors (NN3 network) the best NN configuration has 25 neurons, respectively 15 neurons with "tansig" transfer function on the NN hidden layers. The maximum recorded evaluation error is 2.56% while the average evaluation error is around 0.030% for both testing and training data sets. Figure 16 shows the obtained evaluation error distribution over different error classes:

The implemented NN configurations allow to evaluate the inductance matrix values for any HVPL-MGP problem geometry. Table 4 shows the self and mutual inductance values obtained for a HVPL-MGP problem geometry with a 30 m separation distance between HVPL and MGP; with $\rho_1 = 30\ \Omega$ m, $\rho_2 = \rho_3 = 500\ \Omega$ m and a 20 m width for the middle earth layer. Using the self and mutual inductance values provided by the implemented neural network configurations the equivalent electrical circuit of the investigated HVPL-MGP electromagnetic interference problem could be constructed according to [34, 37].

The *InterfStud* software application developed by the authors [38] automatically creates the above-mentioned equivalent circuit model and evaluates the induced AC

Fig. 16 Percentage error distribution for the optimal NN3 network

Table. 4 Obtained inductive coupling matrix through NN implementation

Self and mutual inductances [μH/m]					
	PhW A	PhW B	PhW C	SkyW	Pipe
PhW A	2.45	1.234	1.110	1.187	0.82
PhW B	1.234	2.45	1.100	1.073	0.84
PhW C	1.110	1.100	2.45	1.073	0.80
SkyW	1.187	1.073	1.073	8.74	0.79
Pipe	0.822	0.842	0.80	0.795	2.28

currents and voltages in the MGP. Figure 17 presents the obtained induced AC voltages for the three different problem geometries [35], considering a 10 km long parallel HVPL-MGP exposure, a 130 MVA power load on HVPL with a 0.94 power factor (a 350 A symmetrical current load):

- ***Geom* 01**: A 30 m separation distance, with soil structure: $\rho_1 = 30\ \Omega$ m, $\rho_2 = \rho_3 = 500\ \Omega$ m, and 20 m middle layer width;
- ***Geom* 02**: A 50 m separation distance, with soil structure: $\rho_1 = 10\ \Omega$ m, $\rho_2 = 100\ \Omega$ m, $\rho_3 = 500\ \Omega$ m, and 30 m middle layer width;

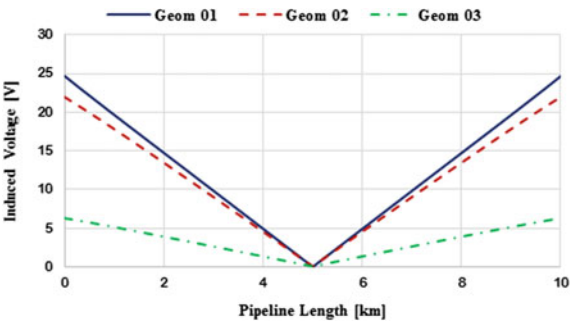


Fig. 17 Induced voltage in MGP for different HVPL-MGP problem geometries

- ***Geom* 03**: A 150 m separation distance, with soil structure: $\rho_1 = \rho_2 = 100\ \Omega$ m $\rho_3 = 1000\ \Omega$ m, and 100 m middle layer width.

7 Conclusions

This chapter starts with a brief introduction to artificial intelligence (AI) based advanced numerical methods applied in engineering, making a summary of the most commonly used AI techniques (Genetic Algorithms, Fuzzy Logic and Neural Networks, Sects. 2–4) and new approaches in the field (through two demonstrative applications).

The first application (Sect. 5) presents a genetic algorithm implementation to determine the equivalent horizontal soil structure based on Wenner on-site soil resistivity measurements. A proper knowledge of the earth structure is required in electrical engineering application like grounding grid design for power substations, cathodic protection design of underground metallic gas or oil pipelines, design of lightning protection.

The presented multi-layer soil structure GA optimization technique was also applied by the authors at archaeological sites in order to identify and establish the trajectory of buried walls, according to the obtained equivalent earth configurations.

In the second presented application (Sect. 5) a neural network based artificial intelligence technique has been implemented to evaluate the inductive coupling matrix of a specific HVPL-MGP electromagnetic interference problem. The proposed neural network approach reduces considerably the required computation time. From Figs. 14–16 it can be observed that the evaluation error produced by the identified optimal NN architecture are usually less than 0.1% in comparison to the finite element results considered as reference. Therefore, the implemented neural network solution to evaluate the self and mutual inductance values is a very effective one, especially if we take into account the fact that the solutions provided by neural networks are obtained almost instantaneously and can be used to evaluate the induced currents and voltages.

References

1. McCarthy J, Minsky ML, Rochester N, Shannon CE (1995) A proposal for the dartmouth summer research project on artificial intelligence. https://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html. Accessed 31 Aug 1995
2. Brunette ES, Flemmer RC, Flemmer CL (2009s) A review of artificial intelligence. 4th International conference on autonomous robots and agents (ICARA), pp 385–392, Wellington, New Leeland, February 10–12 2009
3. Rechenberg I (1971) Evolutionsstrategie: Optimierung technischer Systeme nach PrinLipien der biologischen evolution", Dr. Ing. thesis, Technical University of Berlin, Department of Process Engineering, Berlin

4. Beyer HG, Schwefel HP (2002) Evolution strategies—a comprehensve introduction. Nat Comput 1(1):3–53
5. Holland JH (1975) Adaptation in natural and artificial systems, Ann Arbor: The University of Michigan Press
6. Vose M (1999) The simple genetic algorithm: foundations and theory. MIT Press, Cambrige, Cambridge, MA
7. Sivanandam SN, Deepa SN (2008) Introduction to genetic algorithms, Springer, New York, USA
8. Haupt RL, Werner DH (2007) Genetic algorithms in electromagnetics, IEEE Press, Wiley-Interscience
9. Reeves CR, Rowe JE (2003s) Genetic algorithms: principles and perspectives. A guide to GA theory, Kluwer Academic Publishers, Dordrecht, The Netherlands
10. Yu X, Gen M (2010) Introduction to evolution algorithms, Springer, London, UK
11. KoLa JR (1992) Genetic programing, MIT Press, Cambrige, USA
12. Ladeh LA (1965) Fuzzy sets. Inf Control 8:338–353
13. Dubois D, Prade H (1980) Fuzzy sets and systems: theory and applications, Academic Press, New York, USA
14. Chen G, Pham TT (2001) Introduction to fuzzy sets, fuzzy logic and fuzzy control systems, CRC Press, Boca Raton, Florida, USA
15. Fuzzy Logic Toolbox (2019) User's guide, ver. 2.5, The Mathworks Inc.
16. Ross TJ (2004) Fuzzy logic with engineering applications, 2nd edn, Wiley, Chchester, UK
17. Buckley JJ, Eslami E (2002) An introduction to fuzzy logic and fuzzy sets, Springer-Verlag, Berlin
18. McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biol 5(4):115–133
19. Caudil M, Butler C (1992) Understanding neural networks: computer exploration, vol 1, MA: The MIT Press, Cambrige
20. Hertz J, Krogh A, Palmer R (1995) Introduction to the theory of neural computation. Lecture notes, Santa Fe Institute, Adduson-WesleL Publishing Company
21. Hagan MT, Demuth HB, Beale MH, de Jesus O. Neural network design, 2nd edn, eBook
22. Micu DD, Czumbil L, Christoforidis GC, Simion E (2012) Neural networks applied in electromagnetic interference problems. Revue Roum Des Sci Tech Serie Electrotech et Energetique 57(2):162–171, ISSN: 0035-4066
23. Christodoulou C, Georgiopoulos M (2001) Applications of neural networks in electromagnetics, Artech House, Norwood, USA
24. Beale MH, Hagan MT, Demuth HB (2017) Neural networks toolbox. Users's Guide, ver. 10.0, the Mathworks Inc.
25. Şteţ D, Czumbil L, Micu DD, Ţopa V, Ancăş L () Stream gas pipeline in proximity of high voltage power lines. Part I—soil resistivity evaluation. In: 47th International universities' power egineering conference (UPEC), London, UK, September 4–7 2012
26. Czumbil L, Şteţ D, Micu DD, Ţopa V, Ancăş L (2012) Stream gas pipeline in proximitity of high voltage power lines. Part II—induced voltage evaluation. In: 47th Internatonal universities' conference on power energy(UPEC), London, UK, September 4–7 2012
27. Dawalibi FP, Donoso F (1993) Integrated analysis software for grounding, EMF and EMI. IEEE Comput Appl Power 6(2):19–24
28. Yang H, Yuan J, Long W (2001) Determination of three-layer earth from wenner four-probe test data. IEEE Trans Magn 37(5):3684–3687
29. Munteanu MS, Czumbil L, Micu DD, Braicu ŞF, Nemeti S, Pîslaru M (2017) Measurement of soil resistivity in order to determine the buried walls trajectory. Adv Electr Comput Eng (AECE) 17(1):103–108, ISSN: 1582-7445
30. CIGRE (1995) Guide on the infuence of high voltage AC power systems on metallic pipelines
31. Micu DD, Christoforidis GC, Czumbil L (2013) AC Interference on Pipelines due to Double Circuit Power Lines: A detailed study. Electr Power Syst Res 103:1–8

32. Gupta A, Thomas MJ (2006) Coupling of high voltage AC power line fields to metallic pipelines. In: 9th International conference on electromagnetic interference and compatibility, (INCEMIC), Bangalore, India, February 23–24 2006
33. Christoforidis GC, Labridis DB, Dokopoulos PS (2003) Inductive interference calculation on imperfect coated pipelines due to nearby faulted parallel transmission lines. Electr Power Syst Res 66(2):139–148
34. Christoforidis GC, Labridis DP, Dokopoulos PS (2005) A hybrid method for calculating the inductive interference caused by faulted power lines to nearby buried pipelines. IEEE Trans Power Deliv 20(2):1465–1473
35. zumbil L,Micu DD, Şteţ D, Ceclan A (2015) Inductive coupling between overhead power lines and nearby metallic pipelines. A neural network approach. Carpathian J Electr Eng (CJEE) 9(1):29–44
36. Micu DD, Christoforidis GC, Czumbil L (2012) Artificial intelligence techniques applied to electromagnetic interference problems between power lines and metal pipelines. In: Recurrent neural networks and soft computing, Intech, Ch 12, pp 253–274, Rijeka, Croatia
37. Micu DD,Czumbil L, Christoforidis GC, Ceclan A, Şteţ D (2012) Evaluation of induced AC voltages in underground metallic pipeline. COMPEL: Int J Comput Math Electr Electron Eng 31(4):133–1143
38. Czumbil L, Christoforidis GC, Micu DD, Şteţ D, Ceclan A, Pop O (2011) A user-friendly software application for induced AC interference evaluation. In: 46th International universities' power engineering conference, UPEC, Soest, Germany, September 5–8

Numerical Methods for Solving Nonlinear Equations

Narges Mohammadi, Shahram Mehdipour-Ataei, and Maryam Mohammadi

Abstract The nonlinear description has continuously been crucial in a wide range of disciplines to provide an accurate prediction of a natural phenomenon. Thus, finding a reliable solution method for these nonlinear models is of significant importance since, in most real-life applications, direct solution methods are not feasible, even in linear cases. Moreover, an inefficient method is likely to take additional computational cost and effort. This chapter attempts to provide a fundamental description of various iterative methods for solving nonlinear discretized equations. In the first part, a theoretical account of nonlinear systems with different types of iterative methods are depicted. The second part deals with both one-point and multi-point iterative methods; this includes a description of the method, mathematical formulations, and the weak and strong points. Different iterative methods to solve a system of nonlinear equations are then described. Some discussed methods include the family of conjugate gradient, multi-step, and Newton-like. This part also identifies intricacies regarding a system of nonlinear equations, offering different remedies to solve these issues. Finally, a comparative study of the discussed methods and their applications in solving conventional equations are outlined in brief. The iterative methods mentioned in this chapter can be useful not only in solving nonlinear problems but also in linear problems and optimization.

Keyword Nonlinear equations · Iterative methods · Root-finding · Computational efficiency · Jacobian matrix

N. Mohammadi
Faculty of Mechanical Engineering, Amirkabir University of Technology, Tehran, Iran
e-mail: nargesmohammadi@aut.ac.ir

S. Mehdipour-Ataei (✉) · M. Mohammadi
Faculty of Polymer Science, Iran Polymer and Petrochemical Institute, Tehran, Iran
e-mail: s.mehdipour@ippi.ac.ir

M. Mohammadi
e-mail: ma.mohammadi@ippi.ac.ir

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_5

Nomenclatures

A. Acronyms

IVP	Initial Value Problems
BVP	Boundary Value Problems
ODEs	Ordinary Differential Equations
PDEs	Partial Differential Equations
FE	Finite Element
FV	Finite Volume
FD	Finite Difference
CO	Convergence Order
LU	Lower–Upper

B. Symbols/Parameters

x	Unknown value(s)
i	Iteration index
x^*	Actual root
$E.I.$	Efficiency index
ρ	Convergence order
n	Total number of function evaluation
$[a, b]$	Specific interval
$f(x)$	Nonlinear function
f'	First derivative of the nonlinear function
f''	Second derivative of the nonlinear function
$x^{(0)}$	Initial approximation of the zero of the function
C	Computacional cost
$J^{(i)}$	n-dimensional Jacobian matrix
$F(x)$	Column vector of nonlinear functions
$\frac{\partial}{\partial x}$	Partial derivative with respect to the unknown x
J	Jacobian matrix
h	Incriminate value
∇F	Gradient of $F(x)$
T	Transpose sign
$H(x)$	Hessian matrix
λ	Damping parameter
$\eta^{(i)}$	Forcing term
$S^{(i)}$	Inexact newton step

1 Introduction

There is a wide range of natural phenomena, as well as numerous practical applications that can be accurately simulated through mathematical analysis. The formulation of a real phenomenon in a format of the system of equations endows many benefits, such as validating the results of physical experiments and demonstrating a reliable relationship among variables explaining the features of a phenomenon or a system. A problem modeled in a frame of mathematical equations can be either linear or nonlinear. Linear problems are more straightforward to solve than nonlinear problems, with respect to computational cost and implementation. The linear analysis may be applied for nonlinear problems when errors are too small, or they can be tolerated; but in some problems, according to the required accuracy for a problem and the target of analysis, employing the nonlinear analysis is necessary and unavoidable. Some examples demonstrating the importance of a nonlinear description include the design of components for some special usage such as aerospace and nuclear engineering, obtaining an accurate understanding of a phenomenon, or simulating the behavior of some materials. Hence, complex nonlinear equations, in spite of intricacies stemming from solving them, should be employed to provide a perception of the behavior of these phenomena as well as present a realistic approximation of the response [1, 2].

Nonlinear equations can be either one-variable equations (scalar equations) or multi-variable equations (the system of nonlinear equations). A nonlinear equation can be represented as an equation that does not follow the superposition principle, and the output and input of the system are not directly proportional to each other. Regarding the system of nonlinear equations, it is a set of n simultaneous equations with n unknowns that consists of only one or more nonlinear equations [3]. In contrast to scalar nonlinear equations, the system of nonlinear equations due to their complexity has received less attention.

Solving a nonlinear equation is an intractable task; besides, the uniqueness and existence of the nonlinear solution may also be challenging. To solve these nonlinear problems, there are two numerical and analytical methods. Albeit the analytical method provides high accuracy without much computational cost and effort, in most cases, finding a closed-form solution for a nonlinear system is not feasible; these solutions are restricted to some simplified and exceptional cases and are not suitable for real applications. On the other hand, the growing knowledge and improvements of computer technology have made the use of complicated numerical methods an easy task for analyzing physical systems; by writing new programs or using the existing computer packages. Therefore, during the last decades, researchers have been stimulated to develop new computationally efficient methods to find a numerical method approximating a solution for nonlinear initial value problems (IVP) and boundary value problems (BVP) [1, 4, 5].

Nonlinear equations governing a system can take different forms of algebraic, differential– ordinary differential equations (ODEs), partial differential equations (PDEs)—and integral equations. Regarding differential or integral equations,

initially, one of the well-known numerical methods such as Finite Element (FE), Finite Volume (FV), Finite Difference (FD), or other discretization methods may be employed to convert nonlinear equations to nonlinear algebraic equations. These discretization techniques provide a weakly nonlinear system of equations. Then, an iteration method should be adopted to solve this weakly system of nonlinear equations [1, 6].

In iteration methods, by an initial guess, a particular process like a generalized rule should be followed for every new estimation of the final response until the termination criterion meets. Termination criterion is a condition determining an acceptability level for the final allowable error that should be satisfied to terminate to an iteration process. The termination criterion for a small value ε, and the approximated values x with the iteration index i can be defined in different ways. Three commonly used of these criteria are as:

(1) Two last responses are very close together:

$$\left|x^{(i+1)} - x^{(i)}\right| < \varepsilon \tag{1}$$

(2) The relative difference of two last responses are small:

$$\left|\frac{x^{(i+1)} - x^{(i)}}{x^{(i)}}\right| < \varepsilon \tag{2}$$

(3) The final response is sufficiently small [2, 7, 8]:

$$\left|f(x^{(i)})\right| < \varepsilon \tag{3}$$

The value of ε can be varied according to each problem and the final goal of that problem. As a numerical method provides an estimate of the exact response, the termination criterion is an important factor defining the accuracy and reliability of results. Moreover, this criterion can significantly affect the time of the process.

According to Traub [9], iterative methods to solve nonlinear equations can be divided into two categories viz one-point and multi-point iterative methods, regarding the fact that whether they employ new data from different points or not. These two methods are also divided into multi-point methods with memory or without memory and one-point methods with memory or without memory based on whether old information is reused or not. In one-point methods without memory, the value of a new estimation, say x_i, is computed only based on the information at x_i. In this method, the only way to enhance the convergence of the problem is by increasing the derivative order. For example, to achieve a method of convergence order i, one should employ the $(i-1)$th derivative of a function. Some examples of this method are Newton, Halley, and Cauchy methods. A method is named a one-point method with memory when the next approximation is obtained based on the information of older points such as $x^{(i-1)}, x^{(i-2)}, \ldots, x^{(1)}$ as well as $x^{(i)}$ to estimate the value of $x^{(i)}$.

One well-known example of this case is the Secant method. This type of iterative methods mostly includes derivative-free algorithms.

Regarding multi-point iterative methods, they use the new information at different points, which results in a more computationally efficient method in comparison to one-point methods. These methods can increase the order of convergence without any need to employ a higher-order derivative as well as provide a wider region of convergence [10].

To select an iterative numerical method for a nonlinear problem, there are some criteria that should be considered to ensure the efficient performance of a solver in providing an accurate result. In the first step, the convergence order (CO) should be measured to identify the speed of a numerical method to obtain the final response. According to Eq. (4), for $\rho \geq 1$ and $K > 0$, if a function with the actual root of x^* satisfies this relationship, it is said that the method converges to x^* with order ρ. This relationship can be calculated based on either n-dimensional Taylor expansion or the matrix approach.

$$\lim_{n\to\infty} \left|x^{(i+1)} - x^*\right| \leq K \left|x^{(i)} - x^*\right|^{\rho} \tag{4}$$

In addition, if $\rho = 1$ and $K = 1$, it is said that the method is super-linearly convergent, which is faster than linearly convergent [11].

With respect to the convergence, iterative methods can be either locally convergent or globally convergent methods. In the locally convergent method, the order of convergence is greater than 1. Moreover, to ensure the convergence to the accurate root, this method requires an initial guess sufficiently close to the root. In contrast, a globally convergent method is not restricted to a good initial approximation to converge, but the convergence order of these methods is often lower than locally convergent methods. For example, the Newton method is a second-order locally convergent method highly sensitive to the initial guess; whereas the bisection method linearly converges even for a poor initial approximation [12].

The number of call functions or function evaluations required at each step can also determine the efficiency and the CPU time (running time of an algorithm) required for a process; this can be determined by the number of times that goal function and its supporting functions computed during the process. In a system of nonlinear equations along with the number of function evaluations, other operations such as matrix-vector multiplications, matrix-matrix multiplications, Jacobian evaluations, and Jacobian inverses calculation are deemed factors determining the efficiency of a method. These factors are important since they may cause a very time-consuming process, in particular, in temporal problems in engineering disciplines, which require the repetition of a lengthy iterative procedure to solve the system of nonlinear equations per time step. Another factor is the radius of convergence that is often ignored

since it is difficult to calculate. Lastly, the CPU time used for a process can again determine the efficiency and speed of a method, in particular, for a large nonlinear system. Obviously, a process with less CPU time is a better method. All factors mentioned above can affect each other; for example, to obtain more accuracy, a method should use more function evaluations [7, 13, 14]. On the other hand, considering all these factors for each problem is intractable; therefore, regarding each problem, some of these criteria may be ignored; for example, to reach a higher accuracy, one may have to employ more function evaluations. Hence, it is of great importance to identify various numerical methods and their characters so that the most suitable one can be opted according to the circumstances.

To this aim, the rest of the chapter depicts a fundamental description of various iterative methods for solving nonlinear algebraic equations originating from discretized partial and ordinary derivative or integral equations. In the following parts, initially, both on-point and multi-point iterative methods to solve scalar equations are described; they are compared from a computational viewpoint. Then, different commonly-used methods to solve a system of nonlinear equations are described. Advantages and drawbacks of each method are mentioned, and some alternative methods to overcome their weaknesses are represented. Finally, the application of the methods together with a comparative study are provided, in the last part.

2 One-Variable Nonlinear Equations

In this part, a description of different methods to solve one-variable nonlinear equations is presented. Although nonlinear scalar equations are only employed to some limited cases, a knowledge of them can assist researchers in comprehending possible solutions for a system of nonlinear equations more straightforward. Furthermore, basic ideas of some methods to solve one-variable problems can also be used for nonlinear multi-variable problems. Methods depicted in this part consist of one-point and multi-point methods.

In order to evaluate the efficiency of various scalar iterative methods and compare different offered methods with each other, the efficiency index $E.I.$ can be defined as a function of the convergence order ρ and the total number of function evaluation n as $E.I. = \rho^{\frac{1}{n}}$. The higher efficiency index means that a method is of higher speed with lower computationally cost.

2.1 One-Point Methods

One-point iterative methods are the simplest methods for solving scalar nonlinear equations that are not often adopted for complicated problems. These methods can be used either as a base for more complicated methods, such as the Newton and

Steffensen methods, or as a starting method for locally convergent methods like bisection method.

I. **Bisection Method**

By assuming f as a continuous scalar function at a specific interval, say $[a, b]$, in which the sign of $f(a)$ and $f(b)$ are opposite, the bisection method determines that there is at least one root in this interval based on the intermediate value theorem. This method computes the value of f at the midpoint of the domain $c = \frac{a+b}{2}$. Then, according to the sign of the $f(c)$, the value of c is replaced with either a or b. As a result, a new smaller subinterval around the root of the equation can be established for the next step. This procedure recurs until the termination criterion meets [15].

One of the advantages of the halving method is that it only uses one function evaluation per iteration, except for the first step, it employs two functions that can be overlooked by the final number of functions used at the end of the process. This method also unconditionally converges to the actual root provided that the nonlinear function is continuous in the whole of the domain. Moreover, the number of iterations for a specified accuracy can be obtained in advance; for example, the error after n iteration will be less than $\left|\frac{b-a}{2^n}\right|$. Nevertheless, this method is not widely used for complicated nonlinear equations due to the low rate of convergence (linear convergence) for the same accuracy in comparison with other methods. Additionally, for a function with multiple roots in an interval, this method may not work accurately, since the sign of endpoints may be the same. Therefore, it is proposed to employ this method to find a rough estimate of a function for other faster methods requiring a proper approximate initial value; that is, for the first iterations, this method can be applied to obtain an approximation sufficiently close to the root, and for the next iterations, other faster methods can be used to accelerate the speed of the calculation [8, 15].

II. **Fixed-Point Method**

This method is a simple and stable method that can be employed to solve continuous functions. This method initially converts the function of $f(x) = 0$ to $g(x) - x = 0$. Then, to find the root of $f(x)$, it addresses the iteration equation as $x^{(i+1)} = g(x^{(i)})$; in each iterate, the response of the function is set as the new point. As the bisection method, this method is also linearly convergent.

The rate of convergence of this method may change for a different arrangement, and even for some arrangement, it either may not converge or converge to a false root. There is not a general rule to determine what arrangement is the best; nevertheless, based on the previous research, it can seem that the slope of the curve of the function $g(x)$ can affect the speed of the algorithm [8, 15].

III. **Muller Method**

In contrast with previous methods, this method approximates the function $f(x)$ by a quadratic polynomial as $ax^2 + bx + c$, which can increase the speed of the method.

By using three points of x_0 and x_2 as endpoints and x_1 between two endpoints, near the zero of the function, the coefficients of this quadratic polynomial can be obtained. In the next step, the nearest root to the middle point, x_1, is set as the new point. The resulting value is assigned as a new point for the next approximation along with the two points that are closer to the root.

Muller's method can be used to obtain complex root if the starting point is a complex number. In addition, except for the first iteration that requires three function evaluations, this method only employs one function evaluation per iteration [8, 15].

IV. Newton Method

The well-known Newton method approximates the roots of a nonlinear functions on the basis of the first two terms of the Taylor polynomial expansion of the nonlinear function $f(x)$ expanded about the actual root of the equation x^* as:

$$f(x^*) = f(x^{(0)}) + f(x^* - x^{(0)})f'(x^{(0)}) + \frac{(x^* - x^{(0)})^2}{2} f''(x^{(0)}) \tag{5}$$

where f' and f'' are the first and second derivative of the $f(x)$, respectively, and $x^{(0)}$ is the initial approximation of the zero of the function, which is sufficiently close to it. Since $f(x^*) = 0$ and $\left|x^* - x^{(0)}\right|$ is small enough, the terms with higher order can be ignored, and the iterative method for $n > 1$ iteration can be shown as:

$$x^* \approx x^{(i+1)} = x^{(i)} - (x^{(i+1)} - x^{(i)}) \frac{f(x^{(i)})}{f'(x^{(i)})} \tag{6}$$

According to the above equation, the new point, $x^{(i+1)}$, is considered as the x-intercept of the tangent line to the graph of $f(x)$ at the current approximation x_n.

This method is a highly efficient method that converges quadratically. Additionally, it can work even for complex roots, providing that the initial guess be a complex value. It should be pointed out that this initial root should be sufficiently close to the actual root to avoid this method from being trapped in an endless loop. In comparison with previous methods, this method employs two function evaluations at each step. Additionally, this method requires the second derivative of the nonlinear equation, which may be either unavailable for some problems or burdensome to calculate. Therefore, concerning these potential hurdles, some alternative methods have existed to overcome some. In the following, a number of these methods are presented [16, 17].

V. Secant Method

The secant method is a quasi-Newton method in which the derivative of the function is estimated with the finite-difference approximation. The iteration algorithm of this

method can be shown as:

$$x^{(i+1)} = x^{(i)} - f(x^{(i)}) \frac{(x^{(i+1)} - x^{(i)})}{f(x^{(i+1)}) - f(x^{(i)})} \tag{7}$$

This formulation offers a derivative-free method at the expense of decreasing the convergence order to super linearly convergence. As the Newton method, this method requires to combine with other methods such as the bisection method since this method still requires two good starting points of the zero of the function. The method of False Position can be suggested to prevent this algorithm from diverging. Although this method is of slow convergence and requires more calculations, it always ensures that during the iteration process, the new estimation is always bracketed; that is, the function $f(x)$ changes signs at two x-values [8, 15].

VI. **Steffensen Method**

Unlike previous method, Steffensen method offers a convergence identical to the Newton method; whereas, it does not involve the derivatives of the function. Using divided difference, this derivative-free method offers a proper one-point algorithm for non-differentiable problems as below [18]:

$$x^{(i+1)} = x^{(i)} - \frac{f^2(x^{(i)})}{f(x^{(i)} + f(x^{(i)})) - f(x^{(i)})} \tag{8}$$

VII. **Chebyshev–Halley Methods**

The well-known Chebyshev–Halley methods are improved versions of the Newton method, offering a higher convergence order for the iterative method using the second derivative of the function. This method can be shown as:

$$x^{(i+1)} = x^{(i)} - \left(1 + \frac{L_f(x^{(i)})}{2(1 - \beta L_f(x^{(i)}))}\right) \frac{f(x^{(i)})}{f'(x^{(i)})} \tag{9}$$

$$L_f(x^{(i)}) = \frac{f''(x^{(i)}) f(x^{(i)})}{(f'(x^{(i)}))^2} \tag{10}$$

For different values of β this formulation includes three well-known methods as: Chebyshev method for $\beta = 0$, Halley method for $\beta = \frac{1}{2}$, and Super-Halley method for $\beta = 1$.

As the Newton method, these methods still require a good initial guess in the neighborhood of the root. These well-known methods provide cubic convergence order; however, they are not suggested to use since they require second-order derivative. Generally speaking, in most problems, it is attempted to avoid higher-order derivative because of some potential challenges concerning them. As stated before, methods in which higher-order derivatives are required may be cumbersome to deal. This issue can even be more problematic about a system of nonlinear equations with

multi variables. In these problems, some additional problems such as the calculation of the Jacobian matrix and the difficulty thereof may arise. Hence, the enhancement of the convergence order at the expense of the employing higher derivative is often deemed to be impracticable. As an alternative solution to obtaining higher convergence, multi-point methods can be offered. These methods can improve the convergence order without any need to employ a higher derivative. In the next part, this method is represented [19, 20].

2.2 Multi-point Methods

All methods discussed so far were based on the one-point iterative methods. As stated, these methods are of a lower convergence order; hence, to improve their convergence, multi-point methods are presented. These methods are mainly proposed to enhance the convergence order of existing methods, such as the Newton and Steffensen methods, by adding one or more steps along with employing more variables, without using higher-order derivative.

It should be noted that these methods require more function evaluations; therefore, although every added step can increase the convergence, it may not be computationally efficient. In other words, it can be said that every multi-step method is not suitable to use since some of them result in the same efficiency as other existing methods with lower computational efficiency. Therefore, as stated in previous part, the efficiency index $E.I.$ is used to determine the performance of a method; that is, the number of function evaluations used is whether or not proportional to the corresponding accuracy.

As an example of a multi-point method, consider the following algorithm based on the Newton method in three steps.

$$\begin{cases} y^{(i)} = x^{(i)} - \frac{f(x^{(i)})}{f'(x^{(i)})} \\ z^{(i)} = y^{(i)} - \frac{f(y^{(i)})}{f'(y^{(i)})} \\ x^{(i+1)} = z^{(i)} - \frac{f(z^{(i)})}{f'(z^{(i)})} \end{cases} \tag{11}$$

At first sight, it may seem an efficient method since it increases the convergence order of the Newton method to eight; but, this algorithm requires six function evaluations—three functions and three derivatives of the function—that results in $8^{1/6} = 1.414$efficiency index, which is as equal as the Newton method [10, 21]. Therefore, it should be attempted to employ or offer multi-step methods with optimum convergence order; for example, the Ostrowski method is as an improved Newton method in two steps by adding more variable as:

$$\begin{cases} y^{(i)} = x^{(i)} - \frac{f(x^{(i)})}{f'(x^{(i)})} \\ x^{(i+1)} = y^{(i)} - \frac{f(y^{(i)})(x^{(i)}-y^{(i)})}{f(x^{(i)})-2f(y^{(i)})} \end{cases} \tag{12}$$

This method increases the convergence order of the Newton method to four with employing only three function evaluations at each iteration—two functions and one first derivative evaluation. There are many other methods providing higher optimum accuracy in more steps, such as three-point methods with eighth-order convergence [22, 23] or four-point method with sixteenth-order convergence [24]. To obtain such optimum convergence order, appropriate weight functions can be employed. These weight functions can not only effectively increase the convergence order but also can be used for non-smooth functions. A through discussion of the different weight functions are beyond the scope of this text. Nevertheless, avid readers can find more detailed description in [25, 26].

There are also a number of multi-point iterative methods using the Steffensen method as the base method that may be appropriate for non-differentiable functions. Some of these derivative-free multi-point methods are given in studies [27, 28].

3 System of Nonlinear Equations

In previous parts, iterative methods for solving one-variable problems were described, but, in most real problems, one encounters more complex problems with more than one variable; in these cases, a system of nonlinear equations should be solved, which requires a different approach to one-variable problems. For example, in one-variable problems, attempts are made to provide a higher convergence order with lower function evaluations; whereas, in multi-variable problems, the goal is preventing the Jacobian matrix from being singular or ill-condition, or decreasing the computational evaluation concerning Jacobian matrix. Hence, regarding a system of nonlinear equations, to define the optimal convergence order, besides the number of function evaluations and convergence order, other factors such as Jacobian inversions, vector-vector, matrix–matrix, and matrix-vector multiplications should be considered as options affecting computational cost. Table 1 demonstrates the effect of these parameters on the total cost.

Therefore, the efficiency index of a system of nonlinear equations with n iteration can be defined in terms of the computational cost, C, as $E.I. = \rho^{\frac{1}{C}}$ [29].

Of previous methods to solve scalar nonlinear equations, only Newton and fixed-point methods can be extended for multi-variable nonlinear equations since the notations of other one-variable methods cannot be defined for this type of problems. In this part, other commonly-used methods such as Newton-like and multi-step methods are also proposed.

Table 1 Computational cost attributed to various operations used for solving a system of nonlinear equations [29]

Operations	Computational cost
LU-factorization	
Multiplications	$\frac{n(n-1)(2n-1)}{6}$
Divisions	$\frac{n(n-1)}{2}$
Total cost	$\frac{n(n-1)(2n-1)}{6} + 3\frac{n(n-1)}{2}$
Solution of lower and upper triangular systems	
Multiplications	$n(n-1)$
Divisions	n
Total cost	$n(n-1)+3n$
Scalar-vector multiplication	n
Point-wise vector-vector multiplication	n
Matrix-vector multiplication	n

3.1 Gauss–Seidel Method

As the simplest technique for solving nonlinear equations, in this method, instead of solving n nonlinear equations simultaneously, each nonlinear equation is assigned to one of the unknowns and solved to find that unknown separately. In other words, using an initial guess, say $x^{(0)} = \left[x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\right]^T$, each equation $f_j(x^{(i)})$ is considered as a one-variable equation for each variable $x_j^{(i+1)}$, then it is solved using one of the on-variable methods. The procedure of this method can be shown as:

$$\begin{aligned}
&f_1(x_1^{(i+1)}, x_2^{(i)}, x_3^{(i)}, \ldots, x_{n-1}^{(i)}, x_n^{(i)}) = 0\\
&f_1(x_1^{(i+1)}, x_2^{(i+1)}, x_3^{(i)}, \ldots, x_{n-1}^{(i)}, x_n^{(i)}) = 0\\
&\qquad\qquad\vdots\\
&f_1(x_1^{(i+1)}, x_2^{(i+1)}, x_3^{(i+1)}, \ldots, x_{n-1}^{(i+1)}, x_n^{(i+1)}) = 0
\end{aligned} \tag{13}$$

If the new variables are not updated for the successive equations, the method converts to Gauss-Jacobi algorithm.

Generally, there are no scheme explaining which sequence should be considered for each equation and variable; nonetheless, it is preferable that if an equation contains only one variable, that equation and variable be considered as the first one. Both methods converge considerably slower than other methods. Additionally, this method may not work efficiently if the system is not diagonally dominant or positive-definite [30].

3.2 Fixed-Point Method

As the one-dimensional fixed-point method, initially, each equation is assigned to one variable; then, every equation is split and solve for that assigned variable.

$$\begin{aligned} x_1^{(i+1)} &= g_1(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \ldots, x_n^{(i)}) \\ x_2^{(i+1)} &= g_2(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \ldots, x_n^{(i)}) \\ &\vdots \\ x_3^{(i+1)} &= g_3(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \ldots, x_n^{(i)}) \end{aligned} \tag{14}$$

To simply accelerate the convergence order of this method, each new estimation from previous functions should be replaced in the next functions [8, 30].

It should be stressed that both fixed-point and Guess-Seidel methods neither require the derivative of a function nor involve high-computational complexities; besides, these methods can provide a good rough estimation for complicated problems. However, these methods can be applied for simple problems since these methods give a low convergence order. In other words, if one encounters problems demanding lengthy calculation of functions per iteration, the following faster methods are suggested in order to obtain the final result in fewer iteration [8, 30].

3.3 Newton Method

As the one-dimensional Newton method, by using multi-variable Taylor series expansion, the well-known Newton method for a system of nonlinear equations can be gained in a matrix form as:

$$F(x^{(i+1)}) = F(x^{(i)}) + J(x)(x^{(i+1)} - x^{(i)}) \tag{15}$$

in which $J^{(i)}$ is the n-dimensional Jacobian matrix for n-dimensional column vector of $F(x)$ and x is the column vector of the unknowns as:

$$x^{(i)} = [x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}]^T \tag{16}$$

$$F(x) = [f_1(x), f_2(x), \ldots, f_n(x)]^T \tag{17}$$

$$J^{(i)} = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1^{(i)} & \cdots & \frac{\partial}{\partial x_n} f_1^{(i)} \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_n^{(i)} & \cdots & \frac{\partial}{\partial x_n} f_n^{(i)} \end{bmatrix} \tag{18}$$

where $\frac{\partial}{\partial x}$ is the partial derivative with respect to the unknowns of the problem.

As stated previously, despite being quadratically convergent, this method has many drawbacks, even more than the one-dimensional case. This method is too costly since numerous function evaluations, and the derivative thereof should be calculated at each iteration, in particular, for nonlinear systems with a large matrix. More importantly, the evaluation of the Jacobian matrix J per iteration can be a challenging task with an excess amount of time for each process; during the process, the Jacobian matrix may become ill-conditioned or singular, which cause the iteration procedure to diverge and influence the stability of this method. Another problem concerning this method is that the iteration process is highly sensitive to an accurate initial approximation to ensure convergence and control progress path towards the root of the function, even more than a one-variable case. Hence, this method may deem to be an expensive method to employ. Keeping this in mind, there are many remedies proposed to overcome some obstacles regarding these issues and enhance the efficiency of this method. In the next parts, the most important ones are discussed [31, 32].

3.4 Quasi-Newton Method

To release difficulties regarding the calculation of the Jacobian matrix for each iteration, quasi-Newton methods have been developed to estimate the Jacobian matrix at each iteration based on the previous steps.

The Broyden technique as a generalized secant method for multi-variable problems is a quasi-Newton method that can be used as an alternative method with a less computational cost. This method employs the Jacobian matrix for only the initial guess. Then, using the current approximation of the Jacobian matrix J_{n-1}, this method updates it at each iteration by the Eq. (19):

$$J^{(i+1)} = J^{(i)} + \frac{F^{(i)} - J^{(i)}x^{(i)}}{\left|x^{(i)}\right|^2}\left|x^{(i)}\right|^T \tag{19}$$

where $\Delta x^{(i)} = x^{(i+1)} - x^{(i)}$ and $\Delta F^{(i)} = F^{(i+1)} - F^{(i)}$. In the next step, the successive approximation of the solution can be obtained by substituting the new estimate of the Jacobian matrix in the Newton method. This approximation significantly decreases the computational cost and provides a superlinear convergence [33]. Moreover, unlike the Newton method, these methods are not self-correcting; that is, round-off errors are accumulated during successive iterations.

This technique, using the approximation of the Jacobian matrix instead of the actual value of it, has been developed by other researchers; the correlation method, Davidon-Fletcher-Powell formula, and Broyden-Fletcher-Goldfarb-Shanno method for example, are another method to estimate the Jacobian matrix instead of directly

calculate it. To obtain more insight into these methods, interested readers are referred to [34–36].

3.5 Steepest Descent Method

An effective way to solve a system of nonlinear equations is by transforming a root-finding problem to an optimization problem so that the global optimization methods can be used to solve the problem [37]. One of the well-known methods based on this idea is the first-order steepest descent method or gradient ascent method that ensures convergence even with a poor initial guess for a system of nonlinear equations. This method converts the system of nonlinear equations to a problem of minimizing the sum of the square of all nonlinear functions, $G(x)$, as:

$$G(x) = \sum_{n=1}^{m} f_n^2(x) \tag{20}$$

To calculate the minimum of the function $G(x)$, or zeros of nonlinear functions, this method moves in a direction in which the value of $G(x)$ decreases. To determine this direction, the property of the gradient of a function and the extreme value theorem—the maximum of a function occurs when is parallel to ∇G are employed. Accordingly, this method computes $-\nabla G$ as the direction in which the most decrease in the value of nonlinear functions occurs. Additionally, per Eq. (21), to increase the efficiency of the method, it is attempted to add and select the best value for α, which can improve the speed of the method towards the root.

$$x^{(i+1)} = x^{(i)} - \alpha\left(\nabla G\left(x^{(i)}\right)\right) \tag{21}$$

To find the best value for α resulting in faster convergence, readers are referred to the Ref. [8].

The speed of this algorithm is still slow. On the other hand, the calculation of the gradient of the sum of the square of nonlinear functions is a costly process for each iteration. As a consequence, this method may not be served as a computationally efficient method since it does not reduce the amount of calculation. Nevertheless, this method is suggested to use in combination with other methods, which are sensitive to the starting point, since it does not depend on the initial guess and can provide a good rough initial guess. As in the one-dimensional case, this method is similar to the bisection method used for a scalar equation [38].

3.6 Leven-Marquardt Method

Like the previous method, the Leven-Marquardt method is used to optimize the sum of the square of nonlinear functions, to find zeros of the system of nonlinear equations; except, it estimates the functions in terms of their Taylor expansion. In this method, firstly, the function of S in terms of the sum of the square of nonlinear functions is defined as:

$$S^2(x) = \sum_{n=1}^{m} [f(x)]^2 \tag{22}$$

By using the Taylor expansion (Eq. (16)) of the function $f(x)$, the value of S^2 can be yielded as:

$$S^2(x) = F(x)^T F(x) - 2F^T(x)J(x)h + J^T(x)J(x)h^2 \tag{23}$$

where h is the incriminate value equals $h = x^{(i+1)} - x^{(i)}$. Therefore, the minimum value of S^2 can be obtained by setting the derivative of the square of the nonlinear functions with respect to h equal to zero as:

$$\frac{\partial}{\partial h} S^2(x) = -2F^T(x)J(x) + 2h^T J^T(x)J(x) = 0 \tag{24}$$

where $J^T(x)J(x)$ and $J^T f(x)$ is the Hessian matrix, $H(x)$, and the gradient of $F(x)$, ∇F, respectively. Thus, by rewriting the Eq. (24), the Gauss–Newton iteration algorithm can be gained as:

$$x^{(i+1)} = x^{(i)} + H^{-1}(X)\nabla F \tag{25}$$

The above-mentioned equation, which is called Gauss-Newton method, can only be used when the initial guess is in the vicinity of the roots of the system of nonlinear equations. Therefore, to resolve this issue, the Levenberg–Marquardt method can be replaced. According to Eq. (26), by adding λ as a damping parameter varying between 0 and 1, this method can interpolate between two methods of steepest descent method and Gauss–Newton method. In other words, if the current solution is far from the actual response, the algorithm becomes slow like the steepest descent method and the value of λ is set to be large; while, when the approximated solution is near the accurate response, the speed of the algorithm increases, approaching a Gauss–Newton method. In this case, the value of λ is set to be small.

$$x^{(i+1)} = x^{(i)} - [H(x) + \lambda I]^{-1}\nabla F \tag{26}$$

Note that, in the Levenberg algorithm, if the value of the damping factor becomes large, the Hessian matrix is never used. To avoid this problem, the process of scaling can be done based on the curvature of equations so that where the gradient is small (low curvature), a small value for the damping parameter should be selected to provide a larger movement. The improved Levenberg algorithm, Eq. (27), connects the movement to curvature as [39–41]:

$$x^{(i+1)} = x^{(i)} - [H(x) + \lambda diag(H(x))]^{-1} \nabla F \tag{27}$$

The great advantage of this method over the Newton method is finding all potential roots of a nonlinear system without any requirement for a good initial guess.

3.7 Multi-step Method

This method is an efficient and rapid technique with wide applications in curve-fitting, approximating the derivative and integral of a function, and solving nonlinear equations. Generally speaking, in multi-step methods, the information of previous methods is retained to approximate the value of the next step. Regarding solving nonlinear equations, multi-step methods are of a more prominent place since it can lead to a significant decrease in computational time and effort emanating from computing Jacobian matrix per iteration.

The basic idea behind this method is using the Jacobian matrix obtained in previous steps for the successive steps. As a consequence, the time-consuming process of calculation of the Jacobian matrix and the lower-upper (LU) factorization in each iteration is simply removed by following a method without updating the Jacobian matrix at every single step, which is called the frozen Jacobian method. It can be stated that this technique can convert an existing one-point iterative method without memory, like the Newton method, to a multi-step method with memory.

Multi-step iterative techniques comprise two parts, namely base part and multi-step part. In the base part, the Jacobian matrix and its LU-factorization are calculated and kept for the next part. By reusing the information of the base-part, multi methods solve a system of linear equations. For example, the multi-step Newton method is shown as [42–44]:

$$\text{Base Newton Method} \begin{cases} y^0 \text{ initial guess} \\ F'\left(y^0\right)x^{(i+1)} = F\left(y^{(i)}\right) \\ y^{(i+1)} = y^{(i)} - x^{(i+1)} \end{cases}$$

$$\text{Multistep Newton Method} \begin{cases} \text{for } i = 1 : m-1 \\ F'\left(y^0\right)x^{(i+1)} = F\left(y^{(i)}\right) \\ y^{(i+1)} = y^{(i)} - x^{(i+1)} \\ \text{end} \\ y^{(0)} = y^{(m)} \end{cases} \tag{28}$$

In the multi-step Newton method, for $m \geq 1$ steps, the convergence order of the Newton method is increased to $m + 1$, which means that every additional step increases the convergence order by one. This algorithm may not seem cost-efficient, but, only one evaluation of the Jacobian matrix for one cycle can justify the wide-application of this method, in particular, for a large nonlinear system with sparse matrix. In contrast, for scalar nonlinear equations, this method is not recommended since it does not have the optimal convergence order [6, 45].

Some improvements concerning increasing the convergence of multi-step methods have been made. In these methods, it has become feasible to obtain higher convergence order than $m + 1$ by using the idea of multi-variable iterative methods. More details on these algorithms are given in studies [46, 47].

It has been demonstrated that this method, in spite of being efficient from a computational viewpoint, still requires the derivative of the function, which is not available for every problem. Hence, in such cases, derivative-free methods, which are represented below, can be adopted.

3.8 Picard Method

The main idea of the Picard or direct iteration method is splitting the nonlinear equations into two linear and nonlinear parts to reduce the size of computational work. In this case, the expensive operator of LU decomposition is only applied to the linear part. Picard method is a derivative-free technique of the decomposition technique for solving a system of nonlinear algebraic equations in engineering disciplines containing large scale problems. This method arranges the system of nonlinear equations as:

$$K(x)\mathrm{x} = \mathrm{F} \tag{29}$$

In which x is the vector column of unknowns, K and F are nonlinear coefficient matrix and the column vector of constants of nonlinear equations, respectively. The simple iterative method to solve this equation is as:

$$\{x\}^i = \left\{K(\{x\}^{(i-1)})\right\}^{-1}\{F\} \tag{30}$$

However, according to the above formulation, as all nonlinear equations are solved simultaneously, the coefficient matrix containing nonlinear part should be inversed at each iteration, which is computationally expensive. Therefore, to solve this issue, the decomposition technique can be effectively used by splitting the coefficient matrix into two parts of linear K_L and nonlinear $K_N(x)$ as:

$$K(x) = K_L + K_N(x) \tag{31}$$

Instead of inverting of $K(x)$ in Eq. (30), the alternative Eq. (32) can be followed. In this equation, the operation of inverse is only applied to the linear part. As a consequence, the computational time can considerably decrease.

$$\{x\}^{(i)} = (K_L)^{-1}\left[F - K_N(\{x\}^{(i-1)}).\{x\}^{(i-1)}\right] \tag{32}$$

To accelerate the convergence of this method, an auxiliary relaxation parameter γ, which varies between 0 and 1, can be added. By using this parameter, the new estimation can be substituted by a weighted average of the last two approximations as [2]:

$$\{\overline{x}\}^{(i)} = \gamma\{x\}^{(i-2)} + (1-\gamma)\{x\}^{(i-1)} \tag{33}$$

It should be pointed out that in this method, this method is of a slow speed; besides, it is likely that the decomposition process results in the non-singularity of the coefficient matrix.

3.9 Newton–Krylov Method

Inexact Newton methods are practical methods to approximate the resulting equation of Newton method, instead of exactly solving it. The algorithm of this method can be shown as:

$$\left\| J(x^{(i)})s^{(i)} + F(\mathrm{x}^{(i)}) \right\| \leq \eta^{(i)} \left\| F(x^{(i)}) \right\| \tag{34}$$

where $\eta^{(i)}$ is the forcing term of the i-th iteration varying between [0, 1) and $S^{(i)}$ is the inexact Newton step ($x^{(i+1)} = x^{(i)} + S^{(i)}$). In this method, initially, a suitable value for $\eta^{(i)}$ is selected; then, the linear equation is approximately solved for $S^{(i)}$. Some conventional methods to solve the resulting linear equations are classical splitting methods or the modern Krylov subspace methods. If widely-used Krylov subspace iteration methods are employed for solving the inexact Newton step, this method is called the Newton–Krylov subspace method [48].

The determination of forcing term in this method is of great importance that directly affects the accuracy and convergence of the method as the right-hand side expression is both a linear model of the system and the residual of the Newton method. Therefore, a suitable choice of $\eta^{(i)}$ results in a reduction in the norm of the model along with an increase in the accuracy of the model. For example, for $i = 0$ inexact Newton method converts to the quadratically convergent Newton method [48–50]. To explore some strategies for finding the best value for $\eta^{(i)}$, readers can be referred to [51, 52].

In comparison with the previous method, this method is deemed the most efficient tool for solving a system of nonlinear equations with a large sparse matrix. This

method only involves matrix–vector products instead of expensive operations like the inverse Jacobian matrix and LU decomposition thereof. Hence, this method decreases the computational cost regarding these operations. In various studies, it is shown that for small size problems, although this method cannot compete with other higher order Newton methods; for problems with large sparse matrix, this method considerably decreases the computational cost. Note that, as the Newton method, this method is locally convergent and require a globalization method. One effective method to enhance the global convergence of this method is using the backtracking technique—shortening the interval for an unsatisfactory step.

A downside of this method can be attributed to the calculation of the Jacobian matrix. Hence, for problems in which constructing the Jacobian matrix is difficult, the Jacobian free Newton–Krylov method can be replaced. By contrast with other methods mentioned in previous parts, this method does not calculate or store the Jacobian matrix, even for the initial approximation; instead, an approximation of Jacobian-vector product is constructed using the finite difference method. The step of this finite difference method can significantly affect the accuracy of the Jacobian-vector product [53, 54].

4 Discussion

In this part, an overall comparative explanation of the described methods, as well as some examples of the practical usage of these methods in different disciplines, are presented. As stated previously, the numerical methods for solving scalar equations can be described in two separate parts: one-point methods and multi-point methods. Accordingly, one-point methods are the most straightforward methods offering lower convergence together with ease of implementation. Among different one-point methods, Newton and Steffensen methods are two one-point methods, which are the basis of more advanced approaches, with the optimum convergence order. As these methods require a good initial guess, it is proposed to combine these methods with other globally convergent methods such as the bisection method. Chebyshev and Halley methods are other one-variable methods offering the cubic convergence order by employing the second derivative of a function; however, owing to some difficulties for computing a higher-order derivative, multi-point methods may be better alternatives increasing the convergence order. Using auxiliary variables in more steps, multi-point techniques present higher accuracy for the existing methods. It should be pointed out that every multi-point method with a higher convergence rate cannot deem an efficient method since need more function evaluations with the same efficiency as the previous methods.

Regarding the nonlinear system of equations, Gauss–Seidel, and fixed-point methods as derivative-free algorithms are just suitable for simple problems since they have slow convergence speed. The well-known Newton method is another method with quadratic rate of convergence. This method has some limitations, such

as needing a good initial guess, the time-consuming process of matrix–matrix evaluation, and the evaluation of the Jacobian matrix and its inverse. Nevertheless, this method has widely incorporated with a lot of mathematical software and computational simulation. One of the main problems in using Newton methods is the costly evaluation of the Jacobian matrix per iteration. To overcome such an issue, the estimation of this matrix, instead of directly computing, is suggested by Quasi-Newton methods for problems in which the derivatives of the function are not available, or the calculation of them is time-consuming. The quasi-Newton methods have also been used to solve Maxell's equations and Navier–Stokes equations [5].

Multi-step methods with frozen Jacobian are another technique that can be a remedy to the costly evaluation of the Jacobian matrix per iteration with increasing the convergence order. As scalar equations, the multi-point techniques can also enhance the convergence and effectiveness of these methods. Multi-step methods have also successfully been adopted for solving nonlinear Poisson, heat conduction, and wave equations [29, 42, 43]. Another way to solve multi-variable nonlinear equations is using global optimization methods to find all zeros of a nonlinear system, even with a poor initial guess. Despite being efficient, most of these methods require the computation of derivatives of the functions. These methods are also popular methods that have extensively incorporated in mathematical software. Picard method based on the decomposition technique is also defined as a derivative-free and straightforward method; it, though, may show instability for complex problems. As one of the commonly-used methods, the Newton-like methods can provide lower computational time for problems with a large sparse matrix. Of various techniques, Newton–Krylov method, as a Newton-like method, may be the best option in solving a complicated nonlinear system. The Jacobian-free Newton–Krylov method has also been employed to solve a wide range of problems, including radiation diffusion, Bratu, Navier–Stokes, and Maxwell's equations [49, 50]. Moreover, this method has been incorporated in many computational and simulation software. For the sake of comparison, the numerical characteristics of these numerical methods are also given in the Table 2. To the best of our knowledge, this chapter reviewed some general iterative techniques to solve nonlinear equations. However, there are other numerous developed techniques that are an improved version of the existing methods or a combination of them.

There are also some tangible cases displaying the application and importance of the discussed methods in real-world applications. As some notable examples, the fixed-point method has been extensively employed for solving the magnetic hysteresis field problems [56, 57]. The gradient-family methods have also been used in microwave imaging applications [55, 58, 59]. In the application of piezoelectric material as an energy harvester, the Newton and Picard method is also adopted to address the nonlinear behavior of structures with piezoelectric material actuators [60, 61]. In computational plasma physics, the Newton–Krylov methods are also employed to solve many diverse cases; these applications are all given in a study by Knoll and Keyes [49]. It should be pointed out that the aforementioned examples are only a limited number of cases demonstrating broad applications of iterative methods in energy applications. Obviously, given the potential capability of each technique,

Table 2 A comparative study of numerical methods for solving a nonlinear system of equations

Methods	Derivative free	Convergence speed	Globally convergent	LU factorization
Gauss–Seidel	Yes	Slow	No	0
Fixed point	Yes	Slow	No	0
Newton	No	Fast	No	Number of iterations
Quasi-Newton	No[a]	Slow	No	0
Steepest descent	No	Slow	Yes	Number of iterations
Leven-Marquardt	No	Fast	Yes	Number of iterations
Multi-Step	No[a]	Fast	Yes/No[b]	≥ 1
Picard	Yes	Slow	No	1
Newton–Krylov	No[a]	Fast	No	0

[a]The derivative is only needed for the initial guess
[b]It depends on the algorithm

these iterative methods can be adopted for solving other types of applications and equations.

5 Conclusion

Nonlinear analysis is an essential part of every discipline. In some applications, nonlinear solutions should be unavoidably employed to address some nonlinear phenomena, such as the dynamic behavior of wind turbines, the analysis of the damage within the structures, or the behavior of some novel materials. Keeping the leading role of the nonlinear solution methods in mind, this chapter provides an overview of some fundamental numerical methods and their attributes for solving the nonlinear discretized equations. Having dealt with some iterative numerical methods, in this part, this chapter is enclosed with concise results gained through this investigation, as well as some comments on future and ongoing directions in the development of iteration methods for solving nonlinear discretized problems.

To sum up, it may be concluded that for selecting a proper method for solving nonlinear equations, the equations themselves and the final accuracy always should be examined in selecting a method since employing some complex solution methods may seem unnecessary. On the other hand, there are some general factors specifying the best efficient technique for a problem. In dealing with scalar nonlinear equations, the number of call functions, convergence order, initial guess, and smoothness are determining factors in selecting a method. Regarding a system of nonlinear equations together with the above factors, other criteria should be considered, in particular, if

one is dealing with a complex system of nonlinear equations. For example, matrix–matrix operation, vector–matrix operation, the evaluation of Jacobian matrix and its inverse are other factors that are critical in choosing the best method.

Further development of more efficient and robust methods for nonlinear equations can be directed in the following way. Regarding scalar iterative methods, more efficient methods can be developed by establishing multi-point iterative methods with optimum convergence. Many ongoing researches are focusing on finding a suitable weight function or involved free parameter for the existing method that provides optimum convergence order.

With respect to iterative methods for solving the nonlinear system of equations, these solution methods have a prominent place in the foreseeable future due to the extensive application of them in advanced design and research as well as many intricacies regarding these methods. For example, concerning multi-step methods, there are still numerous continuing researches offering the optimal and higher convergence order. Moreover, developing some techniques to smooth or globalize the existing method and prevent them from diverging is the area of interest in this field. Another promising place can be related to Newton–Krylov methods. According to the merits and potential applications of these methods, a considerable number of researches during the very recent years have been focused on this topic; this area is still in progress. As the last remark, the combination of methods can assist researchers in enjoying the positive attributes of each method, as many techniques have been developed only by combining different existing methods.

References

1. Wang J (2018) A new iterative method for solving nonlinear equation. Dissertation, University of Saskatchewan
2. Reddy JN (2014) An introduction to nonlinear finite element analysis: with applications to heat transfer, fluid mechanics, and solid mechanics. OUP, Oxford
3. Wong EC (2014) Numerical methods for nonlinear systems of equations. Dissertation, Universiti Teknologi Malaysia
4. Mohammadi N, Asadi H, Aghdam MM (2019) An efficient solver for fully coupled solution of interaction between incompressible fluid flow and nanocomposite truncated conical shells. Comput Methods Appl Mech Eng 351:478–500
5. Zienkiewicz OC, Taylor RL, Zhu JZ (2005) The finite element method: its basis and fundamentals. Elsevier
6. Malik ZU (2015) Numerical iterative methods for nonlinear problems. Dissertation, Università degli Studi dell'Insubria
7. Maheshwari AK (2009) A fourth order iterative method for solving nonlinear equations. Appl Math Comput 211(2):383–391
8. Burden RL, Faires JD (2005) Numerical analysis, 8th ed. Thompson Brooks/Cole
9. Traub JF (1982) Iterative methods for the solution of equations, vol 312. American Mathematical Soc
10. Petković MS, Neta B, Petković LD et al (2014) Multipoint methods for solving nonlinear equations: a survey. Appl Math Comput 226:635–660
11. Soleymani F, Lotfi T, Bakhtiari P (2014) A multi-step class of iterative methods for nonlinear systems. Optim Lett 8(3):1001–1015

12. Yamamoto T (2000) Historical developments in convergence analysis for Newton's and Newton-like methods. J Comput Appl Math 124(1–2):1–23
13. Pooley D (2003) Numerical methods for nonlinear equations in option pricing. Dissertation, University of Waterloo
14. Janem GA (2015) Solving nonlinear boundary value problems using the homotopy analysis method. Dissertation, United Arab Emirates University
15. Gerald CF (2004) Applied numerical analysis. Pearson Education India
16. Alefeld G (1981) On the convergence of Halley's method. Am Math Mon 88(7):530–536
17. Argyros IK, Szidarovszky F (2018) The theory and applications of iteration methods. CRC Press
18. Cordero A, Torregrosa JR (2011) A class of Steffensen type methods with optimal order of convergence. Appl Math Comput 217(19):7653–7659
19. Kou J, Li Y (2007) The improvements of Chebyshev-Halley methods with fifth-order convergence. Appl Math Comput 188(1):143–147
20. Chun C (2007) Some second-derivative-free variants of Chebyshev-Halley methods. Appl Math Comput 191(2):410–414
21. Wang X, Liu L (2010) Modified Ostrowski's method with eighth-order convergence and high efficiency index. Appl Math Lett 23(5):549–554
22. Soleymani F, Sharifi M, Mousavi BS (2012) An improvement of Ostrowski's and King's techniques with optimal convergence order eight. J Optim Theory Appl 153(1):225–236
23. Thukral R, Petković MS (2010) A family of three-point methods of optimal order for solving nonlinear equations. J Comput Appl Math 233(9):2278–2284
24. Sharifi S, Salimi M, Siegmund S et al (2016) A new class of optimal four-point methods with convergence order 16 for solving nonlinear equations. Math Comput Simul 119:69–90
25. Zafar F, Cordero A, Torregrosa JR (2019) Stability analysis of a family of optimal fourth-order methods for multiple roots. Numer Algorithms 81(3):947–981
26. Artidiello S, Cordero A, Torregrosa JR et al (2015) Design of high-order iterative methods for nonlinear systems by using weight function procedure. Abstr Appl Anal. https://doi.org/10.1155/2015/289029
27. Soleymani F (2012) Optimized Steffensen-type methods with eighth-order convergence and high efficiency index. Int J Math Math Sci. https://doi.org/10.1155/2012/932420
28. Zheng Q, Li J, Huang F (2011) An optimal Steffensen-type family for solving nonlinear equations. Appl Math Comput 217(23):9592–9597
29. Qasim S, Ali Z, Ahmad F et al (2016) Solving systems of nonlinear equations when the nonlinearity is expensive. Comput Math Appl 71(7):1464–1478
30. Sadiku MN (2000) Numerical techniques in electromagnetics. CRC press
31. Ahmad F, Tohidi E, Carrasco JA (2016) A parameterized multi-step Newton method for solving systems of nonlinear equations. Numer Algorithms 71(3):631–653
32. Nili Ahmadabadi M, Ahmad F, Yuan G et al (2016) Solving systems of nonlinear equations using decomposition technique. J Linear Topol Algebra 5(03):187–198
33. Kelley CT (1995) Iterative methods for linear and nonlinear equations. Siam
34. Shademan A, Farahmand AM, Jägersand M (2010) Robust Jacobian estimation for uncalibrated visual servoing. In: 2010 IEEE International conference on robotics and automation, Anchorage, AK, USA, May 2010. IEEE, p 5564
35. Lapresté JT, Jurie F, Dhome M et al (2004) An efficient method to compute the inverse Jacobian matrix in visual servoing. In: 2004 IEEE International conference on robotics and automation, New Orleans, LA, USA, USA, April 2004. ICRA'04, vol 1. IEEE, p 727
36. Mansard N, Lopes M, Santos-Victor J et al (2006) Jacobian learning methods for tasks sequencing in visual servoing. In: 2006 IEEE/RSJ International conference on intelligent robots and systems, Beijing, China, October 2006. IEEE, p 4284
37. Tsoulos IG, Stavrakoudis A (2010) On locating all roots of systems of nonlinear equations inside bounded domain using global optimization methods. Nonlinear Anal Real World Appl 11(4):2465–2471

38. Johnson R, Zhang T (2013) Accelerating stochastic gradient descent using predictive variance reduction. In: Advances in neural information processing systems, vol 26. Neural information processing systems, Lake Tahoe, Nevada, December 2013, p 315
39. Ranganathan A (2004) The Levenberg-Marquardt algorithm. Tutoral LM Algorithm 11(1):101–110
40. Shukla PK (2010) Levenberg-Marquardt algorithms for nonlinear equations, multi-objective optimization, and complementarity problems. Dissertation, Dresden University of Technology
41. Moré JJ (1978) The Levenberg-Marquardt algorithm: implementation and theory. Numerical analysis. Springer, Heidelberg
42. Ilyas I, Ali Z, Ahmad F et al (2017) Multi-step frozen Jacobian iterative scheme for solving IVPs and BVPs based on higher order Fréchet derivatives. J Math 49(1):125–137
43. Kouser S, Rehman SU, Ahmad F et al (2018) Generalized newton multi-step iterative methods GMN p, m for solving system of nonlinear equations. Int J Comput Math 95(5):881–897
44. Shah FA, Noor MA (2015) Some numerical methods for solving nonlinear equations by using decomposition technique. Appl Math Comput 251:378–386
45. Kung HT, Traub JF (1974) Optimal order of one-point and multipoint iteration. J ACM (JACM) 21(4):643–651
46. Alzahrania EO, Al-Aidarousa ES, Younasa AM et al (2016) A higher order frozen Jacobian iterative method for solving Hamilton-Jacobi equations. Convergence 1000:1
47. Montazeri H, Soleymani F, Shateyi S et al (2012) On a new method for computing the numerical solution of systems of nonlinear equations. J Appl Math. https://doi.org/https://doi.org/10.1155/2012/751975
48. Shin BC, Darvishi MT, Kim CH (2010) A comparison of the Newton-Krylov method with high order Newton-like methods to solve nonlinear systems. Appl Math Comput 217(7):3190–3198
49. Knoll DA, Keyes DE (2004) Jacobian-free Newton-Krylov methods: a survey of approaches and applications. J Comput Phys 193(2):357–397
50. An HB, Wen J, Feng T (2011) On finite difference approximation of a matrix-vector product in the Jacobian-free Newton-Krylov method. J Comput Appl Math 236(6):1399–1409
51. An HB, Mo ZY, Liu XP (2007) A choice of forcing terms in inexact Newton method. J Comput Appl Math 200(1):47–60
52. Gomes-Ruggiero MA, Lopes VLR, Toledo-Benavides JV (2008) A globally convergent inexact Newton method with a new choice for the forcing term. Ann Oper Res 157(1):193–205
53. Eisenstat SC, Walker HF (1994) Globally convergent inexact Newton methods. SIAM J Optim 4(2):393–422
54. Chen Y, Shen C (2006) A Jacobian-free Newton-GMRES (m) method with adaptive preconditioner and its application for power flow calculations. IEEE Trans Power Syst 21(3):1096–1103
55. Mojabi P, LoVetri J (2009) Overview and classification of some regularization techniques for the Gauss-Newton inversion method applied to inverse scattering problems. IEEE Trans Antennas Propag 57(9):2658–2665
56. Dlala E, Arkkio A (2008) Analysis of the convergence of the fixed-point method used for solving nonlinear rotational magnetic field problems. IEEE Trans Magn 44(4):473–478
57. Dlala E, Belahcen A, Arkkio A (2007) Locally convergent fixed-point method for solving time-stepping nonlinear field problems. IEEE Trans Magn 43(11):3969–3975
58. Rubæk T, Meaney PM, Meincke P et al (2007) Nonlinear microwave imaging for breast-cancer screening using Gauss–Newton's method and the CGLS inversion algorithm. IEEE Trans Antennas Propag 55(8):2320–2331
59. De Zaeytijd J, Franchois A, Eyraud C et al (2007) Full-wave three-dimensional microwave imaging with a regularized Gauss-Newton method—theory and experiment. IEEE Trans Antennas Propag 55(11):3279–3292
60. Komijani M, Reddy JN, Eslami MR (2014) Nonlinear analysis of microstructure-dependent functionally graded piezoelectric material actuators. J Mech Phys Solids 63:214–227
61. Yang J, Kitipornchai S, Liew KM (2004) Non-linear analysis of the thermo-electro-mechanical behaviour of shear deformable FGM plates with piezoelectric actuators. Int J Numer Methods Eng 59(12):1605–1632

Theoretical Approach to Element Free Galerkin Method and Its Mathematical Implementation

Bhaumik Nagevadiya, Rameshkumar Bhoraniya, Ramdevsinh Jhala, and Rajendrasinh Jadeja

Abstract Numerical methods such as FVM, FDM, FVM, and BVM are eminent for solving the physical problems in engineering and science. Mentioned numerical methods are based on the predefined topological map, generally called "mesh," Meshes are required to establish the relations between nodes, which becomes vital for the creation of shape functions. The problems with mesh-based methods are (i) They require the qualitative mesh, which is a somewhat tedious, time-consuming & messy task (ii) Meshing & re-meshing for a sizeable computational domain is time consuming, tedious, and costly task also requires the skills (iii) In complex computational domains, the mesh-based method fails in terms of accuracy (iv) Glass hour and shear locking phenomena generally found in the traditional finite element method. In the last two and a half decades, many engineers and mathematicians have proposed a new class of numerical methods known as meshfree methods. Meshfree methods are independent of mesh and approximate the governing PDE based on the set of nodes only. This chapter seeds light on an eminent meshfree method called EFGM. Chapter deals with the introduction and background of meshfree methods, the EFGM method, and its mathematical formulations. The chapter also comprises two elastostatic numerical problems, the 1D problem of a bar with body forces and 2D Timoshenko cantilever beam with traction at the tip, numerical results have been evaluated & compared with exact results. The convergence of both 1D and 2D problems have been discussed. This work built a sound foundation on EFGM and will

B. Nagevadiya · R. Bhoraniya · R. Jhala · R. Jadeja (✉)
Marwadi University, Rajkot, India
e-mail: rajendrasinh.jadeja@marwadieducation.edu.in

B. Nagevadiya
e-mail: nagevadiyabhaumik111@gmail.com

R. Bhoraniya
e-mail: rameshkumar.bhoraniya@marwadieducation.edu.in

R. Jhala
e-mail: ramdevsinh.jhala@marwadieducation.edu.in

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_6

act as a stepping stone for novices in the field of meshfree methods. Keywords: Advanced numerical approach, Mesh-free methods, Element free Galerkin method

Nomenclatures

K	Global stiffness matrix
U	Vector of nodal parameter
F	Global force vector
a	Coefficient vector in moving least square method
p	Basis vector
W(x)	Weight function
A	Moment matrix in moving least square method
t_Γ	Traction on natural boundary (Prescribed)
u_Γ	Displacement on essential boundary (Prescribed)
b	Body force vector
L_d	Matrix differential operator
σ	Stress components

List of Acronym

EFGM	Element free Galerkin method
SPH	Smoothed particles hydrodynamics
PDE	Partial differential equation
MLSM	Moving least square method
FEM	Finite element method
FDM	Finite difference method
FVM	Finite volume method
BVM	Boundary value method
RKPM	Reproducing kernel particle method
MLPG	Meshless local Petrov Galerkin method
LMM	Lagrange multiplier method
RBF	Radial basis functions

1 Introduction

FEM is very popular with the success in engineering and industries to solve complex field problems. The finite element method uses mesh interpolation; in the case of a complex problem domain and large deformations, it leads to error, especially for computing stress and cumbersome remeshing, which consumes more time and money also requires more computational efforts. Even after remeshing and modifying, a method is not sure to give accurate results [1]. Meshfree methods have inherent properties to eliminate the finite element method's difficulties, which is a dependency on a mesh. Meshfree methods are entirely independent of mesh. The shape function construction is performed by the only set of nodes [2].

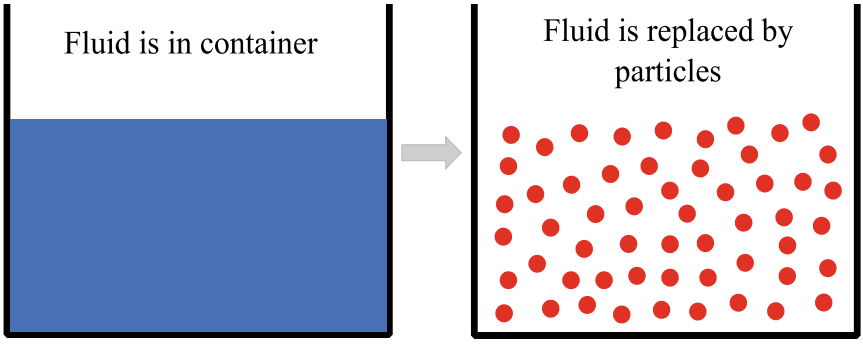


Fig. 1 Schematic diagram of smooth particle hydrodynamics

The invention of the meshfree method was date back from the year 1977, R. A. Gingold, J. J. Monaghan and L.B. Lucy in 1977 proposed the first meshfree method named Smoothed Particle Hydrodynamics. This method is a Lagrangian based method [3]. The critical idea in SPH is to replace the confined volume of fluid with finite particles (Fig. 1) and transforming governing partial differential equations into kernel estimation integrals [4, 5]. The method was developed to model Astrophysics problems. SPH lately widely used for more applications, i.e., continuum mechanics, biomechanics & fluid mechanics.

FEM is very popular with the success in engineering and industries to solve complex field problems. The finite element method uses mesh interpolation; in the case of a complex problem domain and large deformations, it leads to error, especially for computing stress and cumbersome remeshing, which consumes more time and money also requires more computational efforts. Even after remeshing and modifying, a method is not sure to give accurate results [1]. Meshfree methods have inherent properties to eliminate the finite element method's difficulties, which is a dependency on a mesh. Meshfree methods are entirely independent of mesh. The shape function construction is performed by set of nodes only.

In SPH, particles within the smoothing radius *Kh are* considered. To reduce the contribution of the other particles and computation time, where K and h are the constant and smoothing length, respectively. Heart of smoothed particle hydrodynamics is the Monte Carlo technique (Statistical technique) to reduce computational requirements [2]. Nayroles et al. [6] suggested Diffused Element Method to overcome two difficulties of traditional finite element method (FEM), the first difficulty is a discontinuity of derivative at element boundaries (it spoils the accuracy of stress) and the second one is to maintain mesh generation quality that is quite difficult and time-consuming. The key concept in the diffused element method is to replace the interpolation of the traditional finite element method by local weighted least square fittings.

Belytschko et al. [7] refined the diffused element method and proposed a new method in 1994 based on the global weak form called EFGM. The moving least

square method [8] was used for trial and test functions. The essential boundary condition in EFGM was imposed by LMM [9] because the MLSM does not satisfy the Kronecker delta function property. EFGM does not show volumetric locking, and the convergence rate is quite higher than that of FEM. The application of EFMG is widespread nowadays, especially for simulating the fractures problems [10, 11].

In 2009, the article Element Free Galerkin Modelling of Composite Damage was published by I. Guiamatsia et al. The authors' objective was to analyze fracture damage (micro-cracks) of composite laminates with EFGM [12]. Authors considered intra-laminar, and interlaminar fracture of composites as present-day FEM cannot model an intralaminar matrix micro-crack (except the location of the damage is specified). This article represented simulation of delamination (interlaminar) and intralaminar fracture problem with EFGM first time. Different cases of double cantilever composite beam with micro-cracks have been solved with the EFG method. The virtual crack closure technique used for crack propagation in multitrack specimens, for linear analysis results, have good agreements with analytical and FEM for different cases of composite beams.

Ferreira [13] performed a meshfree analysis of a thick composite beam. The authors considered the four most common radial basis functions for review. RBF depends distance between the nodes or, more precisely, RBF is a function of distance to a center point and its Euclidian norm. The author considered first-order shear deformation theory (it is equivalent to the Timoshenko beam theory for isotropic beam) for cases of isotropic, orthotropic, and symmetric laminated composite beams under uniform load and perspective errors had calculated by the root mean square (RMS) method. The author has found that results obtained with a meshfree method with radial basis functions showing good agreement with the existing solutions. The method with radial basis function has massive potential for the solution of structural problems. Similar research work on the mechanics of composite material with a meshfree method can be found from the following literature [14–16].

In 1995, Liu et al. [17] suggested that RKPM is motivated by the theory of wavelets with all advantages as SMHM. One of the most popular meshfree methods MLPG was proposed by Atluri et al. [18]. It uses the local weak form for moving the least square-based method in which local weak forms are produced with overlapping sub-domains, and integration is performed over local sub-domain; hence, background cells are avoided [19, 20].

The main objectives of the authors are to study EFGM along with its mathematical implementation in structural problems. In Sects. 1 and 2, the Introduction and discrete Galerkin forms of EFGM have been discussed. In Sect. 3, a strong form of the mathematical models for 1-D and 2-D structural problems have been introduced. Subsequently, in Sect. 4, EFGM is applied to compute the deflection and stress, and computed deflection and stresses are validated against the exact solution of 1-D and a 2-D structural problem.

2 Mathematical Formulation

In EFGM the region of interest is first discretized with finite number of nodes, whether uniform or random distributed. Then the MLSM is used to estimate the field variable or displacement field u at a particular point x. Then Galerkin weak form are deducted from strong form of governing PDE. The Galerkin weak form [21] for solid mechanics' problems is given by

$$\int_{\Omega} \delta(L_d u)^T (cL_d u) d\Omega - \int_{\Omega} \delta u^T b d\Omega - \int_{\Gamma_t} \delta u t_\Gamma d\Gamma - \int_{\Gamma_u} \delta\lambda^T (u - u_\Gamma) d\Gamma - \int_{\Gamma_u} \delta u^T \lambda d\Gamma = 0 \quad (1)$$

u= Displacement components
b = Body forces
L_d = Matrix of linear operator.
c = Constitutive matrix or elasticity matrix
Γu = Integration over essential boundary
MLSM for trial and test function is given by,

$$u^h(x) = \sum_{I}^{n} \phi_I(x) u_I \quad (2)$$

Φ_I = Shape function at node I
Then, discrete Galerkin form of element free Galerkin equations from 1 is written as,

$$\begin{bmatrix} K & G \\ G^T & 0 \end{bmatrix} \begin{Bmatrix} U \\ \lambda \end{Bmatrix} = \begin{bmatrix} F \\ q \end{bmatrix} \quad (3)$$

K = Global stiffness matrix, λ = Lagrange multiplier (Unknown)
U= Vector of nodal parameters, F = Global force vector

$$k = \begin{bmatrix} k_{11} & k_{12} & k_{13} & k_{1nt} \\ k_{21} & k_{22} & k_{23} & k_{2nt} \\ k_{31} & k_{32} & k_{33} & k_{3nt} \\ k_{nt1} & k_{nt2} & k_{nt3} & k_{ntnt} \end{bmatrix}$$

$$k_{IJ} = \int_{\Omega} B_I c B_J d\Omega$$

$$B_I = L\Phi_I = \phi_{I,x}$$

$$F_I = \int_{\Omega} \Phi_I b_f d\Omega + \int_{\Gamma_T} \Phi_I \bar{t}_r d\Gamma$$

$$G_{IJ} = -\int_{\Gamma_u} N_I^T \Phi_J d\Gamma$$

where,

N_I = Lagrange interpolant
b_f = Body forces
t_r = Traction forces.

2.1 Moving Least Square Method

This method was proposed by Lancaster and Salkausdas [8] in 1981. This method used in pure mathematics for data and surface fittings. This method represents the function in a series manner. Nayroles first used this method for his novel approach titled "Diffused element method" for the construction of the shape functions. In 1994, Ted Belytschko used this method to develop his wonderful method, "EFGM." The fruitful advantage of MLSM is a convenience in order of consistency. In this section, the mathematical formulation of MLSM is given. The Series representation in moving least square method [8, 21] is given as,

$$u^h(x) = \sum_{j}^{m} p_i(x)\, a_j(x) = p^T(x) a(x) \tag{4}$$

a(x) is a vector of coefficient given by, The basis function in 1D, 2D and in 3D are given as,

$$a^T = \left\{ a_0(x)\ a_1(x)\ L\ a_m(x) \right\}$$

$$p^T(x) = \{p_0(x),\ p_1(x),\ \ldots,\ p_m(x)\} = \left\{1, x, x^2, \ldots, x^m\right\}\ (1D)$$

$$p^T(x) = p^T(x, y) = \left\{1, x, y, xy, x^2, y^2, \ldots, x^m, y^m\right\}\ (2D)$$

$$\mathbf{p}^T(\S) = p^T(x, y, z) = \left\{1, x, y, xy, yz, zx, x^2, y^2, z^2, \ldots, x^m, y^m, z^m\right\}\ (3D)$$

The choices of elements in the basis function can be chosen from the Pascal's triangle. Figure 2 presents schematic diagram for the same.

$$u^h(x,x_I) = p^T(x_I)\, a(x) \quad (I = 1,2,\ldots,n)$$

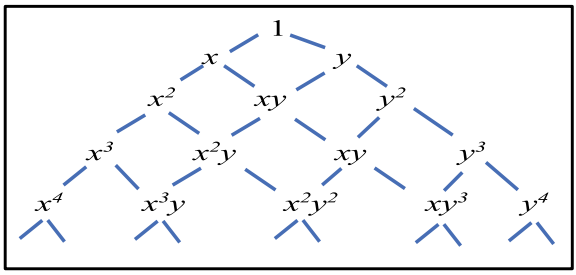


Fig. 2 Schematic diagram of Pascal's triangle

The weighted residual is constructed in ordinary manner,

$$J = \sum_{I}^{n} w(x - x_I)\left[u^h(x, x_I) - u(x_I)\right]^2$$
$$= \sum_{I}^{n} w\,(x - x_I)\left[p^T(x_I)\,a(x) - u_I\right]^2$$

The $w(x - x_I)$ represent weight function, there are clusters of weight function available in mathematics. Next section puts more emphasis on weight function and its physical significance. It is requiring to minimize the residual or weighted residual, hence finding derivative of above function and equate to zero.

$$\frac{\partial J}{\partial a} = 0$$
$$A(x)a(x) = B(x)ds$$
$$A(x) = \sum_{I}^{n} w_I(x)\,p^T(x_I)\,p(x_I)$$
$$B(x) = [B_I B_2 \ldots B_n]$$
$$B_I = w_I(x)\,p(x_I)$$
$$a(x) = A^{-1}(x)\,B(x)\,d_s$$

Hence, approximation function become

$$u^h(x) = \sum_{I}^{n}\sum_{j}^{m} p_j(x)\left(A^{-1}(x)B(x)\right)_{iI} u_I$$

Moving least square shape function at field node "I" can be written as,

$$f_I(x) = \sum_{j}^{m} p_j(x)\left(A^{-1}(x)\,B(x)\right)_{jI} = p^T A^{-1} B_I$$

2.2 Choice of Weight Function

The weight function is a function of the distance between two nodes. Weight function plays a very vital role in EFGM. Weight function (a part of weighted calculus) is a mathematical function used where there is a need to appoint weights, to perform sum, weighted average, and integration. The one-dimensional weight function is expressed in Fig. 3. The two-dimensional weight function is a modification of one-dimensional weight function in each direction (i.e., x and y). Figure 3 shows the behavior of cubic spline weight function (1D case) in the support domain of any field node x in natural co-ordinates (−1 to 1). Cubic Spline weight function [7, 22] can be written as,

$$w(x - x_i)^{\circ} w(\hat{r}) = \begin{cases} \frac{2}{3} - 4\hat{r}^2 + 4\hat{r}^3 & for\ \hat{r}\pounds\frac{1}{2} \\ \frac{4}{3} - 4\hat{r} + 4\hat{r}^2 - \frac{4}{3}\hat{r}^3 & for\ \frac{1}{2}\pounds\hat{r}\pounds 1 \\ 0\ for\ \hat{r}^3 1 \end{cases} \tag{5}$$

The weight function is a function of a distance between two nodes, $r = |x - x_i|$ and $\overline{r} = \frac{r}{r_w}$ r_w is a smoothing length, $r_w = c \times d_{max}$, 'c' is the distance between two nodes and d_{max} is a dimensionless size of the support domain. Generally, for solid mechanics problems $d_{max} = 3$. However, the accuracy of weight function mainly depending upon three parameters, such as choice of weight function, interpolation function, dimensionless size of support domain (Fig. 3).

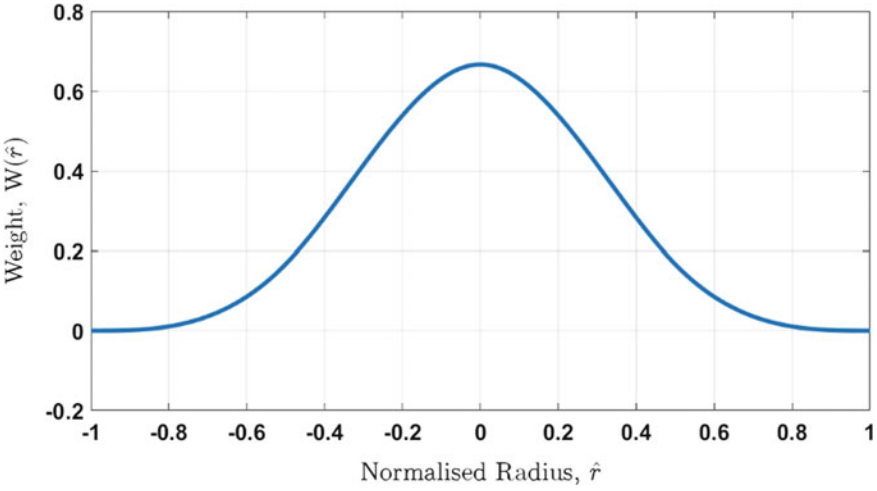


Fig. 3 Distribution of a cubic weight function

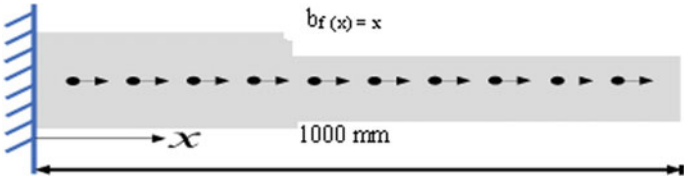


Fig. 4 Schematic diagram of a bar subjected to axial force

2.3 Imposition of Boundary Conditions

It is quite complex to impose boundary conditions in EFGM in comparison to FEM. The main reason to use a special tool such as the Lagrange multiplier method is that MLSM lacks Kronecker delta function property. Ted Belytschko et al. first applied LMM with EFGM to impose necessary boundary conditions. In Galerkin discrete equations, the Lagrange multiplier is treated as unknown functions of coordinates and considered as field variables [23].

$$\lambda^h = \begin{Bmatrix} \lambda_u \\ \lambda_v \end{Bmatrix} = \begin{bmatrix} N_I & 0 & \cdots & N_{n\lambda} & 0 \\ 0 & N_I & \cdots & 0 & N_{n\lambda} \end{bmatrix} = \begin{Bmatrix} \lambda_{u1} \\ \lambda_{v1} \\ \vdots \\ \lambda_{un\lambda} \\ \lambda_{vn\lambda} \end{Bmatrix} \tag{6}$$

Where, n_λ is nodes used for interpolation and N_I is shape function for Ith node at the boundary.

3 Numerical Problems

This section addresses two structural problems, first one is, bar under body forces and second one is, Timoshenko beam subjected to traction load at tip.

3.1 Circular Bar Subjected to Body Force

Considering the second order differential equation (Strong form) [24] for above problem,

$$(AE)\, u_{x,x} + b_f = 0 \qquad \Omega = (0, 1000) \tag{7}$$

Boundary Conditions are,

$$(i)\ u = \bar{u} \quad at\ \ x = \Gamma_u$$
$$(ii)\ E\,u_x = \bar{t}\ \ at\ \ x = \Gamma_t$$

A = Cross- section area,
E = Elastic modulus
b_f = Body Forces,
u = Field variable (Independent variable)
Γ_u = Essential boundary,
Γ_t = Traction boundary
The Eq. 8 presents the analytical solution of the above strong form.

$$u(x) = \frac{1}{AE}\left[\frac{L}{2}x - \frac{x^3}{6}\right] \tag{8}$$

x = Locations (Field points),
u = Displacements.

3.2 Two-Dimensional Timoshenko Beam Subjected to Traction at Tip

Assume a two-dimensional Cartesian coordinate system as shown in Fig. 5. $X = (x, y)$ become any point in the domain. Left hand side of cantilever beam is held fixed $x = 0$, hence right side is free end at $x = x_{max}$. L and W are the length and height of beam respectively, P stands for intensity of force at natural boundary (at free end)

The mathematical relation for parabolic distribution of force along height of the beam can be written as,

$$T_{ry} = \frac{-P}{2I}\left[\frac{W^2}{4} - y^2\right] \tag{9}$$

where,
T_{ry} = Traction force (N/mm^2)
P = Applied load (1500 N)
I = Moment of inertia (mm^4)

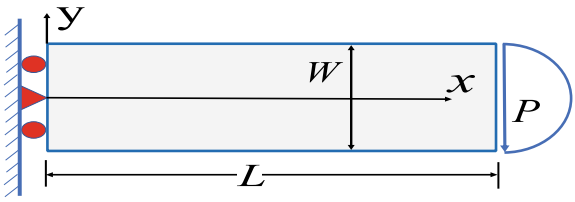


Fig. 5 Two–dimensional Timoshenko cantilever beam

y = Co-ordinates along height (mm)
W = Beam height (50 mm)
L = Beam length (100 mm)
Strong form for problem stated above is,

$$L_d^T \sigma + b_f = 0 \tag{10}$$

Boundary conditions considered for the above problem are,

$$(i)\ \ u = u_\Gamma\ ,\ \text{On essential Boundary}$$
$$(ii)\ \ L_n^T \sigma = t_\Gamma\ ,\ \text{On Traction Boundary}$$

where,
L_d = Matrix of differential operator
σ = Stress components
b_f = Body forces (if present)
t_Γ = Traction on natural boundary
u_Γ = Displacement on essential boundary
L_n = Unit normal on natural boundary in outward direction
Equation 11 shows an exact solution of the above-mentioned problem [25, 26].

$$u_x(x, y) = -\left[\frac{Py\,L}{EI} - \frac{Px^2 y}{2EI} + \frac{Py(2+v)}{6EI}\left(y^2 - \frac{w^2}{4}\right)\right] \tag{11}$$

$$u_y(x, y) = \left[\frac{Py^2 v(L-x)}{2EI} + \frac{PW^2 x}{24EI}(4+5v) + \frac{Px^2}{6EI}(3L-x)\right] \tag{12}$$

u_x and u_y are the displacement in x and y direction respectively at any field node (x, y).

4 Numerical Results and Discussions

The numerical results of 1-D and 2-D solid mechanics problems are presented in this section. The numerical solution for the displacement and normal stress were computed using a MATLAM Programme. The comparison of the numerical solution obtained using EFGM with the exact solution is presented in this section. The percentage error has been computed (Fig. 6).

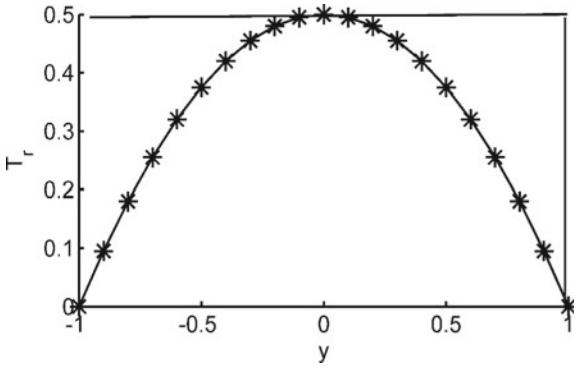


Fig. 6 Parabolic nature of traction along height in natural coordinates

4.1 Circular Bar Subjected to Body Forces

This subsection consist the results related to 1D problem of bar subjected to body forces as mentioned in previous section. As an elementary level, 6 field nodes have been taken under study and values of displacements at each field nodes have been evaluated. The comparison of EFGM and exact solutions is shown in Fig. 7.

One can note that EFGM solutions have good agreement with exact solutions, relative errors have been computed and expressed in Table 1.

4.1.1 Convergence Study

Convergence is a very common term in context to a numerical method. Actually, convergence is a combination of compatibility and smooth ness [27]. Compatibility means capacity or potency of basis function or approximation function to approximate given smooth function at arbitrary accuracy (Table 2).

As shown in Fig. 8, the convergence rate is computed with the help of an error norm relationship. The plot is on log-log paper. In Fig. 8, "h" represents the distance

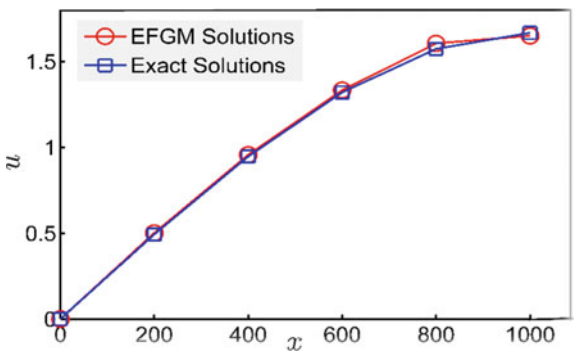


Fig. 7 Comparison of EFGM and exact solution

Table 1 Comparison between EFMGI and exact solution

Sr. no	Co-ordinates (x)	EFGM solution (in mm)	Exact solution	% Relative error
1	0	7.26×10^{-19}	0	–
2	200	0.5011	0.4933	1.58
3	400	0.9564	0.9466	1.03
4	600	1.3325	1.3200	0.94
5	800	1.6054	1.5733	2.03
6	1000	1.6495	1.6667	1.03

Table 2 Number of nodes and error in energy norm

Sr. no	No. of field nodes	Log_{10} (Error in energy norm)
1	6	−6.0439
2	8	−6.3563
3	11	−6.6837
4	15	−6.9885
5	21	−7.3074
6	29	−7.6046
7	41	−7.9160
8	58	−7.9160
9	82	−8.5207
10	116	−8.8149
11	164	−9.1026

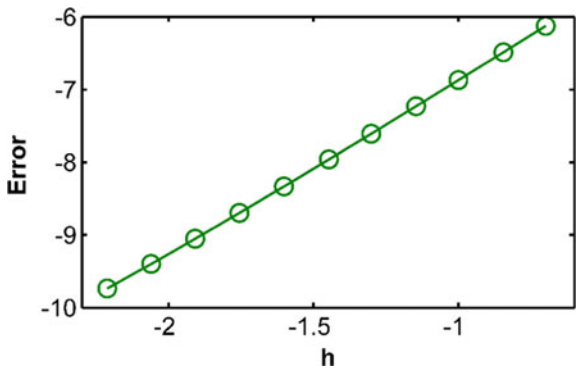


Fig. 8 Convergence rate of the numerical solution

between two nodes. Here, field nodes are increased in geometrical sequence manner with a common ratio that is the square root of two ($\sqrt{2}$). The distance between two nodes are continuously reduced to note down the errors in energy norm, in other words, number of nodes are increased. The mathematical relations for energy norm

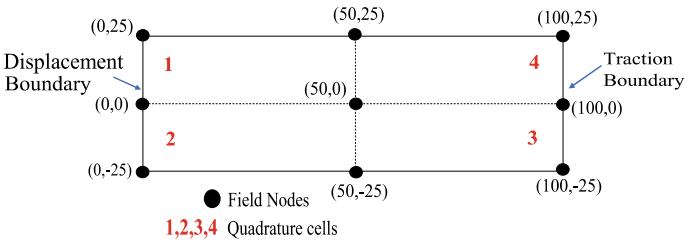


Fig. 9 Field nodes and background quadrature cells for numerical integration

[28] (integration is in natural co-ordinates) can be written as,

$$E_{norn} = \left[\frac{1}{2} \int_0^1 \left(u^{EFGM} - u^{Exact} \right)^T \rfloor \left(u^{EFGM} - u^{Exact} \right) dx \right]^{\frac{1}{2}} \tag{13}$$

4.2 *Two-Dimensional Timoshenko Beam Subjected to Traction at Tip*

This subsection consist the results related to 2D problem of Timoshenko beam as mentioned in previous section. As an elementary level, 9 field nodes have been taken under study and values of displacements (in x and y) and normal stresses at each field nodes have been evaluated. Figure 9 shows the distributions of the field nodes.

Well known numerical integration scheme that "Gauss Quadrature Numerical Integration" has been used here for evaluating integrands in discrete Galerkin weak form. EFGM uses the background integration cells for performing numerical integration. Numerical solutions for displacements have been shown in Table 3 with comparison with exact results.

In bending of beams, maximum tensile and compressive stress at any cross section present at location farthest from neutral axis. In this problem stresses are computed at each Gauss points. At fixed end, the stresses are maximum as compared to any other sections. The distributions of normal stresses at fixed end of the beam are shown in Fig. 10, one can note that upper half portion of beam (0 to 25) endures tension (hence, positive sign) and below half portion (0 to −25) endures compression (hence, negative sign). The distributions of normal stresses in the beam can be seen in Fig. 10 and perspective contour plot of stress distribution in Fig. 11.

Table 3 Displacements in x and y and % relative errors

Sr. no	Field node (x, y)	Displacement in "x"		% Error	Displacement in "y"		% Error
		EFGM	Exact		EFGM	Exact	
1	(0,25)	0.0024	0	-NA-	−0.0187	−0.0185	1.081
2	(0,0)	0.0052	0	-NA-	0	0	0
3	(0,−25)	−0.0024	0	-NA-	−0.0186	−0.0185	0.540
4	(50,25)	0.1408	0.1588	6.863	−0.2155	−0.2365	8.583
5	(50,0)	0.2050	0	-NA-	−0.2066	−0.2272	9.066
6	(50,−25)	0.0035	−0.1588	6.863	−0.2155	−0.2365	8.879
7	(100,25)	0.2050	0.2118	3.220	−0.6267	−0.6662	5.929
8	(100,0)	0.0035	0	-NA-	−0.6231	−0.6662	6.469
9	(100,−25)	−0.2105	−0.2118	3.210	−0.6245	−0.6662	6.259

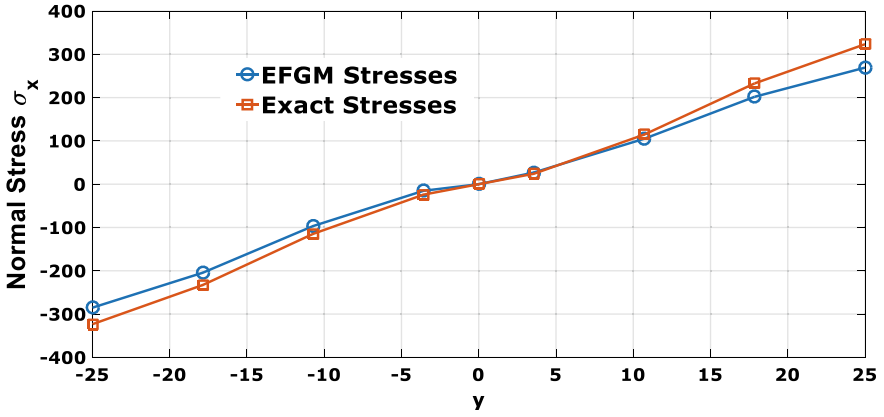


Fig. 10 Normal stress distribution at the fixed end σ_x at x = 0

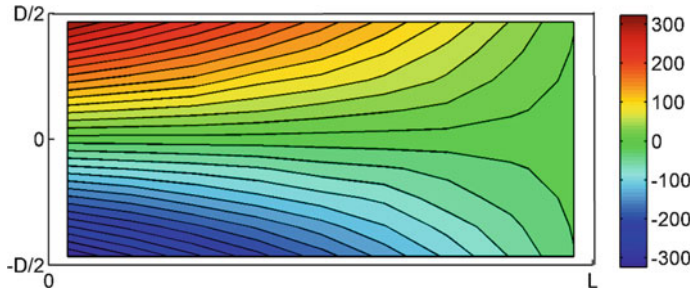


Fig. 11 Contour plot of normal stress distribution in the beam

4.2.1 Convergence Study

This section reveals the method of convergence for two-dimensional beam problem, for convergence, here errors in energy norm is considered for stresses values. Summing up all local energy at nodal level leads to global error in energy. This global energy is normalized as $(u + e)$, here "u" is strain energy, expressed as error in energy normalization. In mathematical operations and treatments of numerical methods like FEM, energy norm is measure of convergence and function of element/mesh sizes. It is well defined by researchers that accuracy in stresses is more vulnerable as compare to displacements [29].

Mathematically,

$$E_{norn} = \left[\frac{1}{2} \int_{\Omega} \left(\varepsilon^{EFGM} - \varepsilon^{Exact} \right)^T E \left(\varepsilon^{EFGM} - \varepsilon^{Exact} \right) d\Omega \right]^{\frac{1}{2}} \tag{14}$$

The Integration is performed over entire domain (Ω)), Where, E = Elasticity matrix, and ε = Computed Strain vectors.

The plot of L_2 error norm and distance between nodes (linear distance in "x") are showed in Fig. 12. The number of nodes selected for investigation are 9, 16, 25, 36, 49, 64, 81, 100, and 121. One can note that convergence is monotonic, which means it is not fluctuating and does not change the sign in-between. The convergence rate is found to be 2.321 using cubic spline weigh function and dimensionless size of support domain was 3, it is better to mentioned again here, that accuracy/convergence

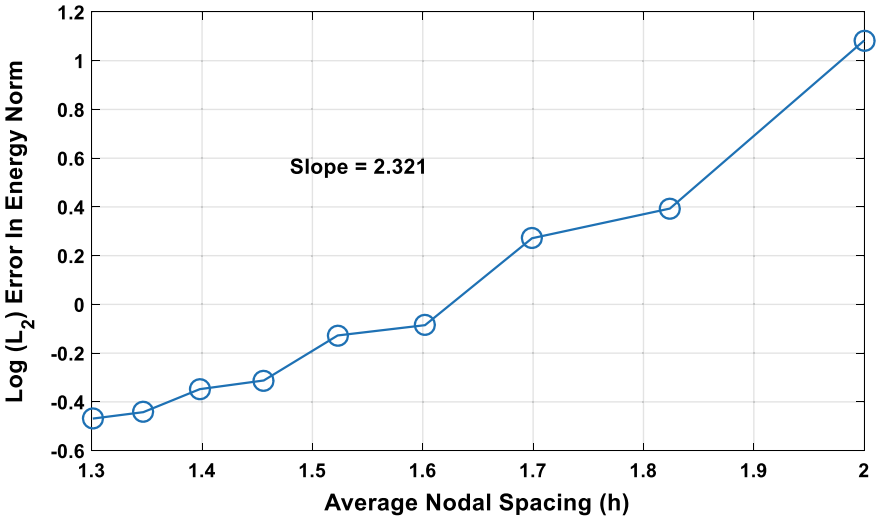


Fig. 12 Convergence rate of the two-dimensional problem

of EFGM is depends on many factors and two key factors are weight function and dimensionless size of support domain.

5 Conclusion

The present study employed EFGM to study two structural problems (1D and 2D), and numerical solutions of both problems have been compared with the exact solution. Numerical results demonstrate that the EFGM results have good agreements with the analytical results in terms of accuracy. Percentage relative errors have been computed, and the convergence study has been performed for both the problem. In the case of the one-dimensional problem, the convergence rate is found to be 2.02, and for a two-dimensional problem, it is 2.31. Thus the convergence rate is found quite near for both the problem. In the EFGM, numerical integration plays a vital role in terms of convenience and accuracy. In both the present cases, the Gauss quadrature method was employed. However, for the one-dimensional problem, the Gauss quadrature method, Trapezoidal rule, and Simpson's 1/3 rule give identical results. However, for the two-dimensional problem domain, a robust method such as Gauss quadrature should be employed. As mentioned, accuracy and convergence of EFGM mainly depend on the choice of weight function, shape function, dimensionless size of support domain, type of node distributions in the domain (uniform or random), and shape of support domain. Change in any of the parameters renders the change in present results. There are no guidelines available for the distribution nodes in the problem domain. Several researchers created some techniques from FEM, such as triangulations and background finite element mesh, for appropriate distributions of nodes in the problem domain. However, the principal difference between FEM and the meshfree approach will be the same.

References

1. Li H, Mulay S (2017) Meshless methods and their numerical properties. CRC Press
2. Nagevadiya B et al (2017) Galerkin meshfree methods: a review and mathematical implementation aspects. Int J Appl Comput Math 5(105). https://doi.org/10.1007/s40819-019-0665-4
3. Gingold R, Monaghan J (1977) Smoothed particle hydrodynamics: theory and application to non-spherical stars. Mon Not R Astron Soc 181:375–389. https://doi.org/10.1093/mnras/181.3.375
4. Shadloo M, Oger G, Touze D (2016) Smoothed particle hydrodynamics method for fluid flows, towards industrial applications: motivations, current state, and challenges. Comput Fluids 136:11–34. https://doi.org/10.1016/j.compfluid.2016.05.029
5. Zhang C, Xiang G et al (2019) a weakly compressible sph method with weno reconstruction. J. Comput Phys 392:1–18. https://doi.org/10.1016/j.jcp.2019.04.03

6. Nayroles B, Touzot G, Villon P (1992) Generalizing the finite element method: diffuse approximation and diffuse elements. Comput Mech 10:307–318. https://doi.org/10.1007/BF00364252
7. Belytschko T, Lu Y, Gu L (1994) Element-free Galerkin methods. Int J Numer Method Eng 37:229–256. https://doi.org/10.1002/nme.1620370205
8. Lancaster P, Salkauskas K (1981) Surfaces generated by moving least squares methods. Math Comput. 37:141–158. https://doi.org/10.1090/S0025-5718-1981-0616367-1
9. Mayers S (2003) An introduction to numerical analysis. Cambridge University Press, Cambridge
10. Belytschko T, Black T (1999) Elastic crack growth in finite elements with minimal remeshing. Int J Numer Method Eng 45:601–620
11. Lu Y, Belytschko T, Gu L, A new implementation of the element free Galerkin method. Comput Method Appl Mech Eng 113(3):397–414. https://doi.org/10.1016/0045-7825(94)90056-6
12. Guiamatsia I, Falzon B, Davies G, Iannucci L (2009) Element-free Galerkin modelling of composite damage. Compos Sci Technol 69:2640–2648. https://doi.org/10.1016/j.compscitech.2009.08.005
13. Ferreira AJM (2010) Thick composite beam analysis using a global Meshless approximation based on radial basis functions. Mech Adv Mater Struct 10:271–284. https://doi.org/10.1080/15376490306743
14. Kahwash F, Shyha I, Maheri A (2018) Dynamic simulation of machining composites using the explicit element-free Galerkin method. Compos Struct 198:156–173. https://doi.org/10.1016/j.compstruct.2018.05.034
15. Watts G, Pradyumna S, Singha M (2017) Nonlinear analysis of quadrilateral composite plates using moving kriging-based element free Galerkin method. Compos Struct 159:719–727. https://doi.org/10.1016/j.compstruct.2016.09.100
16. Watts G, Pradyumna S, Singha M (2018) Nonlinear bending and snap-through instability analyses of conical shell panels using element free Galerkin method. Thin-walled structures 122:452–462. https://doi.org/10.1016/j.tws.2017.10.027
17. Liu W, Jun S, Zhang Y (1995) Reproducing kernel particle methods. Int J Numer Method Fluid 20:1081–1106. https://doi.org/10.1002/fld.1650200824
18. Atluri J, Zhu T (1998) A new Meshless local Petrov-Galerkin (MLPG) approach in computational mechanics. Comput Mech 22:117–127. https://doi.org/10.1007/s004660050346
19. Atluri SN, Zhu TL (2000) The Meshless local Petrov-Galerkin (MLPG) approach for solving problems in elastostatic. Comput Mech 25(2000):169–179. https://doi.org/10.1007/s004660050467
20. Atluri SN (2004) The Meshless method (MLPG) for domain and BIE discretization's. Tech Science Press
21. Liu GR (2009) Meshfree methods: moving beyond the finite element method. CRC Press
22. Knott GD (2002) Interpolating cubic splines. Springer
23. Hoffman JD, Frankel S (2001) Numerical methods for engineers and scientists. CRC Press
24. Reddy JN (1984) Energy and variational methods in applied mechanics. Wiley
25. Timoshenko SP, Goodier JN (2010) Two-dimensional problems in rectangular co-ordinates. Theory of elasticity 3rd edn, Ch 3, pp 41–45, Tata McGraw-Hill Education Private Limited
26. Timoshenko S, Gere J (2009) Beam—column. Theory of elastic stability, 2nd edn, Ch 1, pp 1–12, Dover Publication Inc., New York
27. Zienkiewicz OC, Taylor RL (2013) The finite element method: its basis and fundamentals, Elsevier. https://doi.org/10.1016/C2009-0-24909-9
28. Strang G, Fix G (1973) Approximation, in: an analysis of the finite element method, Prentice-Hall
29. Fish J, Belytschko T (2007) A first course in finite elements, Wiley

Theoretical Approach to Chebyshev Spectral Collocation Method and Its Mathematical Implementation

Rameshkumar Bhoraniya, Pinank Patel, Ramdevsinh Jhala, and Rajendrasinh Jadeja

Abstract This chapter describes the linear stability investigation of the incompressible viscid flow between the two concentric counter-rotating vertical cylinders. The parallel flow assumption was considered for the base flow, and hence it is varying in the radial direction only. The flow is a shear driven, and hence the pressure gradient is zero in the stream wise direction. The Governing stability equations for the disturbance flow quantities are derived in cylindrical polar coordinates by coupling the energy equation with the Navier-stokes formulas. The stability formulas are discretized using CSC method. The discretized stability formulas, combined with appropriate boundary conditions, form a general Eigenvalues problem (EVP). The full spectrum of the eigenvalue problem is computed for the different Reynolds numbers under the effect of viscous heating, different radius ratio, and buoyancy. The axial and radial wave numbers, β and α are taken as $\pi/2$ and zero, respectively. The effect of viscosity variation due to temperature is introduced by Nahme number (Na) and Brinkman number (Br) and effect of buoyancy by Grashof number (Gr). The acute value of Re of the flow is computed for isothermal and non-isothermal Ta–Co flow including the effect of viscid heating and buoyancy at different radius ratio (η). The viscous heating and buoyancy effect destabilize the flow.

Keywords Stability analysis · Chebyshev · Spectral collocation method · Eigenvalue problem

R. Bhoraniya · P. Patel · R. Jhala · R. Jadeja (✉)
Marwadi University, Rajkot, India
e-mail: rajendrasinh.jadeja@marwadieducation.edu.in

R. Bhoraniya
e-mail: rameshkumar.bhoraniya@marwadieduction.edu.in

P. Patel
e-mail: pinank.patel@marwadieduction.edu.in

R. Jhala
e-mail: ramdevsinh.jhala@marwadieducation.edu.in

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_7

Nomenclature

T	Temperature
μ	Viscosity
ρ	Density
C_p	Specific Heat
d	Diameter of Cylinder
g	Gravitational Acceleration
η	Radius Ratio
Pr	Prandtl number
Gr	Grashof number
Re	Reynold's Number
R_{cr}	Critical Reynolds number
Na	Nahme Number
Br	Brinkmen Number
α	Azimuthal wave number
β	Axial wave number

Abbreviations

ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
EVP	Eigenvalue Problem
FDM	Finite Difference MethodFinite Difference Method (FDM)
FEM	Finite element MethodsFinite Element Method (FEM)
CSC	Chebyshev Spectral Collocation
Ta–Co	Taylor–Couette flow

1 Introduction

The (PDEs) and its applications are important in the field of applied mathematics. These are a basic form of equations in the number of applications of physics, natural science and finance. They are used to describe the local properties of the function in the three-dimensional fluid flow problems.

The concept of discretization is the easiest set of rules to approximate the solution of PDEs. In this process, the PDEs are represented as the determinate dimensional problem. At the same time, replacing the Partial differential equation by a distinct model is not insignificant at all, and more often, the choice of the determinate dimensional model to be used depends on the properties behind the mathematical model

itself. The Recent advances in computer technologies have made it easier to determine accurate solutions of the PDEs efficiently, even in the most critical cases of very large systems of PDEs.

The FDM and FEM methods are often used for the mathematical solution of PDEs. However, in the computations of the spatial derivatives, these methods essentially require a massive number of nodal points to provide an accurate numerical solution. The Spectral and Pseudo-spectral methods have been developed as an alternative solution to it in recent years. The spectral methods are different from the FDM and FEM methods, in spectral methods global information is integrated in the calculation of a spatial derivative. The spectral methods can provide greater accuracy for a smooth solution with the use of a very less number of nodes and, therefore, less calculation time as compared to FDM and FEM.

The spectral procedures are widely used for the flow simulations due to higher accuracy. However, it is very difficult to apply it to complex geometries, and generally, it is used for simple geometries. The method of collocation is a numerical solution method for the ODE, PDE and integral equations. In the collocation method, a finite-dimensional space of solution is chosen (most often polynomials), and a number of collocation points are also chosen. Then a solution is selected such that it fulfills the condition of a given equations at the association positions. The orthogonal collocation on finite elements is also used to solve a PDE from Fluid Dynamics. Association locations are selected as the roots of orthogonal polynomials gave better results because of a few striking characteristics of these polynomials.

The main objective of the authors is to study the Spectral collocation method using Chebyshev polynomial and to demonstrate its application for the numerical solution of fluid flow problem. The stability examination of the incompressible flow passing between the two rotating cylinders having same axis of rotation (Ta–Co flow) has been carried out to demonstrate the application of Spectral collocation method. Sections 1 and 2 presents basic introduction, mathematical background and relevant literature review. The governing stability equations, boundary conditions and numerical solution of then eigenvalue problem is discussed in Sect. 3. Section 4 presents solution of base flow temperature profile under the effect of variable viscosity. The validation of the computed results and effect of radius ratio, viscous heating and bouncy on the stability of Ta–Co flow have been discussed in the Sects. 5 and 6.

2 Chebyshev Method

The CSC method is used in discretizing the governing equations and group more grid point at the boundary of a domain. The CSC Method is Global in Nature. In this method Computation at any point depends on information from whole domain. The Chebyshev Spectral Collocation Method provides exponential Convergence rate. This method provides precise results with moderate number of grid points.

Let us consider one ODE, $y'(t) = f(t, (y(t)), y(t_0) = y_0$ The equation is required to be resolved in the interim $[t_0, t_0 + C_k h]$. Choose C_k from $0 \leq c_1 < c_2 < \cdots < c_n \leq 1$. The compliant polynomial association method come close to the result y by the polynomial p of degree n this solution contents the primary condition $qt(0) = y_0$, it also satisfies the variance equation

$$q'(t_0 + h) = f(t_0 + h, q(t_0 + h))$$
$$y_1 = q(t_0 + h)$$

$q'(t_m) = f(t_m, q(t_m))$ at all association points $t_m = (t_0 + C_m h)$ for $m = 1, 2, \ldots$ n. This results in $n + 1$ conditions, which equals the $n + 1$ constraints needed to identify a polynomial of degree n. The association methods used here are implied "Runge Kutta methods". The constants C_m in the "Butcher stand of a Runge Kutta method" are the association points. It may be important to understand that not all implied Runge–Kutta methods are association methods. The association method can be explained with following case. Let us consider two association points $C_1 = 0$ and $C_2 = 1$.

The association conditions are

$$qt(0) = y_0 \tag{1}$$

$$q' = f(t_0, q(t_0)) \tag{2}$$

$$q'(t_0 + h) = f(t_0 + h, q(t_0 + h)). \tag{3}$$

Above mentioned three conditions are used in collocation method, this indicates that p has to be a polynomial of second degree. We can also write the function q as: $q(t) = \alpha(t - t_0)^2 + \beta(t - t_0) + \gamma$ this will help us to reduce the calculations. The coefficients are evaluated by using collocation conditions.

$$\alpha = \frac{1}{2h} = f((t_0 + h, q(t_0 + h) - f(t_0, q_0)), \beta = f(t_0, p_0), \gamma = y_0 \tag{4}$$

The collocation method is now given by

$$y_1 = q(t_0 + h) = y_0 + \frac{1}{2}h(f(t_0 + h), y_1) + f(t_0, y_0)) \tag{5}$$

where, $y_1 = q(t_0 + h)$ is approximate solution at $t = (t_0 + h)$.

The Chebyshev method is used to calculate number of collocation points in a domain. Following figure shows the distribution of collocation points in the given domain.

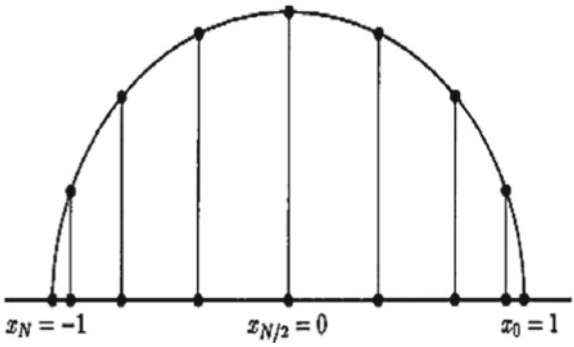


Fig. 1 Chebyshev point distribution

It is evident from the Fig. 1 that Chebyshev locations are placed at equal distance on to the upper half of the unit circle but its projection on to the x-axis are not equally spaced. There are more grid locations which are present at the extreme points in comparison with the center or we can say that there is clustering of grid points at the extreme points in comparison with to the center, it is also observed that the space between grid points at ends is less in comparison to space between grid points at the center so there is finer mesh at the boundary, which will help to have better and accurate results. CSC methods support to characterize a function in best possible way with the help of few representative points.

Periodic Function: The sample points which are evenly spaced throughout the interval are selected to describe a periodic function over an interval. These sample points are selected using N function samples, and they form a trigonometric interpellant comprising of a sum of N sinusoids. This methodology produces merger in N for integration, differentiation, and interpolation.

Non-periodic functions: A Non-periodic function over an interval is characterized with the help of N function examples, the interval is mapped into [−1, 1]. The sample points are chosen based on Chebyshev criteria and a polynomial interplant comprising of a sum of N Chebyshev polynomials is created by the selected points. This method offers convergence exponential in N for integration, differentiation, and interpolation.

Comparison of the Chebyshev Spectral Collocation Method with Analytical Method:

The comparison of CSC Methods is done by using derivative of the Sin (X) with the help of analytical method as well as by CSC method.

The Square in the Fig. 2 represents the results using Analytical Method and star represents the results using spectral method. The results depict that the CSC method provides closely accurate to analytical solution points.

Chebyshev differentiation:

If a vector f_{even} is trajectory of function models considered at equally spaced points in an interval [a, b] i.e. if

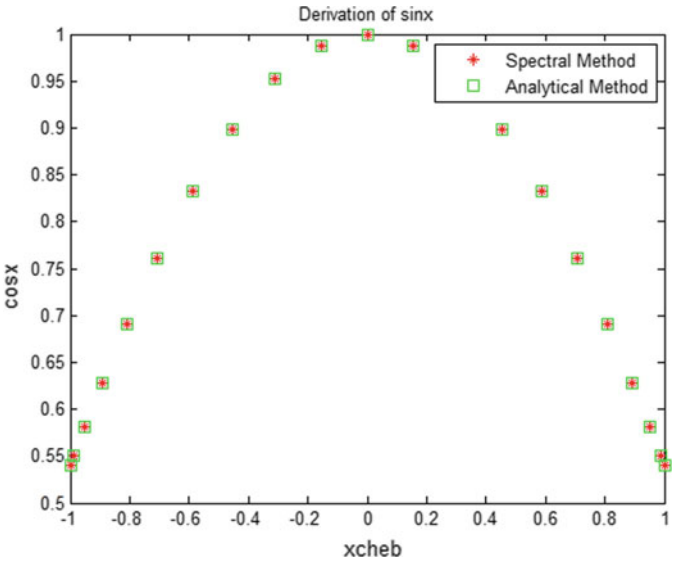


Fig. 2 CSC and analytical method results

$$f_{even} = \begin{pmatrix} f(a) \\ f(a+\Delta) \\ f(a+2\Delta \\ . \\ . \\ f(b) \end{pmatrix} \tag{6}$$

Then, the vector of derived values at choose points can be presented in the form of a centered FDM calculation it will be in the form of a matrix vector product of $f'_{even} = D^{CFD}$

Where,

$$D^{CFD} = \frac{1}{2\Delta} \begin{pmatrix} 0 & 1 & 0 & 0 \cdots 0 & 0 \\ -1 & 0 & 1 & 0 \cdots 0 & 0 \\ 0 & -1 & 0 & 1 \cdots 0 & 0 \\ 0 & 0 & -1 & 0 \cdots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0\,0 & 0 & 1 \\ 0 & 0 & 0 & 0\,0 & -1 & 0 \end{pmatrix} \tag{7}$$

This approximation will converge like $1/N^2$. In Chebyshev spectral method we just need to construct the Nth Chebyshev approximant $\int_{approx}(x))$ to f(x) and differentiate the variable of approximation and considering this as an approximation to the

derivative. The *N*th Chebyshev approximation to f(x) is

$$\int_{approx}(x) = \sum_{m=0}^{N} C_m T_m(x) \tag{8}$$

Differentiating we get

$$f\prime_{approx}(x) = \sum_{m=0}^{N} C_m T\prime_m(x) \tag{9}$$

Now when we calculate the results of this equation at the (N + 1) Chebyshev points: $X_n = \cos \frac{n\pi}{X}$ Where n = 1, 2, … N, we achieve vector $f\prime_{cheb}$. The entries of this vector are estimated values of the derivative of f_{cheb} at the Chebyshev points, and it is related to the vector C of Chebyshev coefficients via a matrix-vector product relationship:

$$N_{Gn} = \frac{\mu_0 U_\infty^2}{k\Delta T_0};\ \mathrm{R}_e = \frac{\rho U_\infty d}{\mu};\ P_r = \frac{\mu_0 C_p}{k};\ G_r = \frac{g\rho^2\beta\Delta T_0 d^3}{\mu_0^2} \tag{10}$$

$f'_{cheb} = T'C$ where T' is (N + 1) x(N + 1) dimensional Matrix.

If we consider the relation C = Λf_{cheb} We receive $f\prime_{cheb} = \mathrm{T}'\Lambda f_{cheb}$ where $Dcheb = T'\Lambda$ matrix that functions on a vector of models at Chebyshev points to produce a vector of f_0 samples at Chebyshev points.

2.1 Application of Spectral Collocation Method in Fluid Dynamic Problems

The bounded flows through channels and other simple geometries are simple in configurations. However, they are too important from the scientific and technological aspects. The vortices generated at corners and in the flow direction, flow transition, and turbulence can be analyzed in the same bounded configurations. To study two-dimensional and three-dimensional flow problems, one can use an arithmetical result of the Na-S equations. The primitive variable and vortices-Stream function approached are used to study viscous fluid flow problems. In the first variable approach, a coupling of pressure and velocity is challenging to satisfy the incompressibility condition.

In the case of a vortices-stream function approach, such a problem, the pressure term is eliminated from the equations. However, the order of the derivative increases in this formulation. In a two-dimensional fluid flow problem, this approach is widely used. However, a straight forward extension of it for three-dimensional flow is not that

easy due to the increased order of derivatives. Thus, the primitive variable approach is more suitable for 3D problems [1].

The non-iterative methods like estimate method and small step method have been developed to prevent the coupling of pressure-velocity difficulty of the primitive variable approach. In the methods mentioned above, no special memory storage is required, and they are suitable for the solution of unsteady flow problems. These methods use the prediction-correction method in which pressure is predicted from the projected velocity in the divergence-free velocity field.

The higher-order temporal scheme with the spectral method is used to improve accuracy by incorporating a variation of the projection method with the pseudo-spectral method. A semi-implicit projection system with second-order accuracy in temporal discretization gives good numerical stability in spectral collocation form. The diagonalization is performed, which is an effective and efficient direct method of solution. This combination of a temporal scheme with spectral-spatial discretization is giving a very fast solution in comparison to the iterative solution procedure. The combination of the two higher order of accurate methods has been validated using square opening flow at Re = 10,000 and backward-facing step in channel flow at Re = 875 [1].

Taylor–Couette flow is a drag driven flow in which flow is produced by the relative circular motion of the cylinders. Therefore, the outside compelling by a pressure gradient is not present [2]. It is an ideal case for studying the flow instability of Newtonian and Viscoelastic fluids. Couette and Mallock are the first to investigate the stability analysis of the viscous incompressible flow [3]. The first mathematical representation of the flow between two concentric, infinite long cylinders having rotary motion in three dimensions was successfully developed by G. I. Taylor and obtained a result that was similar to the experimental results [4]. The work started by Taylor continues the study of two-dimensional flow between the cylinders, which is Taylor's classical problem [5]. There are many theories available to analyze the stability of the flow i.e. energy gradient theory, classical linear theory, Non-linear theory, etc. The majority of the study has been performed on the steadiness of Ta- Co flow without viscid heating effect. Ali and Weidman (1990) theoretically investigated the effect of the temperature gradient in the radial direction on the volatility of the Ta–Co flow [6].

They show that for a given Prandtl number (Pr), the subordinate flow is symmetrical about the axis, and the perilous value of Re surges with the increase in Gr. The resilience force has a steadying result except for large value Pr number fluids. However, the subordinate flow becomes asymmetric about the axis for larger Gr, and Rcr number decreases with increasing Gr [7]. Kolyshkin and Vaillancourt (1993) studied effect of energetics on isothermal Ta–Co flow for a radius ratio from 0.4 to 0.95 and Pr from 1 to 100. They found that the disturbances which are symmetric about an axis are most hazardous turbulences, and flow is disrupted as Pr and Gr increases [8]. The thermal buoyancy effect on the instability was examined by Dah-Chyi Kuo and K. S. Ball (1997). The critical value of Re for the finite annulus is almost equal to the forecast based on linear stability results for continuous length for the isothermal case. The Rcr number for non-isothermal flow is slightly less as compared

to that of isothermal Ta–Co flow [9]. The majority of the research is performed on the steadiness of the Couette flow with the rotary motion of either inside or outside cylinder or with motion about the axis of the inner cylinder depending upon the application [10]. The stability of the channel flow, pipe flow, Couette, and boundary layer flow with viscous heating has been examined by several authors [11, 12]. The experiment of the stability isothermal Ta–Co flow with viscid heating of Newtonian fluid has been studied by White and Mullar [13]. They admitted that the uncertainty of flow is resulted due to a connection between viscid indulgence encouraged temperature stratification and inertial forces. In the case of viscous heating of the fluid, the property viscosity is used the temperature-dependent, and Nahme type viscosity-temperature rise is used to analyze the effect of temperature on the viscosity. Al-Mubaiyedh, Sureshkumar & Khomami (1999) per-formed instability analysis to analyze the effect of viscid indulgence on the Ta–Co flow instability. They considered liquids of high viscidness and large initiation energy with temperature-dependent viscosity [14]. Na provides the degree of thermal compassion of the fluid. Al-Mubaiyedh et al. (2002) studied in detail the thermal influence on the distribution of pressure and kinematics also uneven disturbances to know the flow disruption mechanism due to viscid heating [15]. They predicted a new mode of instability due to viscous heating effect in the absence of buoyancy effect with higher Pr (11000). At high Pr as Na increases, the coupling between radial disturbances and temperature gradient enhances centrifugal instability [16]. In the case of the instability of the Ta–Co flow, the effects of non-isothermal are the result of viscid indulgence and not because of a forced temperature gradient, which has not been studied significantly [17, 18]. In this chapter, we have presented the study of the effect of viscid heating and buoyancy on the stability characteristics of the Ta–Co flow for different radius ratios of the cylinders with uniform i.e., isothermal and non-uniform wall heating i.e., Non-isothermal.

3 Problem Formulation

The incompressible and Newtonian fluid with viscous heating is considered among two concentric rotating cylinders. The walls of the cylinder are kept at the same temperature in the matter of isothermal Ta–Co flow and different invariable temperatures for non-isothermal Taylor–Couette flow. The governing stability equations (Linearized Navier-Stoke equations) are derived using the standard procedure. The space between the two rotating cylinders is very small in comparison to the radii of the inside cylinder. The Reynolds number based on the spaced $= (R_2 - R_1)/2)$ between inside and outside cylinder is considered, where R_1 and R_2 are the inside and outside radii of the cylinder, respectively. The radius of the cylinders is normalized by the gap (d). The base flow is changing in the radial direction only. However, the disturbances are three dimensional. Both the cylinders are rotating in opposite directions with the same angular velocity. Viscid indulgence is encompassed in the energy equation through generation number (NGn), which is described as the maximum change of temperature due to viscid heating regulated by the controlled wall temperature,

which is known as Nahme number (Na) and Brinkman number (Br) for isothermal and non-isothermal Taylor–Couette flow respectively [6]. The following relation is used to determine the variation of viscosity with the change in temperature [19].

$$\mu = \mu_0 \exp\left[\beta \frac{T - T_0}{T_0}\right] \tag{11}$$

μ_0 is the reference viscidness at temperature T_0 and it is a non-dimensional initiation energy parameter that characterize the change in viscosity with reference to temperature variation; where β is a positive number for the liquids. The Generation number (N_{Gn}), Reynolds Number (Re), Prandtl number (Pr) and Grashof number (Gr) are defined as,

$$N_{Gn} = \frac{\mu_0 U_\infty^2}{k\Delta T_0}; \mathrm{R}_e = \frac{\rho U_\infty d}{\mu}; P_r = \frac{\mu_0 C_p}{k}; G_r = \frac{g\rho^2\beta\Delta T_0 d^3}{\mu_0^2} \tag{12}$$

where, ΔT_0 is the reference temperature difference, β is thermal expansion co-effcient at reference temperature T_0. The disturbances are assumed to be in normal mode form with the amplitudes are the functions of a radial coordinate only

$$\overline{U_r} = 0 + u_r; \overline{U}_\theta = U_\theta + \mu_\theta; \overline{U}_z = 0 + u_Z; \overline{P} = P + p; \overline{\mu} = \mu + \mu' \tag{13}$$

$$q(r,t) = \widehat{q}(r) e^{[i(\alpha\theta + \beta Z - \omega t]} \tag{14}$$

where $q = [u_r, \mu_\theta, \mu_\phi, P]$, $\overline{Q} = [\overline{Ur, U\theta, U\phi, P}]$, r is radial coordinate, z is axial coordinate, θ is azimuth coordinate, ω is circular frequency, α is azimuth wave number, and β is axial wave number.

$$\frac{\partial u}{\partial r} + \frac{U_\theta}{r}\frac{\partial u_r}{\partial_\theta} - \frac{2U_\theta}{r}u_\theta = -\frac{\partial P}{\partial r} + \frac{1}{\mathrm{R}_e}\left[\begin{array}{c} 2\mu\dfrac{\partial^2 u_r}{\partial^2 r} + 2\dfrac{\partial u}{\partial r}\dfrac{\partial u_r}{\partial r} + 2\dfrac{\mu}{r}\dfrac{\partial^2 u_r}{\partial^2 r} + \\ \dfrac{\mu}{r}\left(\dfrac{\partial u_\theta}{\partial_r \partial_\theta} - \dfrac{1}{r}\dfrac{\partial u_\theta}{\partial_\theta} + \dfrac{1}{r2}\dfrac{\partial^2 u_r}{\partial^2 \theta}\right) + \\ \dfrac{1}{r}\dfrac{\mu}{\theta}\left(\dfrac{\partial u_\theta}{\partial r}\dfrac{u_\theta}{r}\right) - 2\dfrac{\mu}{r^2}\left(\dfrac{\partial u_\theta}{\partial \theta} + u_r\right) + \\ \left(\mu\left(\dfrac{\partial^2 u_z}{\partial r \partial z} + \dfrac{\partial^2 u_r}{\partial z^2}\right)\right) - \dfrac{G_r}{\mathrm{R}_e^2}T \end{array}\right] \tag{15}$$

$$\frac{\partial u_\theta}{\partial t} + \left(\frac{\partial u_\theta}{\partial r} + \frac{u_\theta}{r}\right)u_r + \frac{u_\theta}{r}\frac{\partial u_\theta}{\partial r} = -\frac{1}{r}\frac{\partial P}{\partial \theta} + \frac{1}{\mathrm{R}_e}$$

$$\left[\begin{array}{r} \dfrac{\partial u}{\partial r}\left(\dfrac{\partial u_\theta}{\partial r}-\dfrac{u_\theta}{r}+\dfrac{1}{r}\dfrac{\partial u_r}{\partial \theta}\right)+\mu\left(\begin{array}{c}\dfrac{\partial^2 u_\theta}{\partial r^2}-\dfrac{1}{r}\dfrac{\partial u_\theta}{\partial r}+\dfrac{u_\theta}{r^2}\\ +\dfrac{1}{r}\dfrac{\partial^2 u_r}{\partial\theta\partial r}-\dfrac{1}{r^2}\dfrac{\partial u_r}{\partial\theta}\end{array}\right)+ \\ \mu\left(\dfrac{\partial u_\theta}{\partial r^2}+\dfrac{u_\theta}{r^2}-\dfrac{1}{r}\dfrac{\partial u_\theta}{\partial r}\right)+\dfrac{2\mu}{r}\left(\dfrac{2}{r}\dfrac{\partial u_r}{\partial\theta}-\dfrac{u_\theta}{r}+\dfrac{\partial u_\theta}{\partial r}+\dfrac{1}{r}\dfrac{\partial^2 u_\theta}{\partial\theta^2}\right) \\ +\dfrac{2\mu}{r}\left(\dfrac{u_\theta}{\partial r}-\dfrac{u_\theta}{r}\right)+\mu\left(\dfrac{\partial u_\theta}{\partial z^2}+\dfrac{\partial^2 u_z}{\partial\theta_z}\right)\end{array}\right] \tag{16}$$

$$\frac{\partial u_z}{\partial t}+\frac{u_\theta}{r}\frac{\partial u_z}{\partial\theta}=\frac{\partial P}{\partial\theta}+\frac{1}{\mathrm{R}_e}\left[\begin{array}{r}\dfrac{\partial u}{\partial r}\left(\dfrac{\partial u_z}{\partial r}+\dfrac{\partial u_r}{\partial z}\right)+\mu\left(\dfrac{\partial^2 u_z}{\partial r^2}+\dfrac{\partial^2 u_z}{\partial z\partial r}\right)+ \\ \dfrac{\mu}{r}\left(\dfrac{\partial u_z}{\partial r}+\dfrac{\partial u_r}{\partial z}\right)-\dfrac{Gr}{\mathrm{Re}^2}T\end{array}\right] \tag{17}$$

$$\frac{\partial T}{\partial t}+u_r\frac{\partial T}{\partial r}+\frac{u_\theta}{r}\frac{\partial T}{\partial\theta}+u_z\frac{\partial T}{\partial z}=\frac{1}{\mathrm{R}_e\mathrm{P}_r}\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}+\frac{1}{r^2}\frac{\partial^2 T}{\partial\theta^2}+\frac{\partial^2 T}{\partial z^2}+ \\ \frac{N_{Gn}}{\mathrm{R}_e P_r}\mu\left(\frac{\partial u_\theta}{\partial r}-\frac{u_\theta}{\partial r}\right)\frac{\partial u_\theta}{\partial r}-\frac{1}{r}\left(\frac{\partial u_\theta}{\partial r}-\frac{u_\theta}{\partial r}\right) \tag{18}$$

$$\frac{\partial u_r}{\partial r}+\frac{u_r}{r}+\frac{1}{r}\frac{\partial u_\theta}{\partial\theta}+\frac{\partial u_z}{\partial_z} \tag{19}$$

3.1 Boundary Conditions

At inside and outside wall of the cylinders, the boundary conditions assumed are that there is no slip and there is no penetration for each disturbance velocity components. At the wall, all disturbance velocity components are zero. Another boundary condition taken into account is that at wall pressure do not exists.

$$u_r=(x,r_{out})=0;\ u_0=(x,t_{out})=0;\ u_z=(x,r_{out})=0 \tag{20}$$

$$u_r=(x,r_{out})=0;\ u_0=(x,t_{out})=0;\ u_z=(x,r_{out})=0 \tag{21}$$

Though, the compatibility regulations resulted from the Linearized N-S equations have been collocated at the solid wall.

$$\frac{1}{R}\left[2\overline{\mu}\frac{\partial^2 u_r}{\partial r^2}+2\frac{\partial\overline{u}}{\partial_r}\frac{\partial u_r}{\partial r}+\frac{2\overline{u}}{r}\frac{\partial u_r}{\partial_r}\right]-\frac{1}{\mathrm{R}_e}\left[\frac{i\alpha\overline{u}}{r}\frac{\partial u_\theta}{\partial_r}\right]-\frac{1}{\mathrm{R}_e}\left[i\beta\overline{u}\frac{\partial u_\theta}{\partial_r}\right]+\frac{\partial P}{\partial r} \tag{22}$$

The CSC technique is used to discretize the governing stability equations [17]. This discretization makes the non-uniform nature of the distribution of the grid points with a greater number of grids towards the end. For boundary value problems, it is a favorable arrangement.

$$r_{cheb} = \cos\left(\frac{\pi i}{n}\right) \text{ where i = 1, 2, 3 . . . n} \tag{23}$$

where n = Total number of collocation points.

The stability equations together with the boundary conditions forms a general EVP of the form,

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} \end{bmatrix} \begin{bmatrix} u_r \\ u_\theta \\ u_z \\ p \\ T \end{bmatrix} = i\omega \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} & B_{15} \\ B_{21} & B_{22} & B_{23} & B_{24} & B_{25} \\ B_{31} & B_{32} & B_{33} & B_{34} & B_{35} \\ B_{41} & B_{42} & B_{43} & B_{44} & B_{45} \\ B_{51} & B_{52} & B_{53} & B_{54} & B_{55} \end{bmatrix} \begin{bmatrix} u_r \\ u_\theta \\ u_z \\ p \\ T \end{bmatrix} \tag{24}$$

$$[A][\phi] = i\omega[B][\phi] \tag{25}$$

where, A and B are real matrices of size 5 × n , iω is an eigenvalue and φ is the eigenvector. The QZ algorithm is employed for the solution of the EVP.

4 Base Flow Solution

The fully-developed steady and parallel incompressible base flow is considered between the concentric rotating cylinders in opposite directions. The base is varying along the radial direction only. The thermal stratification is also considered in the radial direction. Thus, the viscosity is varying and it is dependent on the temperature of the fluid. The energy equation is coupled with the Navier-Stokes equation and the reduced non-dimensional equations for the base flow are as follow,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + N_{Gn}\mu\left[r\frac{\partial}{\partial r}\left(\frac{U_\theta}{r}\right)\right]^2 = 0 \tag{26}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + N_{Gn}\mu\left[r\frac{\partial}{\partial r}\left(\frac{U_\theta}{r}\right)\right]^2 = 0 \tag{27}$$

The conditions (28) and (29) are co nsidered for the solution of above equations.

$$u_\theta(x, r_{in}) = +1, T(x, r_{in}) = r_{t1} \tag{28}$$

$$u_\theta(x, r_{out}) = -1, T(x, r_{out}) = r_{t_2}; \text{ where } r_{t_1} = \frac{T_1 - T_0}{\Delta T_0}; r_{t_2} = \frac{T_2 - T_0}{\Delta T_0} \quad (29)$$

Stream wise (azimuth) velocity, temperature and viscosity respectively. The base flow equations for constant viscosity are simple and solved analytically and it is not presented here because the variation in the base flow velocity profile is very small. The Eqs. 28 and 29 are solved with the help of series solution up to second order accuracy [6, 15].

Figure 3a shows the effect of non-dimensional temperature (r_t) on the temperature profile of the flow. The $r_t = 0$ is the isothermal flow in which the inside and outside cylinder ramparts are maintained at equal temperature. Figure 1b presents the temperature profile for various values of Na. It shows that the increase in Na, increases the variation in temperature. The escalation in Brinkman number (Br) also surges the temperature variation for non-isothermal flow.

5 Code Validation

The results obtained for the present computations are compared with the results of P. Schmid and L. Tuckerman (2002) for $\eta = 0.5$ and $\eta = 0.99$ without viscid heating effect. The azimuthal (α) and axial (β) wave numbers are 0 and $\pi/2$. The eigen spectrum shown in Fig. 2a, b are very similar to the results of Schmid et al. [18]. Thus, the code is validated against the results of P. Schmid et al. (2002) (Fig. 4).

6 Results and Discussions

In the present stability analysis, two cylinders are revolving in contrary directions, and hence the speed ratio Ω_1 is as -1. The range of different values of η is changing from 0.5 to 0.99. The radial length and the velocities are controlled by the space, d $= (R_2 - R_1)$ and $\Omega_1 R_1$ respectively. The magnitude of α and β are zero and unity respectively. The stability analysis for the Ta–Co flow is with the variable viscosity due to the viscous heating effect. The generation numbers like Nehme (Na) number and Brinkman (Br) numbers are defined to incorporate the variation of viscosity and temperature on the steadiness representative of the Ta–Co flow.

6.1 Effect of Radius Ratio (η)

In the case of the adiabatic wall conditions, the fluid viscosity remains constant. The Eigenvalues problem is solved to compute the least stable Eigenvalue and associated

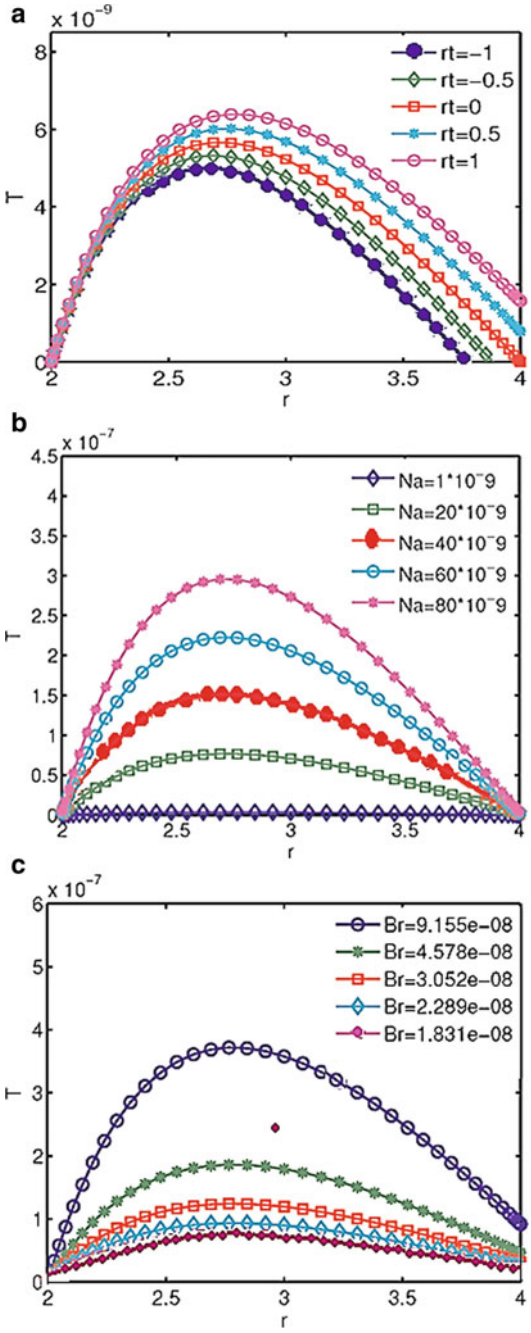


Fig. 3 Effect of **a** temperature (r_t), **b** Nahme number (Na) and **c** Brinkman number (Br) on temperature profile for radius ratio $\eta = 0.5$ speed ratio $\frac{\Omega_1}{\Omega_2} = -1$

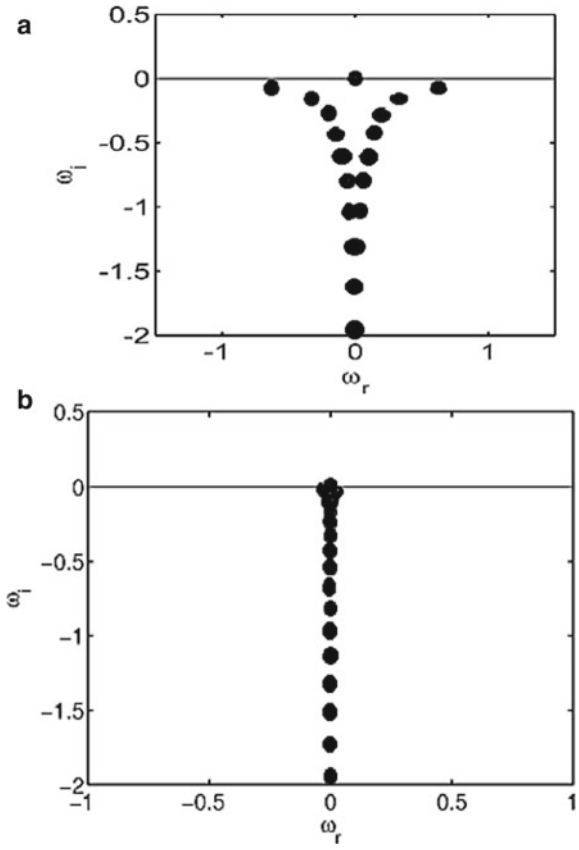


Fig. 4 Eigen spectrum for **a** $\eta = 0{:}5$ and $\mathrm{Re} = 125$ **b** $\eta = 0{:}99$ and $\mathrm{Re} = 350$. The α, β and speed ratio are 0, $\frac{\pi}{2}$ and -1

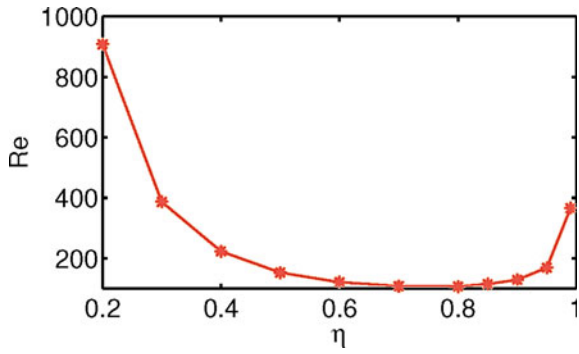


Fig. 5 Effect of radius ratio η on the R_{cr}

Eigenfunctions for different values of η. The perilous value of Re is determined for the range of various values of radius ratio.

Figure 5 shows the change in the value of the critical Reynolds number for different radius ratio η. It shows that R_{cr} number reduces as η increases from 0.2 to 0.7, it is nearly constant in the range of η 0.7 to 0.8 and beyond 0.8 it increases.

6.2 *Effect of Viscous Heating*

To study the viscid heating effect on the instability characteristic of Ta–Co flow, two different cases are considered. In the first case, the inside and outside boundaries of the cylinders are maintained at a fixed temperature of 25 °C, and so it is called isothermal Ta–Co flow. The viscosity of the fluid is variable due to the heating effect. The R_{cr} number is computed for $\eta = 0.5$ and $\eta = 0.99$.

Table 1 shows the assessment of the critical value of Re for $\eta = 0.5$ and $\eta = 0.99$ for adiabatic and isothermal wall conditions of the rotating cylinders. The critical value of Re is lower for isothermal Ta–Co flow compared to adiabatic wall of the cylinders. This indicates that Taylor–Couette flow becomes unstable with the viscous heating effect at a lower Reynolds number.

Figure 6 shows the change in the value of critical Re for different Nahme numbers. Na introduces the viscous heat dissipation. The increase in the speed of rotation increases the heat dissipation effect. The Na number is varied by increasing the rotating speed of the wall of both the cylinders. It has been found that the critical value of Re reduces with the rise in Na Number. Thus, the increase in Nahme number has destabilizing effect on the Taylor–Couette flow. Table 2 shows a comparison of R_{cr} number for isothermal Ta–Co flow for $\eta = 0.5$ and $\eta = 0.99$. It shows that the R_{cr} number reduces with the viscid heating effect. This, in turn, proves that viscous heating has a destabilization effect on the disturbances. In the second case, the inside and outside walls of the cylinders are maintained at different values of constant temperature. The temperature gradient is present in the radial direction. The viscous heating effect is introduced in this case using Brinkman's number. The Brinkman number reduces with the increase in temperature difference of cylinder walls. The variation of critical R_e against Br is shown in Fig. 4b. It shows that the critical value of R_e raised with the rise in the value of Br. Thus, an increase in the Brinkman number stabilizes the flow. Table 3 shows the critical value of Re for different values of ΔT0 for non-isothermal Taylor–Couette flow. It shows that as the ΔT_0 rises, the critical value of Re decreases. It is also witnessed that the critical R_e for non-isothermal Taylor–Couette flow is smaller than that of isothermal Taylor–Couette flow.

Table 1 Assessment of the R_{cr} for isothermal Ta–Co flow

Radius ratio	Critical R_{cr_1}	Critical R_{cr_2}
0.5	153	130
0.99	366	312

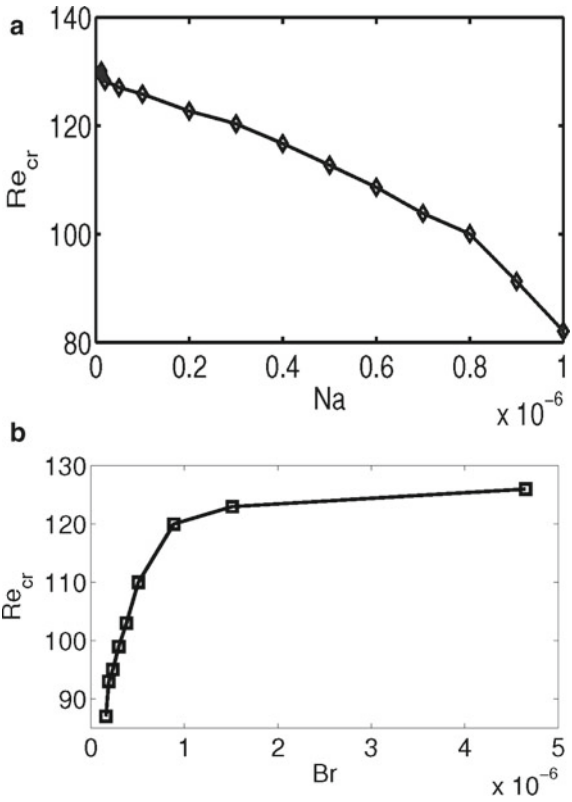


Fig. 6 Effect of **a** Nahme number, Na and **b** Brinkman man number, Br on the R_{cr} number for $\eta = 0{:}5$ and speed ratio $\frac{\Omega_1}{\Omega_2} = -1$

Table 2 Assessment of Rcr number for isothermal Ta–Co flow with buoyancy effect

Radius ratio	Critical R_{cr_1}	Critical R_{cr_2}
0.5	130	128
0.99	312	310

Table 3 Comparison of R_{cr} number for non-isothermal Ta–Co flow with buoyancy effect. Here ΔT_0 is the temperature difference of walls of inner and outer cylinders

Radius ratio		Critical R_{e1}	Critical R_{e2}
0.5	−0.2	131	128
	0.5	120	118
	1.0	103	100
0.99	−0.2	320	310
	0.5	304	300
	1.0	296	292

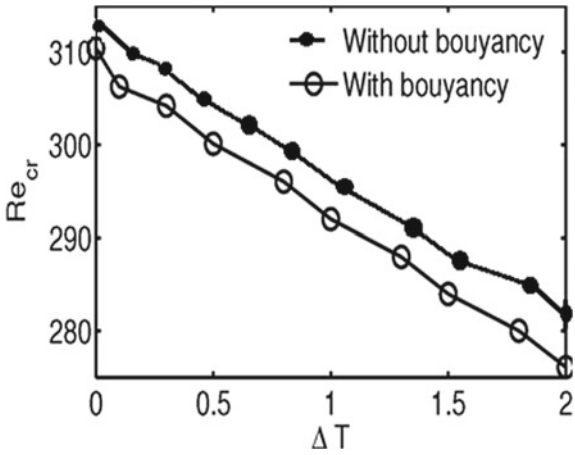


Fig. 7 Variation of Rcr number versus temperature

6.3 Effect of Buoyancy

The effect of buoyancy is studied for isothermal and non-isothermal Ta–Co flow. The effect of buoyancy is combined in the governing equations using Gr number.

From the above results, we observe that in isothermal Ta–Co flow with buoyancy, with the increase in radius ratio from 0.5 to 0.99 there is an increase in R_{cr}.

Figure 7 shows the variation of the Rcr number with the reference temperature difference. It is found that the buoyancy effect reduces the critical Re for the same temperature difference for $\eta = 0.5$ and $\eta = 0.99$.

7 Conclusions

The local temporal Eigenmodes are computed using linear stability theory for Ta–Co flow. The effect of viscid heating and buoyancy are studied on the steadiness characteristics of Ta–Co flow. To study the effect of viscid heating for isothermal and non-isothermal Taylor–Couette flow Nahme and Brinkman numbers are introduced. The general Eigenvalue problem is solved to determine the least stable Eigenmode. For adiabatic wall conditions, it is found that the critical value of Re reduces with the increase in radius ratio up to $\eta = 0.8$, beyond which critical value of Re increases. The isothermal and non-isothermal Taylor–Couette flow is studied by introducing generation number like Nahme and Brinkman number. The critical value of Re increases with the rise in Nahme number while reduces with the increase in Brinkman number. The critical value of Re reduces with the increase in the temperature difference of cylinder walls. It is also observed that the flow becomes unstable at a lower Reynolds

number in case of a non-isothermal flow. The critical value of Re with the buoyancy effect is found smaller for isothermal and non-isothermal. Thus, buoyancy has a stabilizing effect on the disturbances.

References

1. Martinez JJ et al (2007) A Chebyshev collocation spectral method for numerical simulation of incompressible flow problems. J Braz Soc Mech Sci Eng
2. Thomas DG (2004) Thermo-mechanical instabilities in Dean and Taylor–Couette flows mechanisms and scaling laws. J Fluid Mech
3. Drazin P, Raid WH (2004) Hydrodynamic stability. Cambridge University Press. https://doi.org/10.1017/CBO9780511616938
4. Taylor GI (1923) The spectrum of turbulence. Roy Soc London, Ser Math Phys Sci 223:289
5. Belotserkovskii OM, Oparin AM et al (2016) Coherent hydrodynamic structures and vortex dynamics. Math Model Comput Simuls, Math Phys 135–148
6. White JM, Muller SJ (2002) Experimental studies on the stability of Newtonian Taylor–Couette flow in the presence of viscous heating. J Fluid Mech
7. Ali M, Weidman PD (1990) On the stability of circular Couette flow with radial heating. J Fluid Mech 53:220
8. Kolyshkin AA, Vaillancourt R (1997) Convective instability boundary of Couette flow between rotating porous cylinders with axial and radial flows. Phys Fluids 5:3136
9. Kuo DC, Ball KS (1997) Taylor–Couette flow with buoyancy: Onset of spiral flow. Phys Fluids 9:2872
10. Renardy Y, Joseph D (1985) Couette flow of two fluids between concentric cylinders. J Fluid Mech 150:381
11. Mosta S, Sinbanda P (2010) A novel numerical technique for two-dimensional laminar flow between two moving porous walls. In: Mathematical problems in engineering Hindawi Publication
12. Papathanasion TD (1968) Thermomechanical coupling in frictionally heated circular Couette flow. Int J Thermo-Phys 18:825
13. White M, Mullar J (2000) Viscous heating and the stability of newtonian and viscoelastic taylor–couette flows. Am Phys Soc 84:5130
14. Al-Mubaiyedh UA, Sureshkumar R (1999) Influence of energetics on the stability of viscoelastic Taylor–Couette flow. Phys Fluids 11:3217
15. Al-Mubaiyedh UA, Sureshkumar R et al (2002) The effect of viscous heating on the stability of Taylor–Couette flow. J Fluid Mech 46:111
16. Dou H, Khoo B, Yeo K (2008) Instability of Taylor–Couette flow between concentric rotating cylinders. Int J Therm Sci 46:262
17. Johnny de Jesús Martinez (2007) A Chebyshev collocation spectral method for numerical simulation of incompressible flow problems. J Braz Soc Mech Sci Eng 29:3
18. Schmid P, Tuckerman L (2002) Transient growth in Taylor–Couette flow. Phys Fluids 14:3474
19. White JM (2000) Viscous heating and the stability of newtonian and viscoelastic Taylor–Couette flows. Phys Rev Lett

Advanced Numerical Methods Based on Optimization

Marian Gaiceanu, Vasile Solcanu, Theodora Gaiceanu, and Iulian Ghenea

Abstract In this chapter the unconstrained and constrained optimization algorithms for numerical methods are envisaged. The numerical solutions to the fundamental problems in energy systems are provided. The recent heuristic methods used in power systems are highlighted, and the specific algorithms are proven by simulations. The selection of the best solution is one of the authors' concern. Therefore, the optimization algorithms along with numerical examples are delivered in this chapter.

Keywords Numerical methods · Newton · Newton-Raphson · Optimization algorithms · Trust region · Genetic · Mixed integer nonlinear programming · Heuristic method · Ant colony · PSO · Firefly · Simulated annealing · Differential evolution · Bee swarm · Unit-commitment · Very large scale systems

M. Gaiceanu (✉)
Department of Automatic Control and Electrical Engineering, University Dunarea de Jos Galati, Galati, Romania
e-mail: marian.gaiceanu@ugal.ro

V. Solcanu
University Dunarea de Jos Galati, Galati, Romania
e-mail: vasilesolcanu@dedeman.ro

T. Gaiceanu
Faculty of Automatic Control and Computer Engineering, Gheorghe Asachi Technical University of Iasi, Iasi, Romania
e-mail: gaiceanu.theodora@ac.tuiasi.ro

I. Ghenea
Doctoral School of Fundamental and Engineering Sciences, University Dunarea de Jos of Galati, Galati, Romania
e-mail: iulian.ghenea@ugal.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_8

Nomenclature

BFGS	Broyhen-Fletcher-Goldfarb-Shanno
CD	Conjugate Directions
CNO	Constrained Nonlinear Optimization
DFP	Davidon-Fletcher-Powell
FR	Fletcher–Reeves
GA	Genetic Algorithms
HA	Heuristic Algorithms
LQR	Linear Quadratic Regulator
MINLP	Mixed Integer Nonlinear Programming
MMCG	Modified method of conjugate gradients
NM	Newton method
NO	Nonlinear Optimization
NRM	Newton-Raphson method
PR	Polak–Ribiere
PS	Power Systems
PSO	Particle Swarm Optimization
OC	Optimal Control
TSP	Travelling salesman problem
TR	Trust region
SA	Simulated Annealing
SCUC	Security-Constrained Unit CommitmentUnit Commitment
SIP	Semi-infinite programming
SDP	Semi-defined programming
UC	Unit Commitment
UNO	Unconstrained Nonlinear Optimization
VC	Variational Calculus

Symbols

C	convex set
d_k	descent direction
Δ	distance
F	convex function
$f(x)$	real function of variable x
$f'(x)$	first derivative
$grad(f(x^*))$	gradient of the multivariable function
$\nabla f(x^*, y^*)$	the gradient of function f at the minimum point (x^*, y^*)
$\nabla^2 f(x_k) = D2f_k$	Hessian matrix
$\mathbf{K}$ $(\mathbf{x}, t)$	integral cost
$\mathbf{K}$ $(\mathbf{x}_{f,}\ tf)$	final cost

$\mathbf{J}(\mathrm{u})$	performance index
$L(x, u, \tau)$	Lagangian
p	Lagrange multiplier
p_i^C	Cauchy point
$\mathbf{x}$	unknown vector
$s^*(\Delta)$	steepest descent direction
$\|s\| = \Delta$	norm
span	linear combination operator

1 Introduction

Optimization offers the best solution, out of the many possible solutions to a problem, according to a performance criterion imposed, respecting or not certain constraints. In the optimization problems, the performance criterion is chosen according to the proposed objectives, observing certain constraints or not. Multivariable functions are used in variational calculus (VC). VC deals with the functionals, the solution being obtained more often by minimizing the functional. By taking into account the Sect. 1.1 of this book, the evolution of the VC is optimal control (OC). OC is a result of the functional cost minimization, the control being dependent by the state feeback product of the system, the cost function, and a weighting matrix.

The problem of determining the shortest possible route between two points is the oldest optimization problem, the well-known solution being a straight line segment.

Greek mathematicians Zenodorus and Poppus studied the problem of Princess Dido (Elisa) from Tiria or the isoperimetric problem, inspired by the historical story of Vergilius (70-19 BC), Eneida, about the formation of Carthage (850 BC): to find a plane-generated curve of a given straight segment, covering the largest possible area. The Greeks knew the solution (the circle), but it was only in the 29th century rigorously demonstrated.

Heron of Alexandria gave another interpretation of the shortest path possible, in the paper Katoprika (Principles of optics), noting that if a light source emits a light beam reflected in a mirror; it will follow the shortest possible path from source to observer.

The origins of variational calculus date from 1662 with the Fermat principle the fastest path of a light beam which passes through a single optical environment is that of minimum time.

The first minimum time problem formulated and demonstrated was in 1697 ("brachystochrone problem"–"βραρχιστοζ-Short, χρονο-time"), by Johann Bernoulli. In this way the optimal command appeared.

Modern optimal control

The practical point ov view of OC was discovered by Riccati (1676–1754). Jacopo Riccati had obtained a scalar solution of the nowadays linear quadratic problem.

Kalman (1960), Athans and Falb (1966), push forward the development of the modern optimal control.

Dynamic programming and the principle of maximum

Dynamic programming orginated from Bellman (1950s), takes into account the nonlinear optimal command. The main disadvantage of its solution is the large memory requiremnts. The principle of maximum, developed by Pontryagin and the Soviet school, is a natural consequence of Weierstrass's necessary condition with limited command functions. To determine the shortest route, the optimal control maximizes the Hamiltonian within the boundary of definition. The maximum principle is normally found in dynamic programming.

The *nonlinear problem* of determining optimal trajectories was addressed in 1919, by Robert H. Goddard (1882–1945). Using variational calculus, in 1927 Hamel formulates the nonlinear problem, the analytical solution being found by Tsien and Evans in 1951.

The modern optimal control deliver a more robust solution (in frequency domain) through $H\infty$ and H_2 approaches. The robustness of the modern control is related to the systems with structural uncertains, and unmodelled dynamics.

Research in the field of optimization is on-going by using artificial intelligence, Genetic algorithms and metaheuristics approaches.

The objective of this chapter is to provide a methodology to solve the specific problems in power system applications. Moreover, due to the high importance of energy production, the solutions should be the best one. Therefore, the optimization algorithms are also taken into account. In order to facilitate the understanding of some algorithms, the authors provide numerical exmaples.

In this chapter the formulation of the optimization problems are presented. The Sect. 3 contains the classification of the optimal problems. Different types of the optimizations are included in Sect. 4. Unconstrained and constrained optimization algorithms for numerical methods are provided in Sects. 5–13. The recent heuristic methods and the specific algorithms are shown in Sect. 14. The Sect. 15 provides an overview of the genetic algorithms. In Sects. 16–19, different nature inspired optimization algorithms (metaheuristics) are found: ant colony optimization, Simulated Annealing, Particle Swarm Optimization, Bee Swarm, and Firefly model. The last Sects. (20, 21) are dedicated to very large scale neighborhood search, and security-constrained unit commitment. The chapter ends with the conclusions.

2 Formulation of the Optimization Problem

In order to formulate a problem of optimization of a dynamic system, the following steps must be taken into account: (1) the initial and final conditions; (2) the dynamics of the system; (3) the control limits; (3) the objectives of the problem; (4) performance criterion. The optimal solution will be an admissible control delivered by minimization of the peformance criterion [1].

Implementation of the optimal solution

With the analytical or numerical determination of the optimal solution, the natural problem of implementing this solution found for the considered dynamic system is posed.

In the references [1, 2] the optimal command was obtained either by solving a Riccati differential matrix equation (for problems with fixed final time), or from an Riccati algebraic matrix equation (for problems with infinite horizon or final free time). Riccati equations require a reverse integration in time (from the final step to the initial step) and the storage of specific coefficients on this interval, following the application of the optimal command in the direct sense of time (from the initial step to the final step). This procedure requires a large volume of calculus and is sensitive to the variation of the dynamic system parameters. The solutions that eliminate the mentioned disadvantages are presented in the technical papers [1, 2] are described modern methods of implementing the solutions within modern electric drive systems.

Types of constraints

Constraints are applying to control and states of the dynamic system under consideration. At the level of the relations, they must be within an allowable range (*inequality constraints*), and at the state level, as a result of the existence of differential equations that must be solved, their initial and final values (*equality* constraints) must be known. Another variable to consider is *time*. Therefore, *the initial and final time* are all *equality* constraints.

Performance index

The performance index (cost function) is designed by the designer to meet the objectives set. It usually includes a *final cost* and an *integral cost*, corresponding to the stationary and dynamic regimes. From a systemic point of view, the performance index contains system errors or states, and commands. For a system of order n (n states) with m inputs ($m \leq n$), the performance index is defined by the final cost K ($\mathbf{x}$, t) and the integral cost:

$$J(\mathbf{x_0}, t_0, \mathbf{u}, \mathbf{x}, t) = K(\mathbf{x}, t) + \int_{t_0}^{t} L(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau \tag{1}$$

or by knowing the final time, t_f, it becomes:

$$J(\mathbf{x_0}, t_0, \mathbf{u}) = J\left(\mathbf{x_0}, t_0, \mathbf{u}, \mathbf{x}_f, t_f\right) = K\left(\mathbf{x}_f, t_f\right) + \int_{t_0}^{t_f} L(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau, \quad (2)$$

where

- $J(\mathbf{x_0}, t_0, \mathbf{u})$ is the value of the system performance index for the command $\mathbf{u}$ applied,
- t_f the final time,
- $\mathbf{x}_f$ final state,
- K (x $_{f,}$ *t f) terminal cost*.

The performance index can stretch to infinity if no command exists to bring the system to a specified end state:

$$J(\mathbf{x_0}, t_0, \mathbf{u}) \rightarrow \infty. \quad (3)$$

It is noted with:

$$J(\mathbf{u}) = J(\mathbf{x_0}, t_0, \mathbf{u}) \quad (4)$$

if the initial state $\mathbf{x_0}$ is known at time $t_{0.}$

Solving the optimal control problem leads to the determination of *the optimal regulator.* The regulator will minimize the chosen performance index.

Example: If the energy absorbed by a dynamic system should be minimized, then the instantaneous quantities of the absorbed current and the supply voltage will be included in the performance index, and the controller will drive the system so as to minimize the energy absorbed from the power source.

Remarks:

1. Any solution to the optimal driving problem is called *optimal control;*
2. In the optimization problems, the poles of the system are not directly imposed, but result from the minimization of the chosen performance index;
3. Only the determination of the minimum of the performance index can be considered, because the problem of determining the maximum is similar to finding the minimum of the negative function.

3 Classification of the Optimal Problems

Optimal problem with quadratic functional applied to linear systems are well known as *linear quadratic regulators* (LQR) [2]. The LQR problems minimizes the consumed energy into system. LQR problems can be classified as:

- the problem of adjustment according to the state;
- the problem of the adjustment according to the output.

The linear quadratic optimal driving problems lead to solving different types of the Riccati equations: algebraic, differential (smooth case) or in differences (discrete case), solution that acts as a gain factor on the state feedback of the system.

The solution which minimizes the functional cost is named *optimal control* (OC). OC can be with or *without* constraints, with *fixed* or *free time* problems.

Particular cases:

(1) *The tracking problem,* when a certain trajectory becomes an objective to be achieved;
(2) *The problem* of *optimal control* with *fixed time,* when it is constraint to the final time t_f. The performance index becomes:

$$J(\mathbf{u}) = K(T) + \int_{t_0}^{T} L(x, u, \tau) d\tau \tag{5}$$

(3) *The problem* of control *with final state fixed, free final time,* when a desired state is required, $\mathbf{x}_f$.
(4) *The adjustment problem* when the fixed end point satisfy:$\mathbf{x}^* = \lim_{t \to \infty} \mathbf{x}(t)$, i.e. $\mathbf{x}^*$ is a state of equilibrium of the system.

4 Types of Optimizations

Continuous optimization

Models with continuous variables that can take any real value are problems of continuous optimization

Discrete optimization

The variables into the discrete optimization problems are included into a set of the discrete values.

Optimization without constraints—the constraints do not found on model variables.

Optimization with constraints—*the constraints* are presented on the model variables.

Optimization problems with ***constraints*** can be reduced to problems without constraints by replacing them with penalty terms in the chosen objective function.

The types of the optimization problems with constraints can be found by taking into account both the nature of the constraints and property of the functions (differentiable; non-differential). The constraints could be linear, nonlinear, or convex type.

Optimization problems without objective function

In practice there are problems without a special optimization objective, e.g. to find the variable values by satisfying the imposed model constraints.

This type of problem is named the *feasibility* problems

Optimization problems with objective functions

Usually, the one objective function into the optimization problems can be found. Moreover, there are special cases with multiple objective functions.

Problems of complementarity and variational inequalities are ubiquitous in engineering field, combined or not with economic field. The main objective is to satisfy the conditions of complementarity by the deducted solution.

The *optimization* problem with *multi-objective functions* occurs in logistics, or enginnering combined or not with economics. In these types of optimization problems the solution is a compromise of the goals conflict.

However, problems with multiple objectives can be reformulated as a single objective problem by forming a weighted combination of the functions objectives, or changing part of the function objectives by the adequate constraints.

Deterministic optimization

In this type of *optimization*, it is supposed that the required data to solve the problem is precisely known. Many times in practice, the data is missing (the often case is that there are errors in measurements, or for forecasting problems the future data cannot be known accurately).

Stochastic optimization

There are problems with the uncertain parameters into the used model in which the optimization stage is necessary. The robust control field takes into account these type of the processes conducting to a robust optimal solution. The robustness character is justified by controlling the process within a certain limits, but without knowing the all process data (structural uncertains). In stochastic problems, the advantages of knowing the probability distributions of the problem data are used; the objective is to find a solution to optimize the model performances for any data inputs.

Combinatorial optimization

The study of the arrangements of the elements of a *finite* set, according to a given *structure*, leads to *combinatorics*. In *combinatorial problems*, priorities are problems related to *existence* (there is a particular type of arrangement) and *counting* (the number of arrangements that can be formed):

The characteristic of combinatorial problems is that the number of these arrangements is finite.

The combinatorial optimization principle involves comparing the arrangements based on a criterion and selecting the best arrangement according to that criterion. The purpose of combinatorial optimization is to solve specific problems, i.e. to develop methods and algorithms for effectively finding the most suitable "arrangement" from the finite set of all possible arrangements.

Travelling salesman problem (TSP) is one of the combinatorial optimization problem.

Travelling salesman is located in town 0 must visit in localities *1,2,…,n* and finally returning from where he left off. Knowing the distances between localities (or the costs of traveling between them) the solution to find the route with the minimum total length (or the minimum total cost) is a concern. It can be easily deduced that the total number of the possible routes is $n!$, a route being perfectly determined by the order in which the ones will be visited in the localities. Once this order is established, the calculation of the length or costs of the corresponding route is a simple operation of summing the lengths or costs of the sections that make up the route.

The mathematical model of the TSP

By consider *0,1, …, n* the set of the towns, in which 0 is the town from which the travel salesman depart and arrive at the end of the travel. Is denoted by c_{ij} *cost of* displacement from the locality *i* to the $j \neq i$. If there is not direct connection between the *i* and *j* or in the case of $i = j$ then $c_{ij} \to \infty$. A *route* will be described using the *bivalent* variables: if the route starts from city *i* directly to city *j*, $i \neq j$ then $x_i = 1$, otherwise, $x_i = 0$ (or if $i = j$).

In these conditions the following function is defined

$$f = \sum_{i,j=0}^{n} c_{ij} x_{ij}. \tag{6}$$

The constraints of the problem are given by the relations (7–8):

$$\sum_{j=0}^{n} x_{ij} = 1, \quad i = 0, 1, \ldots, n \tag{7}$$

and taking into account that from any city *i* the traveling salesman should straighten to the another single location,

$$\sum_{i=0}^{n} x_{ij} = 1, \quad j = \overline{0, n} \tag{8}$$

and at any time the travel salesman comes from a place previously visited.

Equations (6)–(8) represent the mathematical model of the *minimum tour* problem.

Heuristic methods for solving TSP

Usually, the optimal solutions of a TSP is very hard to find (if not impossible to find) if the number of the visiting cities are higher than of 50. To determine an acceptable solution more heuristic procedures have been developed. These procedures are attractive from two points of view:

- Can give a “guarantee certificate” for the obtained solution in the sense of the possible evaluation of the “maximum exceeds” from the optimal solution;

- The approximately solution can be found with a moderate *computational* effort into a *reasonable time*, that it meant those two parameters are *polynomial* dependent from the size of the problem (i.e. the number of the cities);

The heuristic methods builds a solution *by trial, by making at each iteration the best possible choice.* Unfortunately, this scheme *does not* usually lead on *the best globally solution.*

Next, several heuristics are presented to solve *the Euclidean problem of the travel salesman,* i.e. of the problem in which $c_{ij\,j}$ are *distances* satisfying both symmetry conditions: $c_{ij} = c_{ji}$, and triangle inequality $c_{ik} \leq c_{ij} + c_{jk}$.

Heuristics—*nearest neighbor*

- Departure from the town 0 to the *nearest* town;
- From the last visited village bows by the nearest unmarked village; in case that there is no longer single location to be visited, returns in place of departure;

Most combinatorial optimization problems are modeled using graph theory. Most of them can be described alternatively by linear programs with integer variables, especially bivalents, hence the close link between combinatorial optimization and integer programming.

Among the problems of combinatorial optimization are:

- the problem of the *minimum path* value between two nodes of a graph;
- the problem of *maximum flow*;
- the problem of *minimum tour*.

Another heuristic problem is the **problem of the minimum cost tree**: a number of "points" must be connected to facilitate the transmission of a certain service. Between points there are "potential links" whose realization involves a certain cost. The problem that arises is to see what connections will actually be made in such a way that any two points will be connected—directly or indirectly—in order to use the service, and the sum of the costs of the connections made will be minimal.

5 Unconstrained Optimization Problems

Overview

Numerical methods for solving unconstrained optimization problems are applied to the objective functions of one or more variables and aim to determine the global extremum point of the function.

From the calculated values point of view, the methods of solving the optimal problems without constraints can be classified in [3–8]:

- *Direct methods*—no derivatives of the objective function are required;
- *Indirect methods*—the derivatives of the objective function are used.

The difference between the direct and indirect methods is mainly based on the presence of the derivatives of the objective function. In the direct methods, there are no objective function derivatives.

The methods that use derivatives calculus have the advantage of a high convergence speed, but the volume of calculations increases, and errors can occur.

Depending on the used principle, the methods are divided into:

1. **Exploration** methods
2. **Removal** methods
3. **Search** methods (first order, second order)

1. The first two methods (**exploaration**, and **removal**) are used to find the domain in which the extremum of the function is located, but are used more in the case of functions of two or maximum three variables, since they are difficult to apply to functions of several variables for one-dimensional extremities.

2. ***Removal Methods.***

The objective function must satisfy the unimodality hypothesis (there is only one extreme point on the definition domain).

Principle: the considered area is divided into two parts by a (segment) separation plan; the value of the function is tested in the two sub-domains by a specific procedure and the one that is not of interest is eliminated; the procedure is continued for the remaining domain by dividing with a separation plan, etc., until a sufficiently small domain is reached, depending on the accuracy required.

3. **First order methods. Linear search methods**

They are the most efficient methods of solving unconstrained problems.

The principle of these methods consists in the iterative approximation of the extremum point, the procedure ending when the stop criterion is satisfied.

The general calculation formula is:

$$x^{i+1} = x^i + \theta_i h^i$$

where

h^i—is the direction of travel

θ_i—is a scalar that represents the step of movement in the h^i direction.

3.1 Indirect Search methods or *gradient methods* apply for derivable functions f (x) in relation to all arguments.

Methods of the II-nd order, in which the derivatives of the I-st order of the function are used:

$$x^{i+1} = x^i - \theta_i^* r^i, \quad r^i = \nabla f(x^i).$$

The step can be constant, variable (decreasing) or optim (*optimal gradient* method or *Cauchy method*).

3.1.1 Methods Based on Conjugate Directions Let x_0 the starting point of the algorithm and x^* the minimum point

$$x^* = x_0 + x_s,$$

$x_s \in \mathbb{R}^n$, unknown vector.

x_s is expressed in a base by using the conjugate directions.

Starting from x_0 and making successive steps along the axes of this base, there is the possibility that in n steps x^* to be reached.

The *conjugate directions (CD) method* can be regarded as an intermediate method between the *gradient* and the *Newton algorithms. The gradient method* has access to the values of the first derivatives order. In the Newton method, the second order derivatives are used. The CD method aims to accelerate the slow convergence rate of the gradient method and at the same time avoid the use of Hessian as in the Newton method. The particular case of the *conjugate directions* method is the *conjugate gradients method*, which was initially developed for quadratic problems. By approximation, this technique has been extended to general optimization problems because it can be argued that near a local minimum point the objective function is approximately quadratic.

This method consists of the following steps:

(a) $r_0 = \nabla f(x_0)$, and $d_0 = -\nabla f(x_0)$;

(b) $x_{k+1} = x_k + \alpha_k d_k$ for any $k \in [0, n-1]$, with $\alpha_k = -\frac{r_k^T d_k}{d_k^T \nabla^2 f(x_k) d_k}$;

(c) $d_{k+1} = -\nabla f(x_{k+1}) + \beta_k d_k$, where $\beta_k = \frac{r_{k+1}^T d_k \nabla^2 f(x_k)}{d_k^T \nabla^2 f(x_k) d_k}$;

(d) at each iteration i, the variable x_0 is replaced by x_i. The above mentioned procedure repeats at each iteration.

This method has the following two main disadvanteges:

– requires the calculation of the Hessian objective function for each iteration,
– is not convergent for the general case.

3.1.2 Modified Method of Conjugate Gradients (MMCG) This method attempts to correct the disadvantages of the conjugate gradient method listed above. In this case, α_k it is calculated by using another method (ideal, Wolfe conditions or backtracking), and β_k is calculated with one of the two formulas proposed below:

– Fletcher–Reeves (FR): $\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$,
– Polak–Ribiere (PR): $\beta_k = \frac{(r_{k+1}^T - r_k)^T r_{k+1}}{r_k^T r_k}$.

The PR method behaves better than the FR method.

Thus, for MMCG problems without constraint, the following steps are followed:

(a) $r_0 = \nabla f(x_0)$ and $d_0 = -\nabla f(x_0)$;
(b) $x_{k+1} = x_k + \alpha_k d_k$, for any $k \in [0, n-1]$, with α_k calculated with Wolfe conditions or backtracking;
(c) $d_{k+1} = -\nabla f(x_{k+1}) + \beta_k d_k$, where β_k is calculated using one of the FP or PR methods.
(d) at each iteration k, replace x_0 with x_k and repeat the process described above.

3.2 Second Order Methods

I. Newton Methods

I.1 N method (NM)

The NM (1669) is part of the second order optimization method category. The conditions of sufficient optimality of the second order are: if there is an x^* which satisfies the (a, b) conditions simultaneously:

(a) $\nabla f(x^*) = 0$,

(b) $\nabla^2 f(x^*) > 0$,

then x^* can be considered as a local minimum.

An important aspect of this method is that the descent direction, $d_k (k > 0)$, can be considered if Hessian matrix is positive definite:$\nabla^2 f(x_k) = D2f_k > 0$.

For each iteration of the algorithm, the following recursive equation (quadratic approximation deducted from Taylor series expansion of the objective function) should be taken into account:

$$x_{k+1} = x_k - \theta_k (D2f_k)^{-1} D1f_k,\ D1f_k = \nabla f(x_k),$$

with step $\theta_k = 1$, and descent direction d_k:

$$d_k = -(D2f_k)^{-1} D1f_k.$$

If the starting point x_0 is not placed into the small perturbations arround of the point x^* interval, the NM cannot guarantee the convergence of the method. This is the main disadvantage of the NM. Another disadvantage of the Newton method is that at each iteration the Hessian and its inverse should be calculated.

Therefore, an improvement has been made by choosing the steps $\theta_k \neq 1$ in optimal manner, resulting so-called *Newton's variable step method*.

I.2. Quasi-Newton method

To overcome the above mentioned disadvantages, the *quasi-Newton method* is used, in which the inverse of Hessian $(D2f_k)^{-1}$ is replaced with an approximated matrix

H_k. This matrix can be calculated much easier, and at the same time keeps the fast convergence speed of the NM.

The principle of the Newton methods consists in the successive approximation of the Hessian matrix or its inverse by appealing only to the first order derivatives of the objective function.

I.2.1 Davidon-Fletcher-Powell (DFP) method

Let B be the matrix that approximates the inverse of the Hessian matrix.

Initially, a point x_0 is chosen and a symmetrical matrix $B_0 > 0$ (usually $B_0 = I$).

Then, $\nabla f(x^i)$ and optimal step θ_i in the travel direction $h^i = -B_i \nabla f(x^i)$ are computed.

It is calculated $x^{i+1} = x^i + \theta_i h^i$.

Check if the stop criterion is met.

If it is *yes*, then *stop* the algorithm,

otherwise, $B_{i+1} = B_i + M_i + N_i$ is computed.

It follows the next iteration.

I.2.2 Broyhen-Fletcher-Goldfarb-Shanno Method (BFGS)

Initially, a point x_0 is chosen and a symmetrical matrix $B_0 > 0$ (usually, $B_0 = I$). Following this, $\nabla f\left(x^i\right)$ and optimal step θ_i on the travel direction $h^i = -B_i \nabla f\left(x^i\right)$ are computed.

It is calculated $x^{i+1} = x^i + \theta_i h^i$.

Check if the stop criterion is met, and if it *yes*, *stop* the algorithm. An update of the Hessian matrix value is delivered.

Continue with the next iteration.

Comparatively, the BFGS algorithm is less affected by the errors in the optimal step calculation than the DFP algorithm. As regards the numerical stability, BFGS is considered the most stable method. From the ones presented so far, it can be observed that quasi-Newton methods only require first order information (only gradient type calculations are used).

The NM is a good choice for the multi-variable equation systems solving, and to find the polynomials complex roots. Newton wanted to determine the solution of an "algebraic" given problem: F(x, y) = 0, in which the variable *y* is expressed as a series of powers in x.

III. Newton-Raphson methods (NRM)

The NRM supposes the determination of the second order derivatives of the objective function. The NRM has the main disadvantage the requirement of a large amount calculus, but this is compensated by the speed of convergence. The Raphson procedure is equivalent to the linear approximation. The first analysis of the NM convergence had been done in 1820 by the Cauchy, and Fourier [3, 7, 9].

III.1. Newton-Raphson method

The method is based on the linear approximation of the objective function by the Taylor series expansion. The general recurrence relationship is:

$$x^{i+1} = x^i - H_i^{-1} \nabla f\left(x^i\right)$$

where

H_i-the Hessian matrix of the objective function at iteration i

There are some disadvantages of the recurrence formula in this form: the method can converge to a saddle point or to an extremely relative point. Thus, the recurrence relation can be modified, eliminating the mentioned problems by moving at the optimum pace on the direction of movement. The relationship becomes:

$$x^{i+1} = x^i - \theta_i^* H_i^{-1} \nabla f\left(x^i\right)$$

For quadratic functions, the extremum point is reached in one step. For the other functions, the extreme point is not found in a single step, but convergence is fast.

Even with this new recurrence relation, the method cannot be used in the case of objective functions of many variables, since:

- nxn dimensional matrix is stored at each iteration
- Each iteration is calculated $H_i^{-1} \nabla f\left(x^i\right)$
- In certain situations it is impossible to calculate the elements of the Hessian matrix
- At each iteration, the inverse of the Hessian matrix is calculated.

III.2 The Marquardt method

It combines the advantages of the optimal gradient method (safe convergence and rapid decrease of the function value for the case of choosing the starting point far from the minimum point) and the Newton-Raphson method (rapid convergence if the starting point is near to the minimum point).

IV. Trust region method (TR)

In this method, based on the information gathered about the objective function f an objective function T_i is so built, that near the point x_i, it behaves just like function f. Since objective T_i is not always the best approximation of f we have to force the search for a minimizer of T_i to a certain area around x_i. In this regard, the step s is search such that the next subproblem will be solved by approximation:

$$\min_s \quad T_i(x_i + s),$$

in which $(x_i + s)$ is placed in the TR.

If the solution obtained does not lead to a significant decrease in f, it means that the TR is overgrown; therefore we reduce the step and proceed again to solve the above mentioned subproblem.

The procedure of applying TR method consists of:

1. Maximum distance Δ_i, i.e. the *radius* of the TR, is chosen;
2. Search for a direction d with a step s such as the maximum decreasing rate is obtained;

3. If the obtained result is unsatisfactory, proceed to the choice of a smaller Δ_i, and repeat the algorithm from *step* 1.

By considering the first two terms of the quadratic model of the function T_i, and the first two terms of the f Taylor's series expansion around x_i that are identical for each iteration x_i, the following relationship can be deducted:

$$T_i(s) = f_i + s^T \nabla f_i + \frac{1}{2} s^T B_k s,$$

with symmetric matrix B_k.

At the same time, by applying Taylor's theorem for a continuous and differentiable function f ,the approximation function around s point can be written:

$$f(x_i + s) = f_i + s^T \nabla f_i + \frac{1}{2} s^T \nabla^2 f(x_i + ts)s,$$

with $t \in (0, 1)$, a scalar number.

In this way, the following approximation T_i is adopted:

$$T_i(s) = f_i + s^T \nabla f_i + O(\|s\|^2).$$

The difference between T_i and $f(x_i + s)$ it is just $O(\|s\|^2)$. The approximation error is small if s is small.

As a consequence, the following problem should be solved: norm

$$\min T_i(s) = f_i + s^T \nabla f_i + \frac{1}{2} s^T B_k s, \text{ with norm } \|s\| \leq \Delta_i, \tag{9}$$

$\Delta_i > 0$ being the trust-region radius.

Three strategies are described for identification of approximate solution, based on which at least one reduction of T_i is obtained, as well as the reduction obtained using the Cauchy point. This point is actually a T_i minimizer along the direction that offers the steepest descent—∇f_i within the TR.

V. The Cauchy Point

In the line search methods, the approximation is sometimes coarse from an optimal length of the step. However, this does not affect global convergence. In order to simplify the calculations, the same methods can be used in trust region methods. In other words, instead of finding an optimal solution to (8.93), to identify an approximate solution s_i is a better approach. Therefore, to achieve global convergence, this approximate solution should be placed within the trust region, and to provide a satisfactory reduction of the objective function.

The satisfactory reduction is obtained using the Cauchy point p_i^C which can be defined by using the following algorithm.

Algorithm:

– Looking for a vector p_i^S which is a linear solution of the (8.93) problem

$$p_i^S = \arg \min_{s \in R^n} f_i + s^T \nabla f_i, \quad \|s\| \leq \Delta_i$$

– The $\tau_i > 0$ scalar is determined such that $T_i(\tau p_i^S)$ is minimized, i.e.

$$\tau_i = \arg \min_{\tau > 0} T_i(\tau p_i^S), \text{ for } \left\|\tau p_i^S\right\| \leq \Delta_i;$$

– Set

$$p_i^C = \tau_i p_i^S,$$

results:

$$p_i^C = -\tau_i \frac{\Delta_i}{\|\nabla f_i\|} \nabla f_i,$$

in which

$$\tau_i = 1, \text{for} \nabla f_i^T B_i \nabla f_i \leq 0,$$

otherwise

$$\tau_{\mathrm{i}} = \min(\|\nabla f_{\mathrm{i}}\|^3 / (\Delta_i \nabla f_i^T B_i \nabla f_i), 1).$$

Always taking the Cauchy Point as the step of the method gives the abrupt descent method. For example, this is not the case for *steepest descent method* even if an optimal step size is chosen at each iteration.

Because the matrix B_i is used only to calculate the step length, the Cauchy point has a lower dependence on it.

A fast (superlinear) convergence is obtained if B_i is exactly Hessian $\nabla^2 f(x_i)$ or a quasi-Newton approximation. The matrix B_i is used for both the direction calculation, and the step length.

The *dogleg method* is one way to find the solution to the problem (9). Another ways, are to use *Steihaug method,* or *two-dimensional subspace* minimization.

In order to increase the clarity, the problem (9) can be simplified by considering the following single iteration problem:

$$\min_{\mathrm{s} \in \mathrm{R}^{\mathrm{n}}} T(s) \stackrel{def}{=} f + q^T + \frac{1}{2} s^T B s, \quad \|s\| \leq \Delta \tag{10}$$

having the solution $s^*(\Delta)$.

VI: Dogleg method

The *Dogleg* method is used to find a two line segments approximate solution instead of a curve path for $s^*(\Delta)$. The origin is the initial value of the first path segment to the unconstrained minimizer with the steepest descent direction defined by:

$$s^U = -\frac{q^T q}{q^T B q} q.$$

The second path segment starts from the point s^U to the point s^B having a trajectory defined by $d(\tau)$:

$$d(\tau) = \begin{cases} -\tau s^U, & 0 \le \tau \le 1 \\ s^U + (\tau - 1)(s^B - s^U), & 1 \le \tau \le 2 \end{cases}$$

with $\tau \in [0, 2]$.

If the norm $\|s^B\| \geqslant \Delta$, then the path $d(\tau)$ crosses the boundary of the TR, $\|s\| = \Delta$, at one point (Fig. 1).

VII. Two-dimensional subspace minimization

By extending the search to the entire surface covered by s^U and s^B(*two-dimensional subspace*), a simplification of the *dogleg* method is obtained.

In these conditions, when matrix $B > 0$, has the positive eigenvalues, the sub-problem (10) becomes:

$$\min_s T(s) \stackrel{def}{=} f + q^T s + \frac{1}{2} s^T B s, \ \|s\| \le \Delta, \ s \in span[q, B^{-1}q], \qquad (11)$$

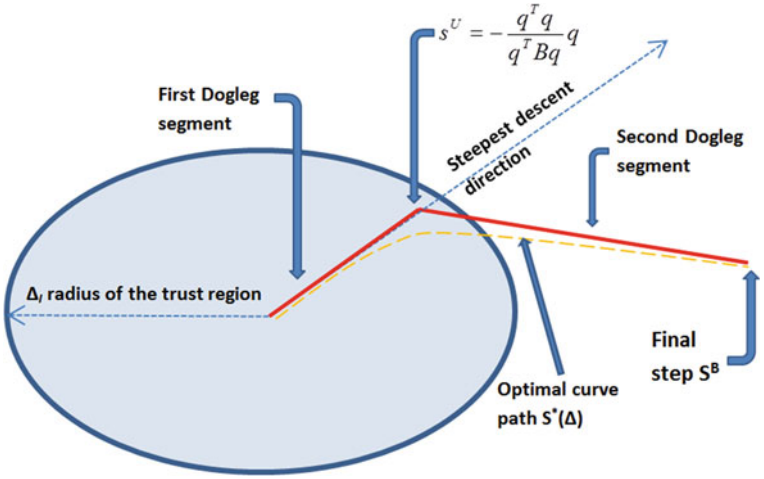


Fig. 1 Illustration of the Dogleg method

in which the span operator means the possible linear combinations set between vectors q and $B^{-1}q$.

The problem (11) can be solved relatively easy. The optimal solution to this subproblem: p_i^C—the Cauchy point.

In case of $B < 0$, the sub-problem (10) is changed to:

$$span[q, (B + kI)^{-1}q], \text{for } k \in (-\lambda_1, -2\lambda_1], \tag{12}$$

where λ_1 indicates the most negative value of the eigenvalue of the matrix B.

The value of k is chosen such that $B + kI$ is positively defined.

In case of $\left\|(B + kI)^{-1}q\right\| \leq \Delta$, the new step is defined as:

$$s = -(B + kI)^{-1}q + a$$

with value of a chosen such that the condition $a^T(B + kI)^{-1}q \leq 0$ is satisfied.

VIII. Steihaug's approach

The above presented two methods (*dogleg, two-dimensional subspace minimization*) assume a linear system whose solutions depend on B or $(B + aI)$. But solving this system when B is big becomes costly. Therefore, in practical situations other techniques must be identified to find a solution to the subproblem (12) that does not depend on the exact solution of the linear system, but which nevertheless leads to an Cauchy point improvement.

In this regard, Steihaug proposed a method similar to the *conjugate gradient algorithm* (as it is previously described).

The main difference between the standard conjugate gradient algorithm and Steihaug's approach is that the algorithm is finalized when either the limits set by the trust region $\|s\| \leq \Delta$ are exceeded or when a negative curvature direction is reached in B.

In the *Newton trust region method*, the B matrix is chosen to be exactly Hessian $\nabla^2 f(x)$ (or approximations thereof). These methods have very good local and global convergence properties.

Unlike the line search methods where a fixed direction d_i and at each iteration a length of step α_i, is determined, in the trust region, firstly a radius of this region is chosen in the form of a maximum distance Δ_i, and, secondly, to both *direction* and *step* are explored such that the best improvement is possible. If the result is not satisfactory, the distance Δ_i is reduced, and a new minimizer should be determined. In general, as the radius of the trust region changes, the direction of the step changes. Very important at each stage is the magnitude of the trust region. Often, the size of the trust region is based on the previous obtained results (if the result is satisfactory) from the iteration algorithm. If the region is not big enough, the algorithm passes a good opportunity to take an increased step that will position the results much closer to the objective function minimizer. If it is too large, the model minimizer may be far

from the objective function minimizer in the region and the size of the TR is reduced (repeat the algorithm).

IX. ***Generic TR Method***

In order to establish the border of the TR, a maximum radius should be chosen $\Delta_{\max} > 0$.

Similiarly, the distance $\Delta_0 \in (0, \Delta_{\max})$, and a fixed η within (0, 1/4) interval, the starting point $x_0 \in R_i$ within the TR R_i, B_0, $\varepsilon > 0$ should be initiated.

While $\|\nabla f(x_i)\|$ norm is larger or equal than ε, do

Compute y_{i+1} as the approximate minimiser of $x_{i+1} \approx \arg\min_{x \in R_i} T_i(x)$;

Determine x_{i+1}:

$$x_{i+1} = y_{i+1}, \text{ if } \frac{f(x_i) - f(y_{i+1})}{T_i(x_i) - T_i(y_{i+1})} > \eta,$$

$x_{i+1} = x_i$ otherwise;

Compute Δ_{i+1}:

$$\Delta_{i+1} = \frac{\Delta_i}{4}, \text{ if } \frac{f(x_i) - f(y_{i+1})}{T_{ki}(x_i) - T_i(y_{i+1})} < \frac{1}{4},$$

$$\Delta_{i+1} = \min(2\Delta_i, \Delta_{\max}), \text{ if } \frac{f(x_i) - f(y_{i+1})}{T_i(x_i) - T_i(y_{i+1})} > \frac{3}{4};$$

$\Delta_{i+1} = \Delta_i$, otherwise;

Build a new model function $T_{i+1}(x)$.

$$i \leftarrow i + 1$$

end.

3.2. Direct Search Methods

These methods are also used in the case of non-divisible functions for determining the values of the objective function. As disadvantages, convergence is slow, based on simple calculations at each iteration.

Direct search methods can be—*iterative methods*—points are getting closer and closer to the minimum.

As algorithm methods can be mentioned:

- Gauss-Seidel,
- Nelder-Mead
- downhill simplex, or amoeba method,
- Rosenbrock and Powell,
- SIMPLEX and COMPLEX.

3.2.1. Nelder-Mead [10]

This optimization method takes part from the direct search optimization methods.

By considering a quadratic function defined by the user as:

```
function y = fnm (x)
  y = 0
for k = 1 : size (x, 2) - 1
y = y + 68 * (x (k + 1) - x (k)) ^ 2 + (1 - x (ik)) ^ 2 - x (k) * x (k + 1);
end
endfunction
```

the problem is to find the optimal value through the Nelder-Mead method.

The optimal searching method is based on construction of new solutions by using a simplex structure and a transforming series. This method does not require the deduction of the gradients.

The detailed optimal structure solution can be found as:

opt = optimset ("Display", "iter");

The structured solution is composed by three elements: number of iteration, the function count, and the minimum value of the function at each iteration. The method can be used to improve the evolution of the optimization process.

Results: the minimum value of the quadratique function is found after 12 iterations as in the following structure (Fig. 2).

Iteration	Func-count	min f (x)
0	3	0
1	3	0
2	7	0
3	11	0
4	15	0
5	19	0
6	23	0
7	27	0
8	31	0
9	35	0
10	39	0
11	43	0
12	47	0

3.2.2. Rosenbrock Method [10]

The problem is to find the optimum value of the nonlinear function of two variables, $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$:

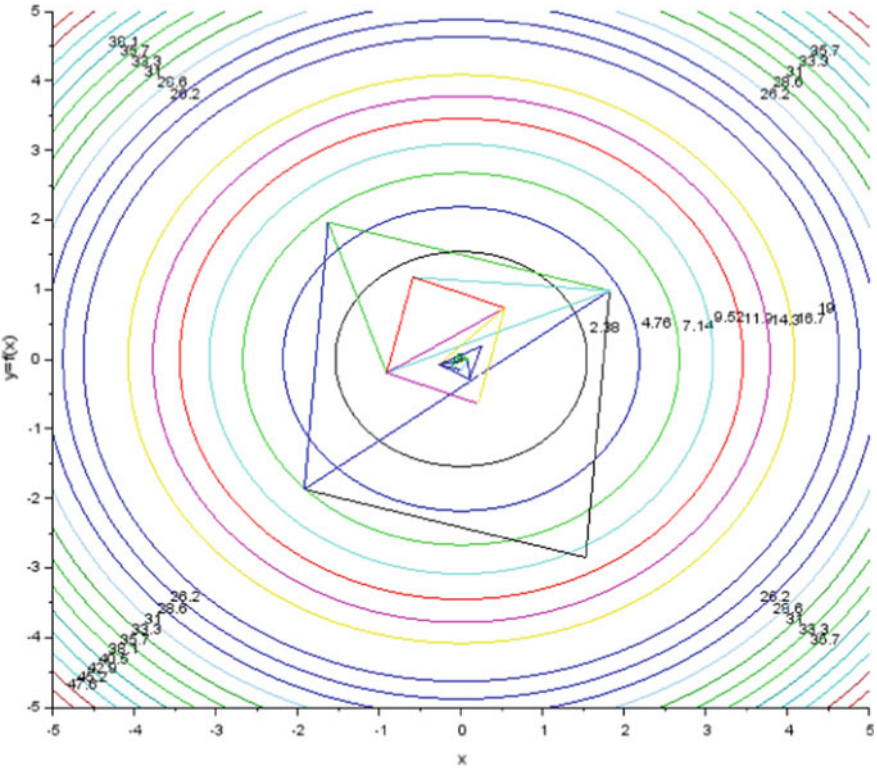


Fig. 2 Nelder Mead method for solving a quadratique function [10]

$$f(x_1, x_2) = (2 - x_1)^2 + 68\left(x_2 - x_1^2\right)^2$$

near to [−1.2 1.0].

The multivariable function mentioned above is known is Rosenbrock function.

This optimization problem takes part of the nonlinear optimization without constraints.

The optimum value is find by using the three basic methods:

(i) Quasi-Newton BFGS,
(ii) Quasi-Newton BFGS with limited memory;
(iii) QN for non-diferentiable objective function (local optimization without using the derivatives).

(i) Quasi-Newton BFGS method is a local optimization based method, and uses the specification of the first derivative respect to each variable.

The software implementation of the above mentioned optimization problem supposes:

– The definition of the index function:
  f = 60 *(x(2) – x(1)^2)^2 + (2 – x(1))^2;
– The first order derivative respect to $x_1(1)$ = D1(1):
  D1(1) = –272. * (x(2) – x(1)^2) * x(1) -2. * (2. –x(1));
– The first order derivative respect to $x_1(2)$ = D1(2):
  D1(2) = 136. * (x(2) – x(1)^2).

After definition of the cost function and its first derivatives, the starting point of the optimization algorithm: in this example the initial point is placed at [−1.2 1.0] blue mark in Fig. 3a. The initial estimate of the solution coordinates should be also be provided: the initial guess is placed at [1.0 1.0] (red point on the Fig. 3a).

Contours of the particular Rosenbrock function of two variables (a = 2, b = 68) are plotted on the Fig. 3a, the optimization process is depicted on the Fig. 3b, and by using the QN-BFGS algorithm the minimum value of the objective function is found at $[x_{1min},x_{2min}] = [2, 4]$, having the scalar value $f_{min}(x_1, x_2) = 0$.

The optimum function value fopt, for the initial guess × 0 can be find as:

fopt1 = 0

placed at the optimum coordinates $[x_{1min},x_{2min}]$ = xopt1

xopt1 = [2. 4.]

(ii) By using the quasi-Newton method based on Broyden-Fletcher-Goldfarb-Shanno (BFGS) with limited memory, the same values are found. This method is applied to large number of variables (more than 100):

fopt2 = 0

at the $[x_{1min},x_{2min}]$ = xopt2

xopt2 = [2. 4.].

(iii) By using the third method, for non-differentiable functions, an approximation of the optimal solution is obtained as:

fopt3 = 0.0004828,

the optimum value of the nonlinear multivariable function being placed at,

xopt3 = [1.9866225 3.9487826].

In order to find the optimal value of the cost function,

The optimization based on the Nelder-Mead search method does not use the information about the gradient of the function. The function value evolution during the unconstrained optimization process can be plotted for the above known function.

In Fig. 3c the dynamic of the optimization process is shown As can be seen, the last half of the iterations maintain the cost function value is almost maintained at the constant value. The optimal value of the cost function is 1.389381e-10 (Fig. 3c).

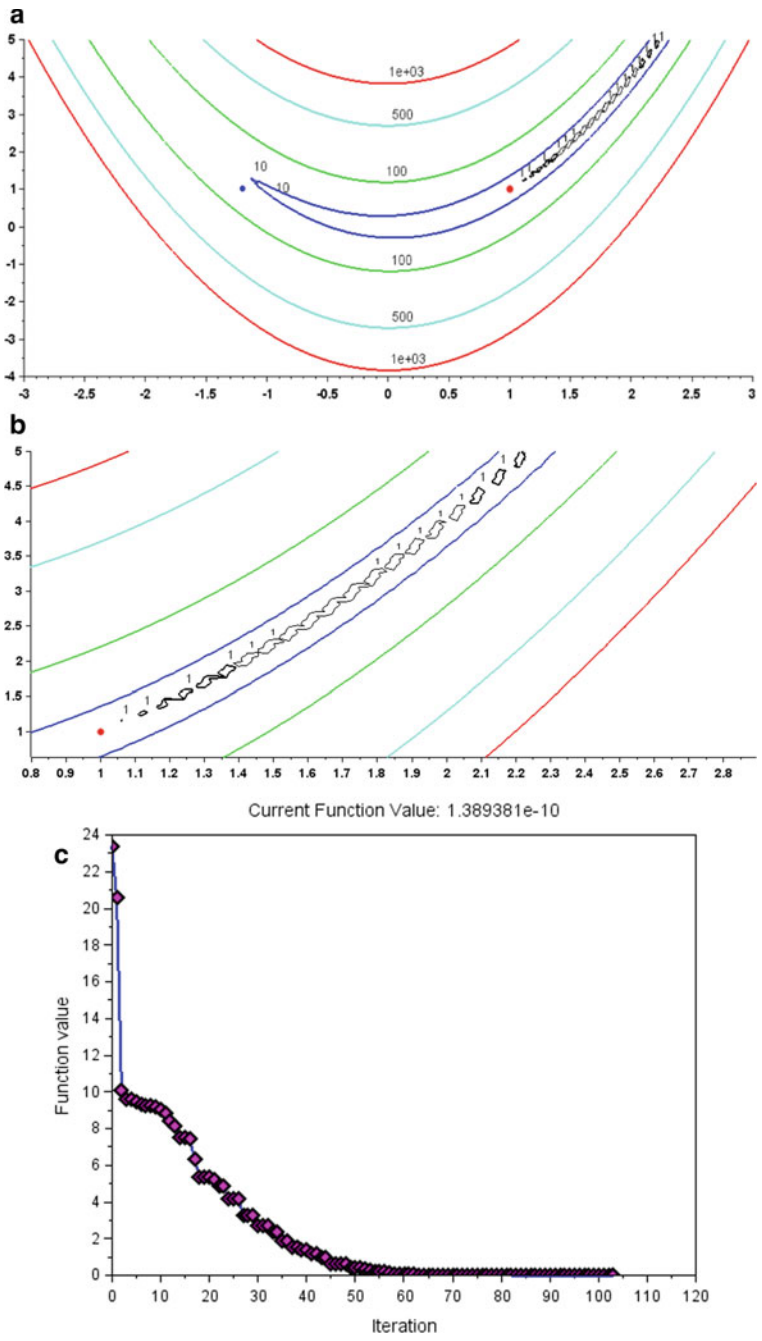


Fig. 3 **a** Contours of the particular Rosenbrock function of two variables (a = 2, b = 68). **b** Plot the initial guess of the optimum variables $[x_{1min},x_{2min}] = [1]$ and the evolution of the optimization process. **c** Cost value: the evolution of the minimization process through the used Nelder-Mead algorithm

The dynamic of the transformations at each iterate, as well as the applied optimization procedure at each step can be viewed:

Iteration	Func-count	min f(x)	Procedure
0	3	23.4048	
1	3	20.5828	initial simplex
2	5	10.054469	expand
3	7	9.6029493	reflect
4	9	9.6029493	contract outside
5	11	9.4738768	contract inside
6	13	9.31917	contract inside
7	15	9.2507973	reflect
8	17	9.2507973	contract inside
9	19	9.1980872	expand
10	21	9.0477781	expand
11	23	8.8547479	expand
12	25	8.3861285	expand
13	27	8.1417887	expand
14	29	7.4969757	reflect
15	31	7.4969757	contract inside
16	33	7.4288297	expand
17	35	6.3523046	expand
18	37	5.3708076	expand
19	39	5.3708076	contract inside
20	40	5.3708076	reflect
21	42	5.2331712	reflect
22	44	4.8457422	reflect
23	45	4.8457422	reflect
24	47	4.1618679	reflect
25	49	4.1618679	contract inside
26	51	4.1442689	reflect
27	53	3.2469009	expand
28	55	3.2469009	contract inside
29	57	3.2469009	contract outside
30	59	2.6898366	expand
31	61	2.6898366	contract outside
32	63	2.6898366	contract outside
33	65	2.334585	expand
34	66	2.334585	reflect
35	68	1.8784037	expand

(continued)

(continued)

Iteration	Func-count	min f(x)	Procedure
36	70	1.8784037	contract inside
37	72	1.5530071	expand
38	74	1.5530071	contract outside
39	76	1.3613402	reflect
40	77	1.3613402	reflect
41	79	1.1890049	reflect
42	81	1.1890049	contract inside
43	83	0.9850215	expand
44	84	0.9850215	reflect
45	86	0.6400976	expand
46	88	0.6400976	contract inside
47	90	0.6400976	contract outside
48	92	0.6241326	reflect
49	94	0.435756	expand
50	96	0.435756	contract inside
51	98	0.435756	contract outside
52	100	0.3462121	expand
53	102	0.2962515	reflect
54	104	0.2115731	reflect
55	106	0.2115731	contract inside
56	108	0.1905142	expand
57	110	0.1339536	reflect
58	112	0.0576853	expand
59	114	0.0576853	contract inside
60	115	0.0576853	reflect
61	117	0.0548968	reflect
62	119	0.043614	contract inside
63	121	0.026894	reflect
64	123	0.026894	contract outside
65	125	0.0176597	contract outside
66	127	0.0176597	contract inside
67	129	0.0158285	contract outside
68	130	0.0158285	reflect
69	132	0.0134937	reflect
70	134	0.0133218	contract inside
71	136	0.0114001	reflect
72	138	0.0114001	contract inside

(continued)

(continued)

Iteration	Func-count	min f(x)	Procedure
73	140	0.0094043	expand
74	142	0.0073118	expand
75	144	0.0034168	expand
76	145	0.0034168	reflect
77	147	0.0003337	expand
78	148	0.0003337	reflect
79	149	0.0003337	reflect
80	151	0.0003337	contract inside
81	153	0.0003007	contract outside
82	155	0.0000446	contract inside
83	157	0.0000446	contract inside
84	159	0.0000178	contract outside
85	161	0.0000071	contract inside
86	163	0.0000058	contract inside
87	165	0.0000028	contract inside
88	167	0.0000006	contract inside
89	169	0.0000006	contract inside
90	171	0.0000005	contract inside
91	173	7.754D-08	contract inside
92	175	7.754D-08	contract inside
93	177	6.537D-08	contract outside
94	179	7.641D-09	contract inside
95	181	7.641D-09	contract inside
96	183	7.641D-09	contract outside
97	185	1.573D-09	contract inside
98	187	1.573D-09	contract inside
99	189	1.325D-09	contract inside
100	191	8.792D-10	reflect
101	193	1.389D-10	contract inside
102	195	1.389D-10	contract inside
103	197	1.389D-10	contract inside

The procedure *contract inside* means that the problem has a solution, and the termination criteria has been attained. In order to check the optimization progress, at each iterate there is a warning message regarding the convergence evolution of the algorithm. If the problem converges to the optimal solution, the following message could be viewed: Optimization terminated.

The variables **x** satisfies the termination criteria using OPTIONS.TolX of 0.0001 (tolerance) and cost function F(x) satisfies the convergence criteria using OPTIONS.TolFun of 0.0001.

By using Nelder-Mead method, the optimal variables pair is obtained,

xopt4 = [2.0000022 4.0000102],

conducting to an optimal cost function of:

fopt4 = 1.389D-10.

6 Nonlinear Optimization Without Constraints

The canonical form of *nonlinear programming* problems is [7] to find the optimal solution vector $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ such that the index function $z = f(x_1, x_2, \ldots, x_n)$ is minimized, having the $g_i(x_1, x_2, \ldots, x_n) \leq 0$, as constraints, with $i = 1, 2, \ldots, m$, and nonnegative conditions $x_1 \geq 0$, $x_2 \geq 0$, $\ldots$ $x_n \geq 0$.

The canonical form can be synthesize as:

$$\begin{cases} \min f(\underline{\mathbf{x}}), & \underline{\mathbf{x}} \in R^n; \\ g_i(\underline{\mathbf{x}}) \leq 0, & i = \overline{1, m}; \\ \underline{\mathbf{x}} \geq 0 \end{cases}$$

In case of the *linear programming* problem, both the *index function*, *z*, and the *constraints*, g_i, are *linear.* Therefore, there is (at least) a general method of solving—for example the simplex method—in the nonlinear case there is no such method. However, substantial progress has been made in some special cases by imposing conditions on the functions f and g_i.

The non-linearity of the objective or some of the constraints leads to *complicating the* task of determining the optimum.

(1) From the beginning we will emphasize that in *nonlinear programming*—with a few exceptions—the methods of solving "theoretically" obtain the optimal solution as the *limit* of a series of solutions. Thus, a *concrete* process *of nonlinear optimization* is *finished* not due to the structure of the problem but by *the user's will* that limits the number of steps according to a whole series of factors such as: complexity of calculation, time available, performance of computing equipment, etc.

(2) it is possible that the objective function in (P) has more *local minima* on the set of admissible solutions A.

The possibility of the existence of several local minima of the objective function represents a serious difficulty in solving a nonlinear program. Indeed, in the formulation itself, such a problem requires the determination of the global minimum of the objective. However, all known non-linear optimization methods fail to determine at

most a local minimum, with no guarantee that it coincides with the global minimum sought.

As we will see, if *A* is convex and the objective function is convex and minimized then it has at most a local minimum which - if any - is automatically global.

(3) Even though the constraints in (P) are *linear* but the objective remains *nonlinear* but *convex*, the optimal solution, although it is on the border of A, is not necessarily a *peak.*
(4) it is possible that the optimal solution is located inside the set A.

Classes of nonlinear problems

6.1 Unconstraint *Optimization Problems*

Unconstraint optimization problems have the general form:
determine $x^* \in R^{n\ n}$ which minimizes the value of the function

$$z = f(x_1, x_2, \ldots, x_n),$$

the minimum being taken after all $x \in R^{n\ n}$ where function f is defined.

6.2 Optimization Problems with Linear Constraints *and* Nonlinear *Objective Function*

In this class a special attention should be on the *quadratic programming* problems. The index function is a *polynomial* of the *second* degree in its variables:

$$f(x_1, x_2, \ldots, x_n) = c_0 + \sum_{i=\overline{1,n}} c_i x_i + \sum_{(i<j)=\overline{1,n}} c_i x_i x_j$$

The problems of quadratic programming are important for the following reasons:

- the fact that it models many practical situations with sufficient accuracy;
- it is solved by methods derived from the simplex method in a finite number of steps;
- solving many problems with linear constraints and nonlinear objective function can be reduced to solving a sequence of quadratic programming problems whose objective functions increasingly approximate the original nonlinear objective.

7 The Problems of Convex Programming

The problems of convex programming are characterized by:

- *convex* objective (index) *function* if it is *minimized* (equivalent: *concave* objective function if it is *maximized*);
- $g_i(x) \leq 0$, the *inequality* constraints are of the form $g_i(x) \leq 0$ in case of *convex* function g (equivalent $g_i(x) \geq 0$, in case of a *concave* function g);
- *equality* constraints are *linear*, a requirement motivated by the fact that linear functions are the only simultaneously convex and concave functions.

Convex problems have the following fundamental properties:

- the set of admissible solutions is *convex*;
- the objective function admits at most an optimal (minimum or maximum) *local*; automatically this will be a *global* optimum and will represent the optimal of the problem;
- if the *free* (*unconstrained*) optimal of the objective function is not an admissible solution then the *constrained* optimal is necessarily located on the *boundary of* set A.

The importance of this class of problems is *very high* due to the following reasons: Convex programming includes linear programming;

- In this area, the greatest research effort was made and the strongest *theoretical* results (such as nonlinear duality theory, Kuhn - Tucker optimality conditions) and *practical* (optimization methods and algorithms) were obtained;
- The whole mathematical formalism of modern economic theory is based on convexity assumptions.

8 The Problems of Separable Programming

The problems of separable programming are characterized by the fact that the index function f as well as the constraints g_I are *separable* within the meaning of the following definition:

The separable function $f(x_1, x_2, \ldots, x_n)$ should satisfy the following relation:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i).$$

Separability is important because it facilitates optimization. For example, the optimization of a separable function without constraints is reduced to the independent optimization of terms.

9 The Problems of Non-convex Programming

The problems of non-convex programming bring together all the problems that do not satisfy the convexity hypotheses. These problems have several local minima. The current methods can cause such an optimum. However, these problems cannot guarantee that the solution is the global optimum. Fortunately, there are several types of non-convex problems, useful in practice that can be solved without special difficulties by special methods, added via *fractional* programming problems.

Sets and convex functions

By considering C as a set of the real *n*- dimensional space ($C \subseteq R^n$), the set is *convex* if for two containing points the segment joining them is inside of the C.

Mathematically,

C is a convex set if $(\forall) \quad x, y \in C,$ for any $\lambda \in [0, 1],$ the following relation is true $(1 - \lambda)x + \lambda y \in C$.

Supposing C is a convex set, and f a numerical function defined at all points of the set C, f is a *convex* function if:

$(\forall) \quad x, y \in C,$ for any $\lambda \in [0, 1],$ the following relation is true

$$f[(1 - \lambda)x + \lambda y] - (1 - \lambda) f(x) + \lambda f(y) \leq 0$$

Oppositely, the concave function should satisfy the folowing:

$$f[(1 - \lambda)x + \lambda y] - (1 - \lambda) f(x) + \lambda f(y) \geq 0.$$

For any different two points, x, y, with $(\forall) x, y \in C$, and $(\forall) \lambda \in (0, 1)$, the function f is named *strictly convex* (*strictly concave*) if the above mentioned relationships are strictely.

10 Nonlinear Optimization Without Constraints. The Convex Case

The general optimization problem:

(P) Find $x^* \in A \subseteq R^n$ with the property that the function value $f(x^*)$ satisfy the relation inf $\{f(x), x \in A\}$,

for any x from the admissible solutions (A) set of the problem (P), A is defined by a set of constraints:

$$gi(x) \leq 0, i \in M = \{1, 2, \ldots, m\}.$$

In order to simplify the explanation, any conditions of non-negativity $x_j \geq 0$ were included in the restriction block in the following form: $-x_j \leq 0$.

Suppose the functions f and $g_1, g_2 \ldots, g_m$ are defined entire the space R^n, ^{n} are convex and differentiable, and at least one of them is non-linear. In this way, the problem (P) is a *programming convex problem.*

Recall that in this context:

- A is a **closed** and **convex set**;
- any **local minimum** of function f on the set A it is a **global minimum**.

11 Optimization Methods with Constraints

Techniques for nonlinear software problems with constraints are superior to those without constraints.

If the constraints are linear, the algorithms are based on methods of optimizing the case without constraints. The basic approach for solving a multivariate constrained nonlinear problem is to reformulate it into a succession of related problems, so that each problem can be solved by simpler methods. As an example, Newton's methods for unconstrained problems based on a local quadratic approximation of the objective function can be developed for the unconstrained problem by restricting the region in which the quadratic model is valid. This constraint is, in principle, similar to the *trust region* method.

These optimization techniques are splitted in three categories:

(1) *methods of gradient*—based on the adaptation of the unconstrained general optimization scheme in case of constraints.
(2) methods based on *penalty functions*: solving the main problem reduces to more unconstrained optimized problems.
(3) methods based on the *plane sections*; its principle, these methods "approximate" by polyhedral set (a set that can be described by a system of linear inequalities); solving the problem is reduced to a sequence (infinite) of linear optimization problems made using the simplex algorithm.

11.1 Convex Programming. Kuhn—Tucker Optimality Conditions

Formulation of the conditions

It is considered the canonical form of a **convex programming:**

(P) Find the $\underline{\mathbf{x}}^* \in R^n$ ^{n} with the property that the value of the function f $(\mathbf{x}^*)$ is minimal:

$$f(x^*) = \mathbf{min} f(x).$$

The minimum value is determined for any $\underline{\mathbf{x}} \in R^n$ by fulfilling the constraints:

$$q_i(\underline{x}) \leq 0, i = \overline{1,m}$$

and the non negativity conditions:

$$\underline{\mathbf{x}} \geq 0, x_j \geq 0 \; j = 1, \ldots, n.$$

It is assumed that the functions f is defined in the space R^n, and q_i ($q_1, q_2, \ldots, qm$) is also defined. Morevover, both of the functions are differentiable.

For each restriction $q_i(\underline{\mathbf{x}}) \leq 0$ corresponds to a variable single nonnegative single u_i. Thus, the **Lagrangian of the** problem is constructed based on:

$$L(\underline{\mathbf{x}},\underline{\mathbf{u}}) = f(\underline{\mathbf{x}}) + \sum_{i=\overline{1,m}} u_i q_i(\underline{\mathbf{x}}),$$

where $\underline{\mathbf{u}} = \overline{u_1, u_m}$ are called **Lagrange multipliers components**.

If $\underline{\mathbf{u}} \geq 0$ ($\underline{\mathbf{u}} \in \mathbb{R}^m$), then L is a convex function and differentiable.

Assuming that the conditions of regularity Slater [11], the set of the admissible solutions is **within the relatively non-empty, the Kuhn - Tucker theorem is as follows:** the necessary and suffiecient conditions such that $\underline{\mathbf{x}}^* \in R^n$ be the minimum value of the problem (P) is to exist $\underline{\mathbf{u}}^* \in R^m$ such that the pair ($\underline{\mathbf{x}}^*$,$\underline{\mathbf{u}}^*$) check the **Kuhn - Tucker optimality** relationship:

$\frac{\partial L}{\partial x_j} \geq 0, \quad j = \overline{1,n}; x_j \cdot \frac{\partial L}{\partial x_j} = 0$, respect with $x_j \geq 0$,

$\frac{\partial L}{\partial u_i} \leq 0, i = \overline{1,m}; u_i \cdot \frac{\partial L}{\partial u_i} = 0$, respect with $u_i \geq 0$.

Constrained optimization. Sufficient conditions for minimum point

Taken into consideration two continuously differentiable real-valued functions f, $g_1, g_2, \ldots, g_m$, if there are vectors $\mathbf{x}_0 \in R^n$, $\lambda_0 \in R^m$, such that:

$$\nabla L(\mathbf{x_0}, \lambda_0) = 0$$

and the positivity condition is respected:

$$(-1)^m \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2}(\mathbf{x}_0, \lambda_0) & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_p}(\mathbf{x}_0, \lambda_0) & \frac{\partial q_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial q_m}{\partial x_1}(\mathbf{x}_0) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 L}{\partial x_p \partial x_1}(\mathbf{x}_0, \lambda_0) & \cdots & \frac{\partial^2 L}{\partial x_p^2}(\mathbf{x}_0, \lambda_0) & \frac{\partial q_1}{\partial x_p}(\mathbf{x}_0) & \cdots & \frac{\partial q_m}{\partial x_p}(\mathbf{x}_0) \\ \frac{\partial q_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial q_1}{\partial x_p}(\mathbf{x}_0) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial q_m}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial q_m}{\partial x_p}(\mathbf{x}_0) & 0 & \cdots & 0 \end{bmatrix} > 0$$

for $p = m + 1, \ldots, n$, then the strictly local minimum of the function f is the *vector* $\mathbf{x_0}$

$$q_i(\mathbf{x}_0) = 0, i = \overline{1,m};$$

if $p = n$, the *bordered Hessian matrix* is obtained from the above last mentioned matrix.

11.2 Semi-defined Programming (SDP)

Semi-infinite programming problems ***(SIPs)*** **or SDPs** are components of the optimization family having an infinite number of variables or of constraints.

A SDP problem is *linear* if two conditions are fulfilled: (1) there is a linear *objective function f*; (2) there are affines constraint functions $g(t)$, for any $t \in T$.

A SDP problem is *convex* if the following two conditions are respected: (1) the objective function f is convex, (2) the constraint g (., t) is concave for all $t \in T$. In this case the considered set is *feasible*.

The function F is convex

$$F = \{\underline{\mathbf{x}} \in \mathrm{R}^{\mathrm{n}} : \mathbf{g}(\underline{\mathbf{x}}, t) \geq 0 \,\text{for all}\, t \in T\}.$$

In this case, at the same time the local and the global minimum are the same.

Semi-defined programming involves *second order cone programming* (SOCP), because SOCP constraints can be written as linear matrix inequalities.

SDP is a generalization of linear programming (LP). SDP is reduced to LP problem only in case of the diagonal matrices [8].

12 Mixed Integer Nonlinear Programming (MINLP)

The variables in the most used optimization problems are integer or discrete form. These can be modeled as MINLPs. The modeling variables can be integers (number of buildings), or binary (0 or 1) type (decision modelling). In addition, there may be continuous variables that may represent, for example, speed or torques. The nonlinearities can be found in the mathematical model of the studied process (the hysteresis of the magnetic materials), or in the decision variables. The objective function of the MILPs can be the costs minimization or the profits maximization.

13 Heuristic Methods

Heuristic Methods (HM) are faster than traditionally optimization algorithms

Classification of the metaheuristic algorithms (MA) for local and global search:

(a) The general structure of the local optimization algorithms

(b) Deterministic for local search. There are two mainly types of searching algorithms: Pattern, and Nelder Mead, respectively;
(c) Random local search: Matyas, and Solis-Wets, respectively;
(d) MA for global search: Local search with restarting; Local iterated search, and Simulated Annealing.

The use of the HM in MINLP

The HM is recommended to be used instead of the large computational burning time of the deterministic algorithms. Many times, these methods cannot guarantee that the solution can be found. However, HM are much faster than the conventional algorithms and simplify the optimization problems. Many of the HM mimic the known continuous methods.

Recently, there are many developed HM aimed to solve the practical problems of MINLP. These HM include procedure for rounding [12] and an attempt to generalize a method SQP [13, 14] the method of gradient descent [15–18], penalty function method [19, 20] and adaptive random search [21].

There are two mainly factors to succeed a HM: speed and reliability to find the suboptimal solution. The term of the suboptimal solution is used for any solution very closed to the optimal solution. This type of solution is very attractive in practice. From the NLP relaxation method a satisfable solution can be determined. The simplest HM are based on rounding. The method of solving the MINLP problem is based on the NLP technique combined with the relaxation of all the integer constraints. Finally, the value of the solution is rounded to a near integer point. There are two disadvantages of using thi method: 1. the value of the solution cannot be feasible, and 2. the obtained value for the objective could be too far from the value obtained in the deterministic case. In the paper [12] a "smart rounding" procedure was developed and a discrete line layout scheme was proposed. Recently, the developed smooth landings methods becomes very attractive. In the paper [15], for example, the authors generalize the ablest method of descending with the MINLP problem line search. They reformulate the optimization problem by using an inverse barrier function to obtain the MINLP problem without constraints.

Local optimization

Local optimization methods search the optimal solution x^* in the vicinity of the studied point V (x^*), i.e.

$$f\left(x^*\right) \leq f(x+e), \text{ with } \varepsilon > 0, \text{ sufficiently small.}$$

Remark: the initial approximation of the optimal solution should be known

The solution can be found in discrete space or in continuous space.

In the first case, the neighbourhood of an element is a finite set that can be thoroughly explored.

In the second case, there are two derived branches, function of the derivability property of the objective function:

(a) in case that the derivability feature is fulfilled by the objective function, two basic methods can be used: gradient, and Newton.
(b) in case that the derivability feature is not fulfilled by the objective function, the direct search methods (e.g. Nelder Mead), or the methods based on small random perturbations are recommended.

Global optimization:

- identifying the global optimality of a function: for a minimization problem, the following property should be attained $f(x*) \leq f(x)$, x^* is considered the global minimum value if the above mentioned property is true for all x;
- by using local search methods, if there are local optimimum points, the minimum value of the objective function can be blocked into the local optimum point.

14 Genetic Algorithms (GAs)

Genetic algorithms are some of the most popular evolutionary computing strategies. Among other applications, these have been successfully used for difficult optimization problems, with multimodal, discontinuous and non-differential objective functions [22]. Traditional optimization algorithms often fail in such cases.

Algorithm

Generate the initial population: *N* random individuals
while stop criterion not fulfilled,
Selection stage of the individuals fittest for the next stage
Reproduction stage: by applying crossover and mutation operators, new individuals are created
Recombination stage: new population are created
end while

15 Ant Colony Optimization

Inspiration from the intelligence of the colonies led to some very successful optimization algorithms.

– Ant colony optimization—a way to solve optimization problems based on the way ants communicate indirectly with each other.

Ant colony optimization algorithms.
Ants are agents that [23]:

– move along the nodes in a graph
– ants choose where to go based on the power of pheromones.

The path of an ant is a specific solution of the candidate.
When an ant has finished a solution, the pheromone is placed in its path, depending on the quality of the solution.
This pheromone pathway affects the behavior of other ants through "stigmergy"

Optimization methods

1. Differential evolution (DE)

The general algorithm from the GA method has been applied in the case of DE. After population initialisation, the generation of descendant population through crossover and mutation is obtained (the first step is the reproduction). The next step is the selection of the best element in the population. By using DE method, a new candidate is building. As evolutionary strategy, the multi-modal test function has been used.

Taking 2 elements in the population having the size n = 1 for each element, considering 50 generations, choosing [−0,5 0] the domain of the function to be minimized with crossover probability of 0.55, the Differential evolution algorithm provides the minimum function value (Fig. 4).

2. Particle cluster optimization
3. Clonal selection
4. Ant colony-type optimization
5. Hill climbing and simulated riding

Hill climbing algorithm

Ideas similar to the ascending gradient method if the objective function is maximized

– They start from a randomly selected point in the search space p_0
– Current point $pc \leftarrow p_0$
– One or more neighboring points are generated, pv
– If the objective function in a neighboring point is better than the current one, then $pc \leftarrow pv$;
– Choose the first best neighbor (*Greedy*, *Simple HC*)
– The best neighbor is chosen (Steepest *Ascent HC*)

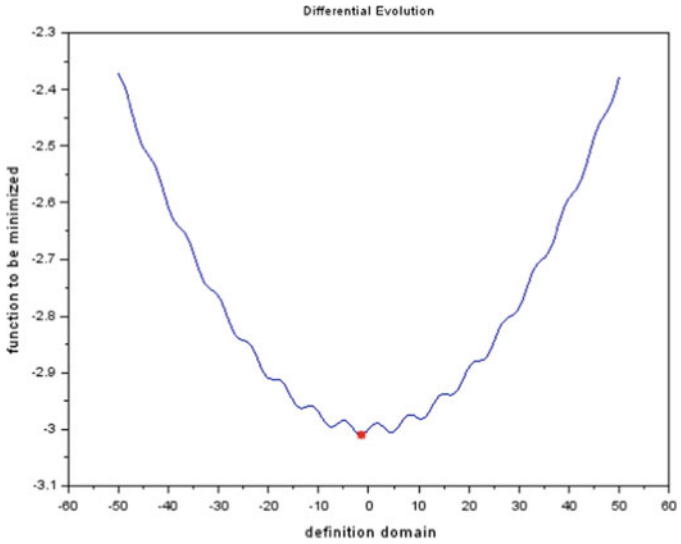


Fig. 4 Graphical representation of the function and elements of the population. Displaying the value of the best element in the population [10]

16 Simulated Annealing (SA)

Stochastic algorithm inspired by metallurgical quenching. Heating and then controlled cooling of a material increases the size of the crystals and reduces defects.

SA algorithm

Suppose a minimization problem, with the objective function E

- If the neighbor is better ($E_v < Ec$), then $pc \leftarrow pv$
- If the current state is better than the next one, the neighbor ($E_v > Ec$), then ΔE is caluclated;
- The difference of the objective functions is calculated: $\Delta E = Ev - Ec$
- It is considered the current temperature T, high at the beginning and decreasing in time.
- Probability of accepting the transition to the lower state is: $P = \exp(-\Delta E/T)$

Example: The function to optimize is the Rastrigin function by using SA method.

In Figs. 5–6, by taking two parameters of Rastrigin function, A = 1, n = 2, the dynamic process of the optimization, and the contour plot of the Rastrigin function during SA optimization are shown (Fig. 5).

The dynamic process of the optimization by using SA.

The intial random initialisation:

$\times 0 = -0.0060798\ 0.0412484$

$y0 = 0.0358653$.

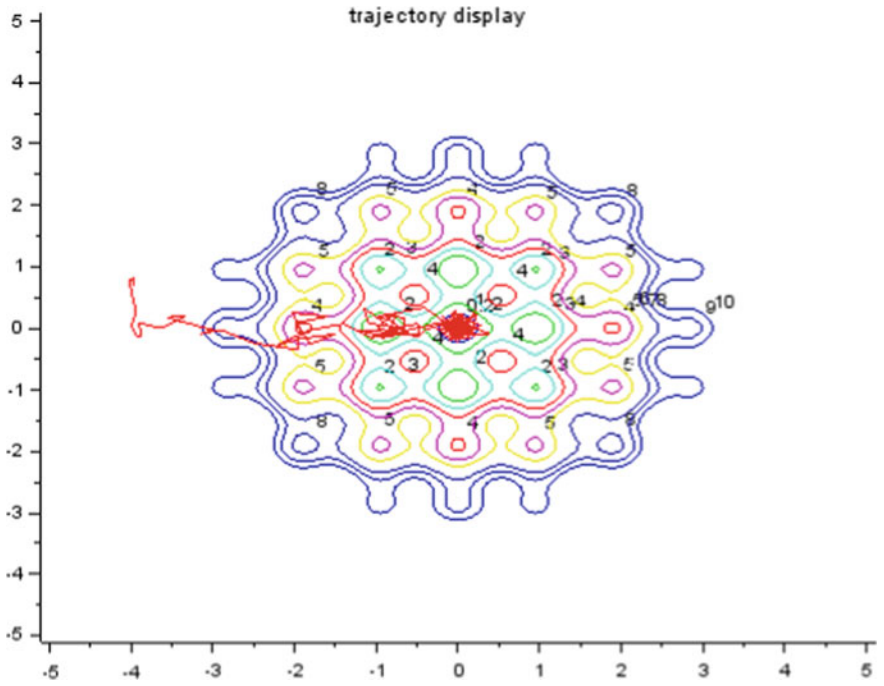


Fig. 5 Contour plot of the Rastrigin function during SA optimization and the dynamic process of the optimization process evolution (red line) [10]

After 2000 number of iterations, by using SA method of optimization (Fig. 7), the coordinates of minimum function obtained at (0.0023824 −0.0028886), corresponding value by Rastrigin function is 0.0002908 (Fig. 6).

By choose a large number of iterations more than 2000), the function has a global minimum at $\mathbf{x} = (x_1, x_2) = (0,0)$, and the value $f(\mathbf{x}) = 0$.

17 Particle Swarm Optimization (PSO)

Swarm intelligence [24] domain that includes intelligent techniques based on the collective behavior of systems with self-organization and without centralized control

Model of the particle assembly

The Particle Swarm Optimization (PSO) [4, 25] technique was proposed by James Kennedy and Russell Eberhart for nonlinear function optimization (1995).

The source of inspiration: the behavior of bird flocks, fish banks, and bees swarms, these are assimilated to a set of particles that move in the search space to identify the optimum [25].

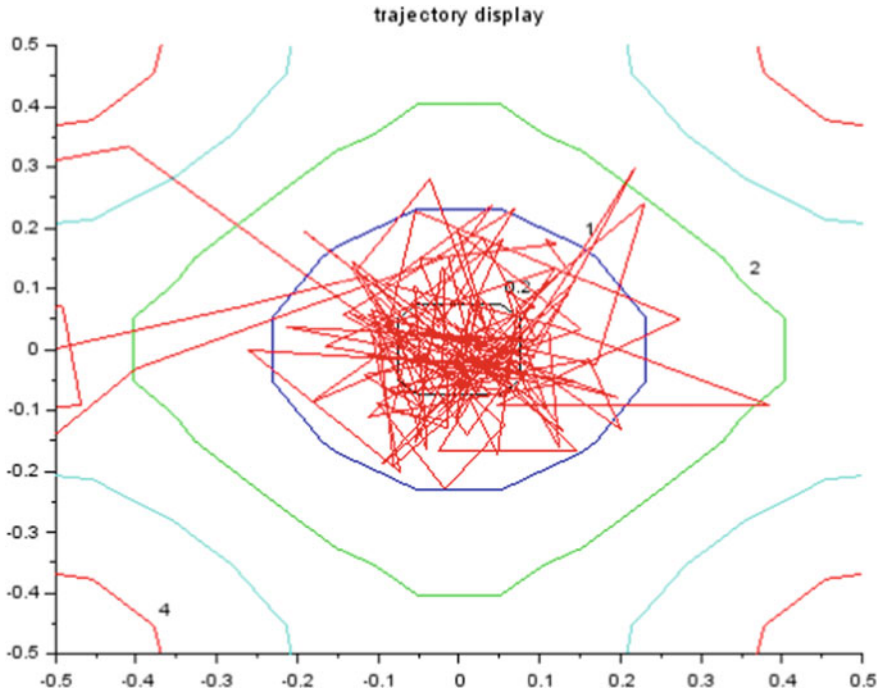


Fig. 6 The dynamic process of optimization (red line) in the specified [−0.5, 0.5] interval

It uses a set of particles whose positions are in the field of objective function and which are modified by an iterative process.

At each iteration, the new position of each particle is determined according to: the current position of the particle.

The best position encountered by the particle is the *local best solution*

Best position found by the whole is the *global best*.

PSO Algorithm:

Initialization of particle positions

REPEAT

speed calculation
update positions

UNTIL < stop condition>

The PSO algorithm is applied to the Rastrigin function.

The same significance of the optimization process (for Rastrigin function) results are obtained as in SA method, discussed above (Figs. 8, 9, 10).

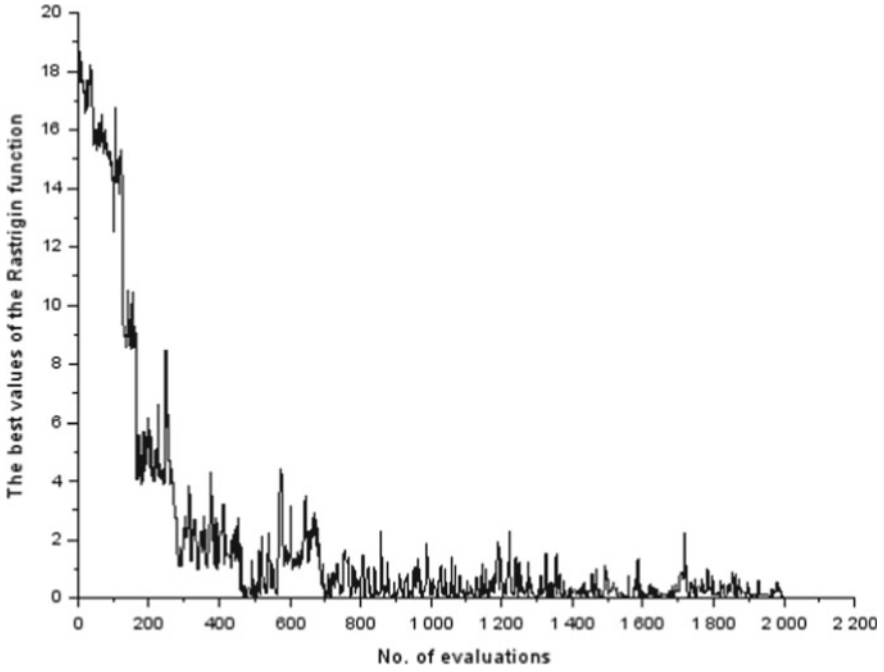


Fig. 7 Optimization process of Rastrigin function of two variables

18 The Model of the Bee Swarm

The source of inspiration: the intelligent behavior of bees in the process of identifying food sources (nectar)

It uses a population of "bees" consisting of three categories: Bees "allocated" to a food source (workers)

Observer bees
Scorpion bees

The model of the bee swarm

Step 1: Initialize the locations in the search space where the worker bees are placed

Step 2: How long the continuation condition is satisfied: The working bees transmit information about the quality of the location where they are to the observer bees; each bee observer selects a location; the selection is based on a probability distribution determined by the values of the associated scores;

Working bees explore the vicinity of their location and move to another neighboring location if it is better; if a worker bee does not discover a better configuration in a limited number of steps then it is relocated to a position determined by a bee researcher. Scorpion bees randomly change their position.

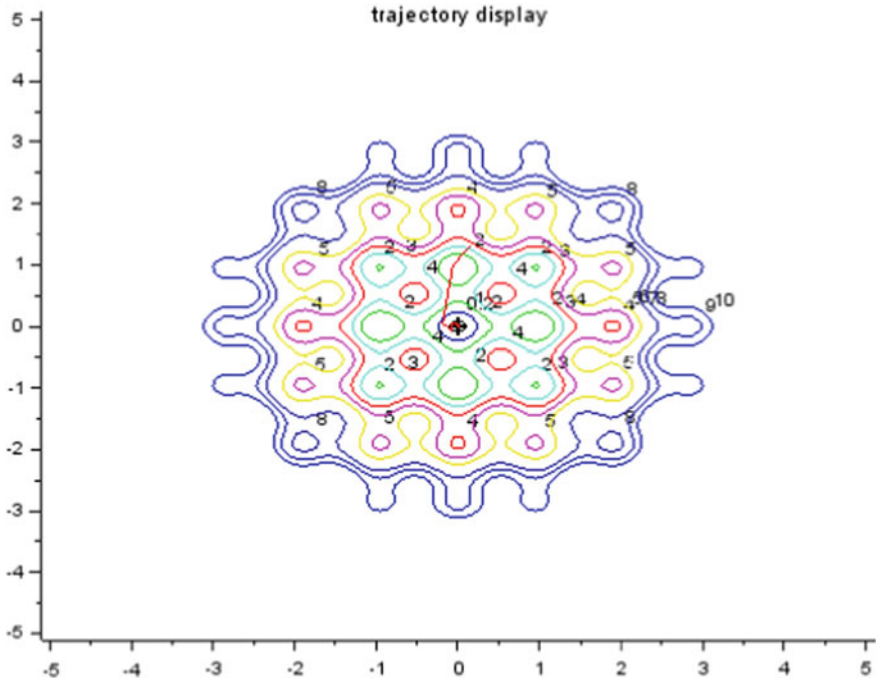


Fig. 8 Contour plot of the Rastrigin function during PSO optimization and the dynamic process of the optimization process evolution (red line) [10]

19 The Firefly Model

Firefly algorithm [26]

The source of inspiration: interactions between firefly based on the light signals they emit.

Main idea of implementation

- Each element of the population corresponds to the position of a firefly
- Each degree of firefly is associated with a degree of brightness (correlated with the value of the objective function associated with the corresponding element in the population).
- The movement of the firefly is guided both by the distance between their positions and the value of the brightness

– Position x_i is shifted to position x_j (if x_j has higher brightness) using the specific parameters (alpha, beta and gamma are control parameters and epsilon is a random value with normal distribution)

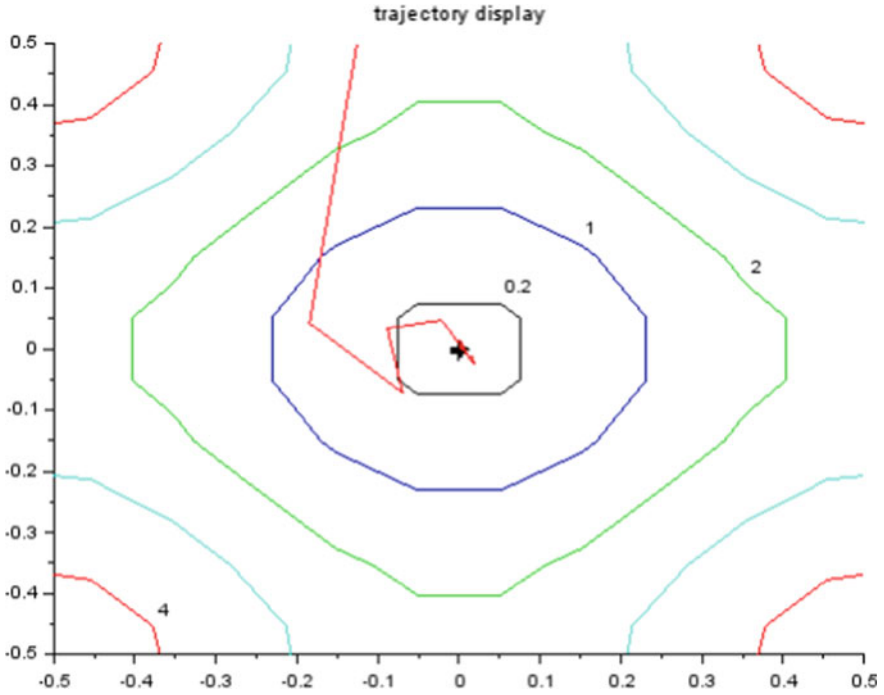


Fig. 9 The PSO dynamic process of optimization (red line) in the specified [-0.5, 0.5] interval [10]

Conclusions:

- Differential evolution has the same genetic operators as classical evolutionary algorithms, but their order and mode of operation is different
- There are many other optimization algorithms inspired by nature
- Particle cluster optimization is inspired by the behavior of birds
- The clonal selection is inspired by the immune system
- Ants colony-type optimization is inspired by the way they search for food
- The simulated ride is inspired by metallurgy.

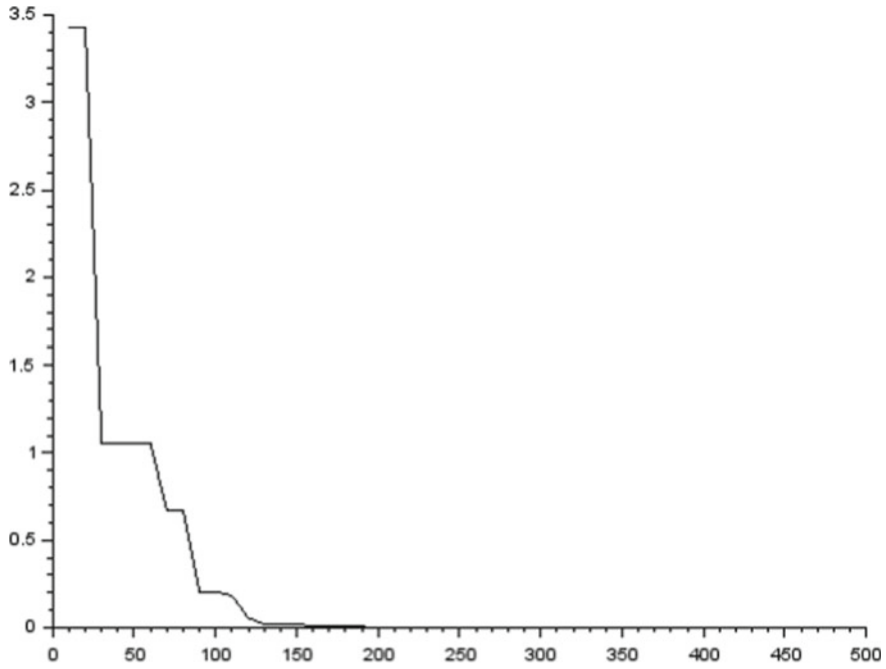


Fig. 10 Particle Swarm Optimization process of Rastrigin function of two variables. Convergence of the optimization process [10]

20 Very Large Scale Neighborhood Search

The main cause of generating complicated difficulties in solving real optimization problems is the **dimension**: such a problem complicated is **too high**. In mathematical programming the size of the problem is relatively, depending on the following parameters:

- Number of variables and the number of constraints;
  - Complexity of the function expressions and the constraints;
  - Performance of the numerical calculus equipment: hardware and software

  In principle, methods for solving large scale problem divides into:
- Direct Methods: they specializes a general procedure to the specifics of a particular class of optimization problems.

Let take into consideration he case of a linear program with upperr bounded variable:

$$\begin{cases} \sum_{j=1}^{n} a_{ij} x_j = b_i \quad , i = \overline{1, .m} \\ 0 \le x_j \le u_j \quad , j = \overline{1, .n} \\ (\max) f = \sum_{j=1}^{n} c_j x_j \end{cases}$$

The classical solution method involves the transformation of the upper conditions in equality:

$$\begin{cases} \sum_{j=1}^{n} a_{ij} x_j = b_i \quad , i = \overline{1, .m} \\ x_j + x_{n+j} = u_j, \quad j = \overline{1, .n} \\ x_j \ge 0 \quad , \\ (\max) f = \sum_{j=1}^{n} c_j x_j \end{cases}$$

The result consists of the program with m + n restriction and 2n variables of those bases were of the order m + n matrices.

- *Indirect methods*, based on the **decomposition large** problem into smaller, **interconnected** sub-problems. Subproblems may be solved **independently** (and if it is even possible **at the same time**), but the fact that the subproblems interacts involves existence of a **coordination** mechanism. Thus, solving the original problem is made in two levels.

  - at the First level—lower—subproblems are solved and the results are communicated:
  - Second level—higher—which analyzes the results and transmit the new parameters to the lower level.

At level one there is a resumption of calculations (re-optimization); new results sent to the upper level for analysing, so on.

Important is the fact that this iterative process converges.

The nonlinear-optimization (NO) is divided into: Unconstrained (UNO) and Constrained (CNO).

The algorithms used to solve *Nonlinear programming problems in Very Large Scale Systems* are as follows:

- Augmented Lagrangian Methods
- Sequential Quadratic Programming
- Feasible Sequential Quadratic Programming
- Reduced-Gradient Methods

Constrained Nonlinear Optimization

Constrained nonlinear optimization problems can be solved by using one of the following algorithms:

- **Interior-point (IP).** The method is used for large-scale problems.
  The IP algorithm estimates the Hessian matrices by using:
  - BFGS (dense)—Limited-memory Broyden Fletcher Goldfarb Shanno
  - Limited memory BFGS
  - Hessian-multiply function
  - Current Hessian (sparse or dense)
  - Finite difference of gradients. This case does not require the knowledge of sparsity structure
- **SQP** algorithm.
  This algorithm is specified to general nonlinear optimization problem.
- The **trust-region reflective (TRRA)** algorithm can be succesfull used in case of the bound constrained problems or linear equals only.

  For the TRRA, the Hessian matrices can be obtained by using:

- Finite difference of gradients, sparsity structure of the Hessian
- Current Hessian (sparse or dense)
- Hessian-multiply function

  Both methods, the IP and TRRA use lower memory usage.

21 Security-Constrained Unit Commitment (SCUC)

The ***Unit Commitment*** (UC) problem is based on the optimization algorithms and it used in power sytems. The idea is to coordinate a set of the electrical generators such that the energy demand with minimal costs is attained, or maximize the profits from energy production.

Coordination of generating units is a difficult task for several reasons:

– there are a large number of units;
– the costs of the energy production from different type of generators can varying. At the same time, different constraints conditions for each generator can be met (due to the different technologies of energy production);
– the generators are spread over a large area into a country. For this reason, the response capacity of the electrical network should be considered. The complex power flow data should be determined to assure the load demand.

SCUC commitment consists of two components, system security and economic dispatching. The objective of the problem is focused exclusively on the economic dispatch of generators with tender segments, without loading costs, starting costs

and other costs incurred during operation. The lowest cost is desired depending on the system operators.

22 Conclusion

The complexity of the mathematical problems into the power systems is very high. Different methods should be used to find the adequate solution.

The class of the numerical methods is very usefull to increase the speed of finding the solution with high precision.

To facilitate the development of numerical problems, the need to establish procedures is stringent, this leads to algorithms development. However, the solution involves the following steps; modeling, choice of numerical methods, operation, results, and interpretation of the obtained results. In power systems, power is known rather than current; thus, the equations resulting in power term are more appropriately. This type of equations (power flow equations), are nonlinears. The solution of them is obtained by using iterative techniques from the numerical algorithms.

The authors of this chapter made an incursion in optimization problems and their solution through the numerical methods. In order to implement the optimization algorithms different programming languages have been used, as Matlab [27], and Scilab [10] Starting with the formulation of the optimization problem, the authors gives different solutions to the exposed topics: optimization of a real variable functions, the extremes of the functions defined over an interval, extremes of functions for which the derivative does not exist, the method of small pertutbations (variations), extremes of the multivariable functions, the minimum of a function of two variables with constraints, the Lagrange multiplier method for determining the minimum of a constraint function, types of optimizations, unconstrained optimization problems, nonlinear optimization without constraints, optimization problems with linear constraints and nonlinear objective function, the problems of convex programming, the problems of separable programming, the problems of non-convex programming, optimization methods with constraints, Mixed Integer Nonlinear Programming (MINLP), Very Large Scale Neighborhood Search, Security-Constrained Unit Commitment (SCUC).

Advanced optimization algorithms are introduced and the simulation results are provided (Genetic Algorithms, Particle Swarm Optimization, and Simulated Annealing).

References

1. Rosu E, Nichita C, Bivol I, Gaiceanu M (1999) Optimizarea energetica a sistemelor de conversie electromecanica (Energetic Optimization of the Electromechanical Conversion Systems). Technical Press

2. Gaiceanu M (2009) Optimal control of the electric drive systems. Galati Univerity Press (University)
3. Metode de optimizare numerica, Ion Necoara, acse.pub.ro
4. https://press.princeton.edu/sites/default/files/inline-files/Absil_Chap7.pdf.Last Accessed 2019
5. http://www.numerical.rl.ac.uk/people/nimg/oupartc/lectures/raphael/lectures/lec6slides.pdf. Last accessed 2019
6. Gould N (2006) An introduction to algorithms for continuous optimization. Oxford University Computing Laboratory and Rutherford Appleton Laboratory Copyright 2006 by Nicholas Ian Mark Gould
7. Botan C (2007) Optimization techniques, Politehnium
8. http://neos-guide.org. Last accessed 2019
9. www.math.ubc.ca. Last accessed 2019
10. http://www.scilab.org/. Last accessed 2019
11. Bolte J, Hochart A, Pauwels E (2018) Qualification conditions in semialgebraic programming. SIAM J Optim 28(2), 1867–1891. https://doi.org/10.1137/16m1133889
12. Olsen GR, Vanderplaats GN (1989) Method for nonlinear optimization with discrete design variables. AIAA J 27:1584–1589
13. Cha JZ, Mayne RW (1989) Optimization with discrete variables via recursive quadratic programming: Part 1—concepts and definitions. Trans ASME, J Mech, Transm, Autom Des 111:124–129
14. Cha JZ, Mayne RW (1989) Optimization with discrete variables via recursive quadratic programming: Part 2—algorithms and results. Trans ASME, J Mech, Transm, Autom Des 111:130–136
15. Amir HM, Hasegawa T (1989) Nonlinear mixed-discrete structural optimization. J Struct Eng 115:626–646
16. Bremicker M, Papalambros PY, Loh HT (1990) Solution of mixed—discrete structural optimization p roblems with a new sequential linearization algorithm. Comput Struct 37:451–461
17. Loh Han Tong, Papalambros PY (1991) "A sequential linearization approach for solving mixed-discrete nonlinear design optimization problems" Transactions of the ASME. J Mech Des 113:325–334
18. Loh Han Tong, Papalambros PY (1991) "Computational Implementation and tests of a sequential linearization algorithm for mixed-discrete nonlinear design problems" Transactions of the ASME. J Mech Des 113:335–345
19. Davydov EG, Sigal IKh (1972) Application of the penalty function method in integer programming problems. Eng Cybern 10:21–24
20. Li Han-Lin (1992) An approximate method for local optima for nonlinear mixed integer programming problems. Comput Oper Res 19:435–444
21. Salcedo RL (1992) Solving nonconvex nonlinear programming and mixed – integer nonlinear programming problems with adaptive random search. Ind Eng Chem Res 31:262–273
22. Houck CR, Joines JA, Kay MG (1996) Utilizing lamarckian evolution and the baldwin effect in hybrid genetic algorithms. meta-heuristic research and applications group. NCSU-IE Technical Report 96-01. Department of Industrial Engineering, North Carolina State University
23. Chandrasekhar A, Gordon DM, Navlakha S (2018) A distributed algorithm to maintain and repair the trail networks of arboreal ants. Sci Rep. 8(1), 9297. Published 2018 Jun 18. https://doi.org/10.1038/s41598-018-27160-3
24. Beni G, Wang J (1989) Swarm intelligence in cellular robotics systems. In: Proceeding of NATO advanced workshop on robots and biological system
25. Kennedy J, Eberhart R (n.d.) Particle swarm optimization. In: Proceedings of ICNN'95—international conference on neural networks. https://doi.org/10.1109/icnn.1995.488968
26. Karaboga D, Basturk B (n.d.) Artificial Bee Colony (ABC) optimization algorithm for solving constrained optimization problems. Found Fuzzy Log Soft Comput 789–798. https://doi.org/10.1007/978-3-540-72950-1_77
27. https://www.mathworks.com. Last accessed 2019

28. http://www.particleswarm.info/. Last accessed 2019
29. Yang XS (2009) Firefly algorithms for multimodal optimization. Lect Notes Comput Sci 169–178 https://doi.org/10.1007/978-3-642-04944-6_14

Ill-Posed Inverse Problems in Electrical Engineering Applications

Andrei Ceclan, Dan D. Micu, Levente Czumbil, Horia Andrei, Marian Gaiceanu, Marilena Stanculescu, and Paul Cristian Andrei

Abstract In this chapter some ill-posed inverse electromagnetic and power engineering problems are introduced, both at a theoretical introductory and mathematical modelling level and detailed regarding their numerical solving case studies based, by the application of several regularization techniques starting from classical Tikhonov approach up to singular values decomposition procedures. Fredholm integral equation mathematical modelling is presented in the physical definition of the inverse electromagnetic and/or power engineering problems, accompanied by explanations regarding the physical significance.

Keywords Ill-posed · Inverse problem · Regularization · Eigenvalue · Singular value · Integral equations

A. Ceclan · D. D. Micu (✉) · L. Czumbil
Numerical Methods Research Center, Electrical Engineering Department, Technical University of Cluj-Napoca, Cluj-Napoca, Romania
e-mail: dan.micu@ethm.utcluj.ro

A. Ceclan
e-mail: andrei.ceclan@ethm.utcluj.ro

L. Czumbil
e-mail: levente.czumbil@ethm.utcluj.ro

H. Andrei
Doctoral School of Engineering Sciences, University Valahia, Targoviste, Romania
e-mail: hr_andrei@yahoo.com

M. Gaiceanu · M. Stanculescu · P. C. Andrei
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania
e-mail: marilena.stanculescu@upb.ro

P. C. Andrei
e-mail: paul.andrei@upb.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_9

Abbreviations

ART	Algebraic reconstruction technique
IEP	Inverse electromagnetic problem
IPIEP	Ill posed inverse electromagnetic problem
HVPL	High voltage power line
SVD	Singular value decomposition
TSVD	Truncated singular value decomposition

1 Introduction

The chapter structure is as following: a first section introduces the Fredholm integral equations applied for ill-posed inverse problems including in the modelling of the electromagnetic field problems. Next, the main key performance indicators related to the inverse problems and regularization techniques to solve them are presented. In the electrical engineering domain, a power-flow case study is then introduced as a practical case study application to be numerically solved using regularization of the ill-posed problem. In this respect, three main categories of numerical methods are presented: a section dealing the so-called intermediate methods, a next section of so called special regularization methods and a last section of the chapter presenting an original and simple approach called harmonic regularization numerical method.

The established objectives are: to present how the inverse electromagnetic or power engineering problems are modelled and when their numerical solving, if ill-posed should be addressed with regularization techniques.

The first section introduces the Fredholm integral equation as a mathematical modelling tool. Then, the main key performance indicators in ill-posed inverse problems are presented using the introduction of a power flow case study. The proposed practical application is continued in the next sections, by showing different numerical regularization techniques applied on the problem. The chapter ends with useful conclusions and further proposed actions and contributions.

2 Fredholm Integral Equation

As it was indicated, a series of applications of electric or magnetic field synthesis are modelled using Fredholm Integral Equation, as incorrectly formulated inverse problems.

An integral equation is an equation in which the unknown function appears under the integral. There is no general method for solving integral equation. Their solutions and their existence depend on the particular forms under they appear.

An integral equation is called linear if the unknown function is separated by the operator under the integral, which is applied to it. The general form of a linear integral equation is:

$$f(y) \cdot \phi(y) = g(y) + \lambda \cdot \int_{a(y)}^{b(y)} K(x, y) \cdot \phi(x) dx, \quad c \leq y \leq d \tag{1}$$

The unknown function is $\varphi(x)$, the operator function $K(x, y)$ is called core (it can be integral square) and g(y) is called effect function (imposed or evaluated). These functions can also have complex values. The integral equation is considered homogeneous if the effect function g(y) = 0 on the indicated domain, and if g(y) is different from 0, the equation is inhomogeneous. λ is in general a parameter, called own value of the integral equation.

If the integral limits are independent of y, so as to $a(y) = a$ and $b(y) = b$, express (1) is called Fredholm Integral Equation (2a, 2b). If $a(y) = a$ and $b(y) = y$, the integral equation is called Volterra type (3a, 3b).

If the unknown function $\varphi(x)$ appears only under the integral and f(y) = 0, we have an integral equation of the first case (2a, 3a). If $f(y) = 1$, the integral equation is of the second case (2b, 3b):

$$0 = g(y) + \lambda \cdot \int_{a}^{b} K(x, y) \cdot \phi(x) dx \tag{2a}$$

$$\phi(y) = g(y) + \lambda \cdot \int_{a}^{b} K(x, y) \cdot \phi(x) dx \tag{2b}$$

$$0 = g(y) + \lambda \cdot \int_{a}^{y} K(x, y) \cdot \phi(x) dx \tag{3a}$$

$$\phi(y) = g(y) + \lambda \cdot \int_{a}^{y} K(x, y) \cdot \phi(x) dx \tag{3b}$$

In the specific bibliography [1], it is specified that the linear Fredholm Integral Equations appear in the modelling of synthesis problems with causal functions, unknown, of electro/magnetic nature, and the nonlinear ones, in the modelling of synthesis problems with causal functions dependent on the material or nature geometry, the shape of the boundaries of the PIE domain [2].

No linear homogeneous Fredholm integral equation has been identified. Their appearance could mean a formulation in the absence of a desirable effect (effect zero identical effect function), which logically has no practical interest.

For the first kind Fredholm integral equation, if a disturbance is applied to the solution function $\varphi(x)$ of the integral equation, dependent on a constant amplitude ε and an increasing string of values $p = 1, 2\ldots$ such that:

$$\Delta\phi(x) = \varepsilon \cdot \sin(2 \cdot \pi \cdot p \cdot x) \tag{4}$$

Then the corresponding disturbance in the effect function is evaluated:

$$\Delta g(y) = \varepsilon \cdot \int_a^b K(x, y) \cdot \sin(2 \cdot \pi \cdot p \cdot x) dx \tag{5}$$

According to the Riemann-Lebesgue lemma [3], the perturbation of the effect tends to zero $\Delta g \to 0$, as $p \to \infty$, so that a high value of the $\|\Delta\varphi\|/\|\Delta g\|$ is reached with the increase of the problem domain size.

On the other hand, the idea of the disturbance in the causal function reflects the uniqueness of the solution and its instability even when the disturbance of the effect is almost unnoticeable. This fact shows the character of the mathematical model of ill-posed (incorrectly formulated) inverse problem of the first kind Fredholm integral equation.

In order to solve the Fredholm equation, in relation with the core expression, the following methods are possible [4]:

(1) Separation of the core variables, $K(x, y) = \sum \alpha_k(x) \cdot \beta_k(y)$ and expressing the solution function as a development after functions of the same variable $\varphi(x) = \sum A_k \cdot \alpha_k(x)$, where the coefficients A_k imply the solution of a linear algebraic system of equations.
(2) The serial expansion of the cause function, after a set of independent linear functions, which forms an orthogonal base—a method called collocation; the unknowns are the weighting coefficient of these basic functions in the series development and are deduced by solving an algebraic system of linear equations.
(3) The discretization of the independent variables (x_i, y_i) so that an algebraic system of linear equations is formed, whose solution directly gives a string of values for the cause function. In order to obtain as accurately as possible, the unknown values, here is also important the method of approximation of the integral (by simple summation, with the trapezoid rule, Simpson, of Gaussian quadrature), as well as the method of interpolation of the solution.
(4) Successive approximation of the solution, after the iterative relation:

$$\varphi_n(y) = g(y) + \lambda \cdot \int_a^b K(x, y) \cdot \varphi_{n-1}(x) dx; \quad n = 1, 2, \ldots; \quad \varphi_0(x) = g(x) \tag{6}$$

For the latter method to be convergent, the core of the integral equation and the causal function must be linear.

As it was indicated, for the first three methods listed, obtaining the solution involves solving an algebraic system of linear equations. For a square system of equations deduced from the second kind of equation, from the point of view of conditioning the solution, in the specialized bibliography no disadvantage is reported. Regarding the uniqueness of the solution, this goal can be fulfilled only if the own values of the integral equation of the second kind do not remove the determinant of the matrix operator.

For consistent details on the Fredholm integral equations, other studied works are recommended [4–8].

3 Key Performance Indicators in Inverse Problems Regularization

This paragraph introduces a set of indicators to characterize the level of instability of an incorrectly formulated inverse problem. The definition takes place in a wider frame than the strict one of the inverse field synthesis problems, with small examples of the author's own contributions in the energy field.

It starts from the linear algebraic system as a matrix equation ($A \cdot z = u$), corresponding to the firs kind Fredholm integral equation. To the matrix of coefficients A, decomposition is applied according to singular values [9], or the decomposition according to the own values [10], corresponding to the relations:

$$A = U \cdot \Sigma \cdot V^T,\ A = P \cdot \Psi \cdot P^2.$$

The orthogonal matrices U, V are called single vectors, and the matrix P is called by passage and contains its own vectors. The matrices Σ and Ψ are diagonal and contain the singular values, respectively those corresponding to the coefficient matrix.

In the previous paragraph it was shown that the incorrect formulation of the inverse problems integrally modelled with Fredholm equations of first kind, appears phenomenologically due to the Green core function.

In the numerical model, this represents even the coefficient matrix. Based on such a hypothesis, it is considered that before the beginning of the solution it may be useful to characterize the level of instability of the solution.

3.1 Own Observations on the Spectrum of Singular Values

The following observations are made:

(1) The multiplication of the cores with specific constants, properly amplifies the ranges of variation of the singular and own values.

(2) The singular values of a matrix, as a Green core from a synthesis problem, are positive and spread over an extended range.
(3) The increase of the discretization level implies a marked widening of the level on which the singular values spread.
(4) The own values of a matrix, as a Green core from a synthesis problem, may coincide with the singular values of the same matrix, may be multiple, or differ substantially from them, taking even complex values.

Without drawing any further conclusions on these observations, it is recalled that solving the system (6) by decomposition according to the singular values, acts as an amplification of summed vectors, as the singular values extend on a wider level, namely:

$$z = \sum_{i=0}^{N} \frac{U_i^T \cdot u}{\sigma_i} \cdot V_i \tag{7}$$

where U_i, V_i, represents the singular vectors extracted from the orthogonal matrices, u the right member of the system of equations, as an effect vector, and σ_i the individual singular values obtained by decomposition.

The direct solution by the relation (7) leads to physically inadmissible solutions and a high level of instability to the disturbances that may occur [3, 9, 11].

At this point, it can be specified that a wide spread of the single values represents a sufficient indicator to characterize the level of physical instability and inconsistency of the solution of an inverse problem modelled by (6). However, this does not appear to be a novelty aspect, but it must be remembered that it works regardless of the results obtained for decomposition according to its own values.

3.2 Characterization of Stability as Disturbances with the Spectrum of Singular Values

Further, an application from the field of electricity is analysed, characterizing the natural stability of a power grid, at small perturbations that may occur due to the coupling/decoupling of some consumers [12]. The mathematical model corresponding to this case consists of a system of ordinary differential equations.

The matrix of coefficients specific to the system is subject to decomposition after it own values, and then after the singular values.

As has been observed so far, no particular emphasis is placed on numerical methods for evaluating the own values (based on factorizations, Leverrier-Fadeev, Krylov, Rayleigh, Jacobi, Givens), starting from the characteristic equation, nor on the determination of the singular values.

The existing programs used in this paper, contain algorithms with enough performance to not go into specific details. Figure 1 shows the configuration of the initially

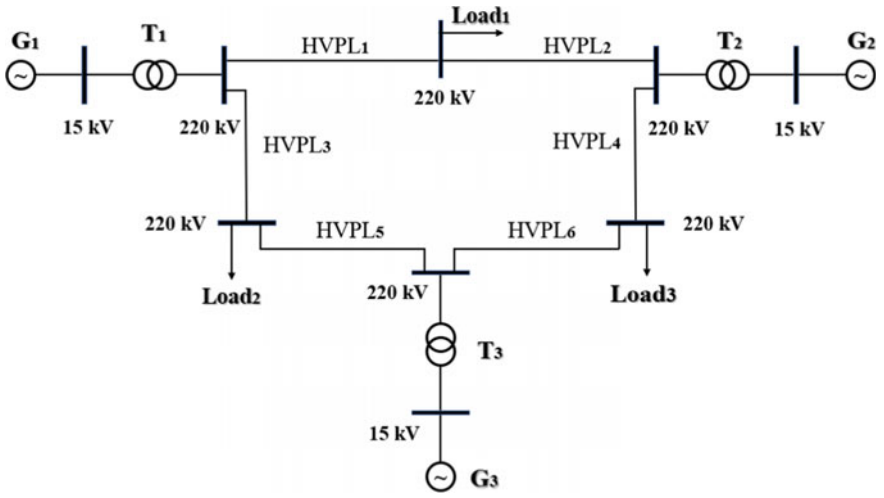


Fig. 1 Configuration of the analysed electrical system

studied electrical grid, in which all the parameters of the system elements and the characteristics of the concrete operating regime are specified. To solve the system, it is possible to apply the decomposition method according to the singular values.

After the appropriate evaluations, the coefficient matrix looks like this:

$$A = \begin{bmatrix} -0.00416 & -153.46 & 0 & -33.841 \\ -0.0057 & 0 & 1 & 0 \\ 0 & -59.524 & -0.00532 & -104.096 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{8}$$

The own values and the singular ones evaluated for this matrix are expressed in Table 1, together with their physical significance for the electrical system. For the studied application, the fact that there is an own value with the real side as positive, indicates an instability of the system. In addition, the imaginary part of the own values gives information about the natural frequencies of the oscillation system.

Table 1 Interpretation of the corresponding own and singular values

Eigenvalues	Significance	Singular values
−0.70028	–	182.223
0.69612	If positive eigenvalue, then there is a natural instability to perturbations in the power system [12]	76.611
−0.00266 + 12.7762i	–	1.414
−0.00266 − 12.7762i	–	0.00403

What intervenes further, is as shown in Table 1, the correlation with the single values, whose range appears to be sufficiently wide.

A series of other coefficients matrices deduced from the configurations of electrical systems were afterwards tested on the same chain. In this way, without claiming a generality, the following were concluded: each time own values appear with the real side as positive, the singular values have a wide range of variation; when all the own values have the real side as negative (stability of the electrical system), the singular values fit in a restricted range. The correlation started from the Green core coefficient matrices.

The validity of characterizing the range of instability, by spreading the singular values, in the energy field can be assumed without demonstration for this class of problems, as an alternative to the indications offered by the own values.

Based on these mentions, the proposal to evaluate the spread of singular values as an indicator of stability on incorrectly formulated problems, or poorly conditioned is shown to be justified.

3.3 Picard's Condition as Performance Indicator

Decomposition of the coefficient matrix after singular values is an efficient method of studying the character of IPIEP [13, 14]. Specialty bibliography proposes as an **indicator** for the solution's degree of instability Picard's condition. For the discrete model from (7) this condition consists in graphical representation the variance of singular values as well as the variance of the scalar product $U_i^T \cdot u$ repeated for $i = 0 \ldots N$.

If on the graph the curve of the product of the singular vectors with the effect vector, does not decrease in mean by at least as fast as the singular values, the problem is considered incorrectly formulated, with an unstable solution. The explanation stems from (7), in which between the two variances exists a discrepancy namely, the denominator becomes of ever lower values, then the single vectors V_i are amplified in the sum, thus irremediably distorting the solution.

For the same core expressions in Table 1, the graphs of the Picard indicator type are presented, in the situation of defining unit value effect functions (Fig. 2).

The Picard indicator clearly shows in all four test cases that the solution is unstable. As an addition to the current bibliography, it is proposed to evaluate the Picard condition indicator and the regularization model, and after solving it for the inverse problem studied. Thus, the performance of the degree of regularization applied can be evaluated.

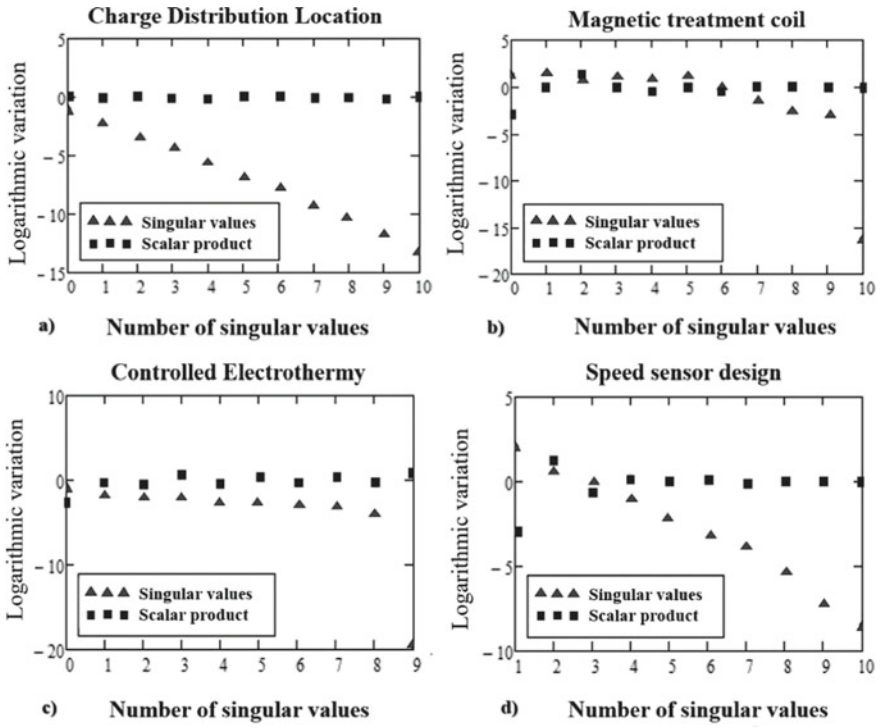


Fig. 2 Verification of the Picard condition, as a performance indicator of PIE for different applications

3.4 Conditioning Number as a Performance Indicator

A third performance indicator in solving IEP with integral model is, as mentioned in the previous chapter, the **conditioning number**. Extensive treatments of this can be found in [9, 10, 14, 15].

The definition of the conditioning number based on the ratio between the coefficient matrix norms, respectively its inverse. These norms represent in fact the maximum singular value, respectively the minimum value of the matrix decomposition:

$$K_A = \|A\|_2 \cdot \left\|A^{-1}\right\|_2 = \frac{\sigma_{\max}}{\sigma_{\min}} \tag{9}$$

In expression (9) the spectral norm was defined as this is the most commonly used. As can be seen, in relation to the results obtained with the previous indicators, for a matrix that causes instability, the value of the conditioning number is high. However, to clearly establish a link between the instability of the solution and the conditioning number indicator, Eqs. (9) and (10) of the link between the maximum percentage

variation of the solution are used, in relation to the disturbances that may occur in the right side, or in the core matrix structure. The deduction of these relationships is elegantly described in [9].

The disturbance acts on the effect vector:

$$\varepsilon_z[\%] = \frac{\left\|z' - z\right\|_2}{\|z\|_2} \leq K_A \cdot \frac{\left\|u' - u\right\|_2}{\|u\|_2} = K_A \cdot \varepsilon_u[\%] \tag{10}$$

where u'—the disturbed effect vector; z'—solution resulting from the effect of the disturbance.

It is obvious that a high value of the conditioning number allows a small error in the effect vector to lead to exaggerated variations of the solution. The issue is due exclusively to the core matrix of the coefficients. There is an expression on the instability of the solution, highlighted by the conditioning number: "A flapping of a butterfly in New York, causes a typhoon in Tokyo."

The disturbances that affect the effect vector can be physical—measurement errors, or numerical—errors of rounding, truncation. Given that the relative error of the effect vector is within a percentage range of (1–5)%, if the conditioning number as an amplification factor has a value over 10^2 means that the initial solution may be affected by a percentage instability up to 500%.

The disturbance acts on the coefficient matrix (on the problem structure):

$$\varepsilon_z = \frac{\left\|z' - z\right\|_2}{\|z\|_2} \leq \frac{K_A \cdot \frac{\|P\|_2}{\|A\|_2}}{1 - K_A \cdot \frac{\|P\|_2}{\|A\|_2}} \tag{11}$$

with P—disturbance matrix, which modifies the initial structure of the coefficient matrix.

The lower the disturbance norm than the coefficient matrix norm, and their ratio is closer to the value of the conditioning number, the weaker the immunity to disturbances of the solution. It is considered that in this case, the disturbance on the matrix structure influences the immunity of the solution to a greater degree.

The disturbances that appear in the matrix structure can be physical in nature—modification of the environmental characteristics, of the boundaries of the problem domain, or of numerical nature—rounding, truncation.

In this context, in which the conditioning number enters in direct correlation with the disturbances that may occur in the equation system (7), the author considers this instrument as an indicator. Moreover, what has not been found in the bibliography, based on the imposition of a maximum limit of the disturbance that may occur, together with an imposition of the admissible limit of the solution modification, can be estimated a value of the conditioning number, from which the problem is susceptible to unstable solutions. Inequalities (10) and (11) respectively are used, with interpretations:

$$K_A \geq \varepsilon_z / \varepsilon_u \tag{12}$$

when the disturbance affects the effect vector, the maximum disturbance limits are imposed ε_u and maximum allowable solution variation ε_z, the value of the conditioning number from which the system is considered to be unstable is given by the report (12).

If for the disturbances affecting the sources, the effect vector, the interpretation is immediate, in the case of those that intervene in the matrix structure, the interpretation on a minimum conditioning number, from which the instability of the solution can be considered does not appear to be relevant.

In the case of the relationship (11) it is considered appropriate to establish an admissible classification of the norm disturbance $\|P\|_2$. It is necessary to consider two cases, which verify the inequality relation (11):

Case I: $(1\text{-}K_A \cdot \|P\|_2/\|A\|_2) > 0$ implies a limitation of the perturbation norm in relation to the minimum singular value, by making the following simplifications:

$$1 > \|A\|_2 \cdot \left\|A^{-1}\right\|_2 \cdot \frac{\|P\|_2}{\|A\|_2} = \frac{\|P\|_2}{\sigma_{\min}}; \quad \|P\|_2 < \sigma_{\min} \tag{13}$$

that is, an upper limitation of the disturbance below the minimum singular value of the coefficient matrix. The closer the disturbance norm is to the minimum singular value, the more likely the system is to be affected by the disturbance.

Case II: $(1\text{-}K_A \cdot \|P\|_2/\|A\|_2) \leq 0$, which would involve analogously $\|P\|_2 \geq \sigma_{\min}$, it is not possible in the inequality (11), because the relative error (percentage) cannot take negative values, being calculated on the basis of spectral norms. This observation has logical veracity, because a marked increase of the disturbance above the minimum singular value would mean a reduction of the effect of the disturbance on the system.

Basically, in this paper is admitted σ_{min} as the maximum threshold up to which the perturbation norm for the matrix can affect the stability of the solution, in relation to the value of the conditioning number. Above this threshold, we consider that the impact on the matrix is no longer considered to have a disturbance effect, but only with a structural modification of the coefficient matrix (see the negative part of the graph in Fig. 1).

On the other hand, also from the relation (11) we indicate the minimum threshold value below which, if the perturbation norm decreases, the inequality is no longer valid (becomes irrelevant):

$$\frac{\varepsilon_z}{1+\varepsilon_z} \cdot \sigma_{\min} \leq \|P\|_2 < \sigma_{\min} \tag{14}$$

The tests performed on both well-conditioned matrices and on the Green core matrix confirm this classification of the perturbation norm. For the situation where the disturbances affect the effect vector, according to the inequality (12), the conditioning number as a performance indicator on the inverse problem, appears as a factor of maximum amplification of the error of the effect size.

For the situation where the disturbance affects the matrix of the coefficients, according to the inequality (13) and the classification of the disturbance in the

interval (14), it is possible to follow a maximum amplification of the solution, in close correlation with the value of its norm of P.

Regarding the concept of conditioning number, the preconditioning methods have been outlined, that is to improve the stability of the solutions, through the physical [16–18] or numerical reformulation of the problem.

4 Conditioning and Regularization in Power Flow Case Study

Starting from the correlation between the properties of a numerical method, with the value of the conditioning number associated with a matrix in the system, in the field of electricity, the author of this paper has identified a method of improving the performance of solving nonlinear systems, when these by linearization become weakly conditioned.

Therefore, a configuration of the electric power system is introduced, in Fig. 3, having the active and reactive powers, generated and consumed, defined in nodes.

The voltage and current in node (4), are considered as balancing. Table 2 contains the sizes determined in relative units.

The unknowns to be identified are: the active and reactive power injected into the balancing node; the voltages (amplitude and phase) in the other nodes, the power flow in the network and the power losses in the system.

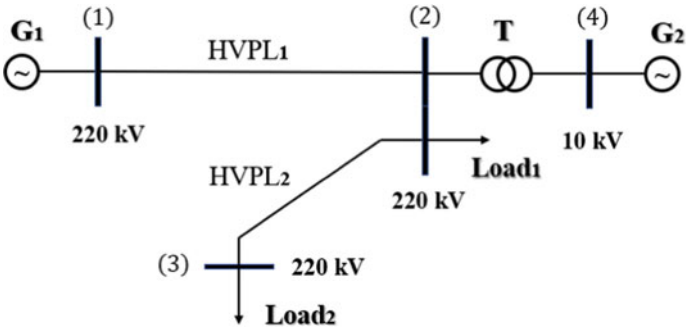


Fig. 3 Configuration of the power system

Table 2 Parameters of network elements

P_1	P_2	P_3	P_4	U_4
0.75	−0.50	−1.15	0.93	1.1
Q_1	Q_2	Q_3	Q_4	δ_4
0.20	−0.20	−0.45	0.44	0

The corresponding mathematical model consists of a system with complex coefficients, reduced to 6 nonlinear equations, deduced from the expression of power balances in nodes. The maximum error of $\varepsilon = 10^{-4}$ u.r, and the initialization of the solution is done with "flat" values [12].

As a method of calculation in the cited reference, it is also proposed to use the classical Newton method, then decoupled, respectively fast decoupled. Basically, a linear system of corrections in the iterative method is repeatedly solved, with Jacobian expressed as a block of 4 submatrices. If in classical Newton, the convergence of the numerical process is high—this is accompanied by a significant computation effort; which is why the decoupled variants give up the matrices on the secondary diagonal, these being of an order of magnitude smaller than those on the main diagonal. We show the expression of the system of corrections equations (15), respectively of the iterative relation (16), after which the solutions of the nonlinear model are evaluated:

$$\begin{bmatrix} J_{p\delta} & J_{pu} \\ J_{q\delta} & J_{qu} \end{bmatrix} \cdot \begin{bmatrix} h_{\delta} \\ h_{u} \end{bmatrix} = -\begin{bmatrix} f_p \\ f_q \end{bmatrix} \tag{15}$$

$$x^k = x^{k-1} + h^{k-1} \tag{16}$$

The indices p, q, u, δ constitute abbreviations coming from the active, reactive power, the amplitude of the voltage and the voltage phase of each node respectively, and their combinations express the dependencies of power-voltage type, specific to the variations that appear in the power systems. What has been particularly noticeable in relation to the reported interpretations, is that the two matrices on the secondary diagonal have high-value conditioning numbers:

$$J_{pu} = \begin{bmatrix} 4.65 & -4.65 & 0 \\ -4.65 & 6.98 & -2.33 \\ 0 & -2.33 & 2.33 \end{bmatrix} \quad J_{q\delta} = \begin{bmatrix} -4.65 & 4.65 & 0 \\ 4.65 & -6.98 & 2.33 \\ 0 & 2.33 & -2.33 \end{bmatrix} \tag{17}$$

$$k(J) = \frac{\sigma_{\max}}{\sigma_{\min}};\ k\left(J_{pu}\right) = k\left(J_{q\delta}\right) = 7.38 \cdot 10^{16} \tag{18}$$

The solution variant addressed in [19], instead of giving up on the poorly conditioned submatrices, applies a local regularization to each iteration. Thus, as a **contribution to the characterization of stability** with the conditioning number indicator, as well as to **obtaining precise and faster convergent solutions** than the decoupled Newton methods, we **propose the partitioned regularization** at each iteration of the nonlinear system.

The regularization method was chosen to be the decomposition by the single truncated values, with the secondary diagonal matrices called abbreviated, SVD:

$$\begin{bmatrix} J_{p\delta} & SVD_{pu} \\ SVD_{q\delta} & J_{qu} \end{bmatrix} \cdot \begin{bmatrix} h_{\delta} \\ h_{u} \end{bmatrix} = -\begin{bmatrix} f_p \\ f_q \end{bmatrix} \tag{19}$$

Table 3 The Newton method decoupled with the Jacobian evaluation only at the first iteration

Iteration number	Maximum error	h_{U_i}	h_{δ_i}	U_i	δ_I (rad)
1	1.15×10^{0}	−0.021127	−0.030726	0.978873	−0.030726
		−0.034672	−0.062337	0.965328	−0.062337
		−0.089921	−0.145974	0.910079	−0.145974
2	8.82×10^{-3}	0.002176	−0.017078	0.981049	−0.047805
		0.000352	−0.016352	0.96568	−0.078689
		−0.003682	−0.031543	0.906397	−0.177517
3	6.18×10^{-3}	0.001749	−0.001851	0.982797	−0.049655
		0.000938	−0.001512	0.966617	−0.080201
		0.000731	−0.004179	0.907128	−0.181696
4	5.11×10^{-3}	0.000203	0.000089	0.983000	−0.049566
		0.000091	0.000167	0.966708	−0.080034
		0.000106	0.000020	0.907235	−0.181676
5	5.41×10^{-4}	−0.000017	0.000041	0.982983	−0.049525
		−0.000017	0.000047	0.966691	−0.079988
		−0.000025	0.000060	0.907209	−0.181616
6	5.26×10^{-5}	–	–	–	–

The results obtained using the partitioned regularization, confirm the relevance of the conditioning number as an indicator of instability, as well as the validity of the regularization method.

Table 3 shows the iterations and solutions when applying the simplified, partitioned regularized Newton method, with the Jacobian calculation only at the first iteration.

As shown in the result table, for the last case, partition regularization, the number of iterations is lower than when applying the decoupled Newton method, so a faster convergence of the solution is obtained. Compared to the classical Newton method, even if more iterations are performed through the partitioned regularization, the computation effort per iteration is significantly lower, as a whole, the method proposed in this paper presents higher performance than the classical one.

By the truncated decomposition after the singular values of the matrices on the secondary diagonal, the values of the conditioning numbers expressed by (19) are significantly reduced.

The method may be of interest in the analysis of larger power systems, where the computational effort and the stability of the solution can be an expensive issue.

In this framework, we propose firstly the use of the conditioning number, as an indicator of the degree of instability of the approximation solution from the forecast, and then the involvement of effective adjustment techniques. The research is open in this area.

5 Intermediate Methods

5.1 Normal Pseudo Solution

In the monograph [9] it is shown that for a linear algebraic system of equations, the minimum error solution, also called normal pseudo solution is obtained by *orthogonalizing the matrix equation*, a process analogous to the minimization of a function of quadratic mean deviation:

$$A^T \cdot A \cdot z = A^T \cdot u \tag{20a}$$

Regardless of the subsequent numerical methods of solving, applied to this system, especially in the case of large arrays, there are rounding errors due to the multiplication of the two matrices.

Another disadvantage is related to the fact that the new matrix $A^T \cdot A$, has a conditioning number $K(A^T \cdot A) = K^2(A)$, aspect deduced from the decomposition by the singular values. The normal pseudo solution is thus to a greater degree likely to be unstable in the disturbances.

It is highlighted by the relationships:

$$A^T \cdot A = \left(U \cdot \Sigma \cdot V^T\right)^T \cdot \left(U \cdot \Sigma \cdot V^T\right) = V \cdot \Sigma^T \cdot U^T \cdot U \cdot \Sigma \cdot V^T = V \cdot \Sigma^2 \cdot V^T \tag{20b}$$

$$\Sigma = diag(\sigma_{\max}, \ldots, \sigma_{\min}) \tag{20c}$$

$$cond(A) = \sigma_{\max}/\sigma_{\min} \tag{20d}$$

$$cond\left(A^T \cdot A\right) = cond\left(V \cdot \Sigma^2 \cdot V^T\right) = cond(V) \cdot cond\left(\Sigma^2\right) \cdot cond\left(V^T\right) = \ldots$$
$$\ldots = cond\left(\Sigma^2\right) = \sigma_{\max}^2/\sigma_{\min}^2 \tag{20e}$$

It is known that the conditioning number of an orthogonal matrix has the unit value [10].

If the core matrix is positively defined symmetric, then the squares of the singular values are the eigenvalues of the normal matrix. This observation suggests that the eigenvalues of the normal matrix must always be positive and real.

5.2 Method of Collocation

The collocation method, as a particular variant of the weighted residue method [6], implies the replacement of the unknown causal vector by a sum of vectors. $f^{<k>}$, which form an orthogonal basis. In the analytical variant of the problem, the unknown

function is approximated by a sum of basic, linearly independent functions [15]. Equations (21) and (22) show the principle of the method:

$$z \approx \sum_{k=1}^{n} c_k \cdot f^{<k>} \tag{21}$$

$$A \cdot F \cdot [c] = u \tag{22}$$

where A^{NxN} is the coefficient matrix; $F \epsilon \, \mathbb{R}^{Nxn}$ matrix formed from the considered base vectors; $[c] \epsilon \, \mathbb{R}^{Nxn}$ the matrix of the weighting coefficients, become the unknown of the new system (20a).

Thus, the problem is reduced to determining a set of coefficients. It is of interest in the numerical evaluation of the method, the conditioning of the new matrix of the system $A \cdot F$, that is, whether the collocation method can be considered as a preconditioning role.

5.3 Tikhonov Regularization Methods

The matrix expression of Tikhonov regularization was already introduced and the mathematical deduction can be achieved starting from the full analytic model, included in Tikhonov functional [9, 10, 20] or from matrix system with requirements of minimal error applied, respectively to solution stabilization.

Depending on the considered order of regularization, the two formulations are equivalent although in the paper [10], first version is named Tikhonov regularization and the second version is named linear degenerated system regularization.

We will concentrate on the second version of the formulation with focus on the followed steps in order to obtain the solution. So, in general case, where the system is undetermined, the orthogonalization method is being applied, in order to arrive at the normal pseudo-solution, which is the first step in regularization.

Further, stabilization matrix $L \epsilon \, \mathbb{R}^{NxN}$ is added (identity matrix, bidiagonal and tridiagonal matrix, the last two corresponding to a discretization of first and second order derivatives), weighted by α parameter, so the solution should be stable to perturbations and also reliable.

$$A^T \cdot A \cdot z + \alpha \cdot L \cdot z = A^T \cdot u + \alpha \cdot L \cdot z \tag{23a}$$

By an inverse matrix procedure $M = (A^T \cdot A + \alpha \cdot L)^{-1}$ the expression becomes:

$$z = M \cdot A^T \cdot u + \alpha \cdot M \cdot L \cdot z \tag{23b}$$

which, afterwards, may be transformed in a recurrent formula (24) or developed in series, by partial overlapping of formula.

$$z^{<k+1>} = M \cdot A^T \cdot u + \alpha \cdot M \cdot L \cdot z^{<k>} \tag{24}$$

The successive approximation formula is repeated until the difference between two consecutive approximations is smaller than an imposed threshold, and the approximation error is also below a threshold.

$$\begin{aligned} \bar{z} &= \left(I + \alpha \cdot M \cdot L + \alpha^2 \cdot M^2 \cdot L^2 + \cdots + \alpha^m \cdot M^m \cdot L^m\right) \cdot M \cdot A^T \cdot u \\ \bar{z} &= \left[(\alpha \cdot M \cdot L)^{m+1} - I\right] \cdot (\alpha \cdot M \cdot L - I)^{-1} \cdot M \cdot A^T \cdot u \end{aligned} \tag{25}$$

The order until the development goes well must take into account the computational effort required to raise the order of matrices, but also the fact that from a certain order, if the value of the regularization parameter is close to zero, the development terms can be neglected.

In all the expressions in which the matrix appears explicitly M, for large matrices it is not recommended to calculate the inverse. We call the relation (24) of iterative Tikhonov regularization, and the approximation relation of the solution $\bar{z}$, (25) of the Tikhonov regularization through series development. Keeping only the first term in the development is equivalent to the classic Tikhonov regularization method.

Another version of Tikhonov successive approximation is expressed by the relations, with the final conditions mentioned in algorithm, but for the Eq. (24):

$$z^{<k+1>} = z^{<k>} + \left(A^T \cdot A + \alpha \cdot L\right)^{-1} \cdot A^T \cdot r^{<k>}; \qquad r^{<k>} = u - A \cdot z^{<k>} \tag{26}$$

There are approaches according to which regularization should be treated iteratively as predictor–corrector [21].

On the conditioning through the Tikhonov regularization, the effect of the regularization parameter is emphasized, for the zero-order regularization:

$$A^T \cdot A + \alpha \cdot I = V \cdot \Sigma \cdot V^T + \alpha \cdot V \cdot I \cdot V^T = V \cdot \left(\Sigma^2 + \alpha \cdot I\right) \cdot V^T \tag{27a}$$

$$cond\left(A^T \cdot A + \alpha \cdot I\right) = cond\left(\Sigma^2 + \alpha \cdot I\right) = \frac{\sigma_{\max}^2 + \alpha}{\sigma_{\min}^2 + \alpha} \leq 1 + \frac{\sigma_{\max}^2}{\alpha} \approx \frac{\sigma_{\max}^2}{\alpha} \tag{27b}$$

Therefore, if the regularization parameter is greater than the minimum singular value, then the conditioning improves. Moreover, it can be considered as the minimum threshold for the regularization parameter, even the minimum singular value of the coefficient matrix A.

Regarding the numerical methods of effectively solving the regularized expressions, they can be any of the ones presented in the first section.

Finally, an original parallel approach is presented between the expression of Tikhonov regularization (26) and the discretized expression of the Fredholm integral equation of the second kind (27b), in the hypothesis of using the same number of discretization points for both variables $N_x = N_y$:

$$(A + \alpha \cdot L) \cdot z = u \tag{28}$$

$$(\lambda \cdot A + I) \cdot z = u \tag{29}$$

From the physical and mathematical point of view, the Fredholm equation of the second kind is considered to be a correctly formulated and well-conditioned problem, if the eigen value λ do not override the determinant of the operating matrix of the system (28) [5, 20, 22].

For the particular case where in the relation (24) is allowed as regulation parameter $\alpha = 1$ and a stabilization matrix $L = I$ and in (25) an own unit value $\lambda = 1$, then the discrete model of the Fredholm equation of the second kind can be considered as a complete limit of regularization, in the sense of implicitly obtaining a stable and achievable solution.

Basically, adding a unit value on the main diagonal of the core matrix A, improves the single values spectrum and thus significantly reduces the value of the conditioning number.

However, we avoid indicating a unit value of the adjustment parameter. Even if the parallel made converges towards a common denominator, the two formulations—Fredholm of the second kind, respectively the Tikhonov regularization—are different. The Fredholm II model appears only when the field problem specifies the sources, and the potentials imposed (or measured) are part of the same portion of the domain, as the causal potentials, unknown.

6 Special Methods of Regularization

6.1 Mixed Tikhonov—TSVD

We continue to present a mixed version of regularization, of type Tikhonov—TSVD. This implies that in the penalty function, the core matrix appears truncated. In this case, the filtering factor included in the diagonal matrix of the singular values, is defined complementarily, by the group of relations (30a–30c):

$$f_i = \sigma_i \ \ if \ \ i \leq k \ \ \vee \ \ 0 \ \ if \ \ i > k \quad \text{or} \tag{30a}$$

$$f_i = \sigma_i \ \ if \ \ i \leq k \ \ \vee \ \ 1 \ \ if \ \ i > k \quad \text{or} \tag{30b}$$

$$f_i = \sigma_i/(\sigma_i + \alpha) \ \ if \ \ i \leq k \ \ \vee \ \ i > k \tag{30c}$$

Thus, the truncated core matrix $A^{(k)}$ will be evaluated by TSVD decomposition:

$$A^{(k)} = U \cdot \begin{bmatrix} f_1 & 0 & \dots & 0 \\ 0 & f_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f_{N_x} \end{bmatrix} \cdot V^T \tag{31}$$

Finally, Tikhonov's expression is shown in Eq. (32), equivalent to (28):

$$\left(A^{(k)^T} \cdot A^{(k)} + \alpha \cdot L\right) \cdot z = A^{(k)^T} \cdot u \tag{32}$$

Based on the arguments presented for the two types of methods, of penalty, respectively of projection, it is considered that this method guarantees a limitation of the excessive amplification of the solution, as well as a framing in the constraint imposed by the stabilization matrix operator L.

6.2 SVD Preconditioning

If from the TSVDM method the truncated matrix of the single vectors to the right is retained, $V^{(k)} \epsilon \, \mathcal{R}^{Nxk}$, we find that its A conjugate application, $(V^T \cdot A \cdot V)$, leads to a reconditioning of the matrix Eq. (1).

The procedure is illustrated in the following sequence of relations, related to Eq. (9):

$$\begin{aligned} & V^{(k)^T} \cdot A \cdot V^{(k)} \cdot V^{(k)^T} \cdot z = V^{(k)^T} \cdot u \\ & M^{(k)} = V^{(k)^T} \cdot A \cdot V^{(k)}; \quad y = V^{(k)^T} \cdot z; \quad b = V^{(k)^T} \cdot u \\ & M \cdot y = b; \quad V^{(k)^T} \cdot z = y \end{aligned} \tag{33}$$

In this form, the equivalent system in (33) is presented as undetermined, which returns us to the initial problem.

If, instead, the matrix of the coefficient's $A \epsilon \, \mathcal{R}^{NxN}$, is replaced by its decomposition by the singular values, and the decomposition matrices are orthonormal between them, the following equivalent formulation of the system is obtained, taking into account the orthogonality property of the single matrix to the left and to the right:

$$\begin{aligned} & V^{(k)^T}[k \times N] \cdot U[N \times N] \cdot \Sigma[N \times N] \\ & \quad \cdot V^T[N \times N] \cdot V^{(k)}[N \times k] \cdot V^{(k)^T}[k \times N] \cdot z[N \times 1] = \dots \\ & \dots = V^{(k)^T}[k \times N] \cdot u[N \times 1] \end{aligned}$$

$$\begin{aligned} & I[k \times N] \cdot \Sigma[N \times N] \cdot I[N \times k] \cdot V^{(k)^T}[k \times N] \cdot z[N \times 1] \\ & \quad = V^{(k)^T}[k \times N] \cdot u[N \times 1] \end{aligned}$$

$$\Sigma[k \times k] \cdot V^{(k)^T} \cdot z = V^{(k)^T} \cdot u$$

which under the writing of the solution z, becomes:

$$z^{(k)} = \sum_{i=1}^{k} \frac{u}{\sigma_i} \cdot V^{<i>(k)} \tag{34}$$

Therefore, a truncated expression of the solution has been reached which depends on: the effect vector, the matrix of the singular values on the right and the first k singular values; and a controlled amplification of the effect.

We also observe the preconditioning/truncation property of the matrix M, which is actually the truncated diagonal matrix of the singular values. The lower the order k of retention of single vectors $\text{cond}(M) = \sigma_1/\sigma_k$ is, the lower the conditional number.

6.3 Conjugate Gradient Regularization Method

Returning to the conjugate gradient method, according to the assessments and demonstration in [23], it is indicated that it is possible to apply the Tikhonov regularization, or DVS. Moreover, we admit that conjugate gradient appears as an alternative to numerically solving a regularized system of equations.

6.4 Algebraic Reconstruction Technique (ART)

We propose a last numerical version, regarding its application to the incorrectly formulated equation systems. In the specialized literature it is called the algebraic reconstruction technique (ART) [24], or Kaczmarz's method and this consists in the sequential iterative solution—for each equation—of the system (1).

In the mentioned work, it is shown that the process of obtaining the solution gains exponential convergence, if the equation of the system is randomly performed, after a sequence of generation proportional to the norms of the lines a_i in the core matrix of the coefficients A.

The algorithm of the method is exposed, in its own version:

$rand(i) \in \{0; N\}$ \\ randomly generating a series of lines of the coefficient matrix, A, with probability distribution $\|a_i\|_2$;

1. $z^{<0>}$ \\ initializing the solution vector;

for $k = 0$

2. $z^{<k+1>} \leftarrow z^{<k>} + \dfrac{u_{rand(i)} - a_{rand(i)} \cdot z^{<k>}}{\left\| a_{rand(i)} \right\|_2} \cdot a_{rand(i)}^{T}$

\\ the sequential iterative evaluation formula of the solution;

3. $k \leftarrow k+1$ *repeat until* $\left\| u - A \cdot z^{<k>} \right\|_2 \le \varepsilon \quad \wedge \quad \left\| z^{<k+1>} - z^{<k>} \right\|_2 \le \delta$

\\ condition of completion of the iterative process;

7 Regularization as a Harmonic Reconstruction of Signals

In the case of each synthetic application detailed in this chapter, graphical representations for the single vectors on the left have been drawn. According to the relation (35), the solution of the inverse problem is given by the sum of these vectors, weighted by a set of coefficients that depend on the singular values.

In most cases, the analysis of these graphical representations reveals an oscillatory variation (around the Ox axis) of the single vectors. For some cores, these oscillations are close to sinusoidal functions of different pulses. Figure 4 demonstrates a

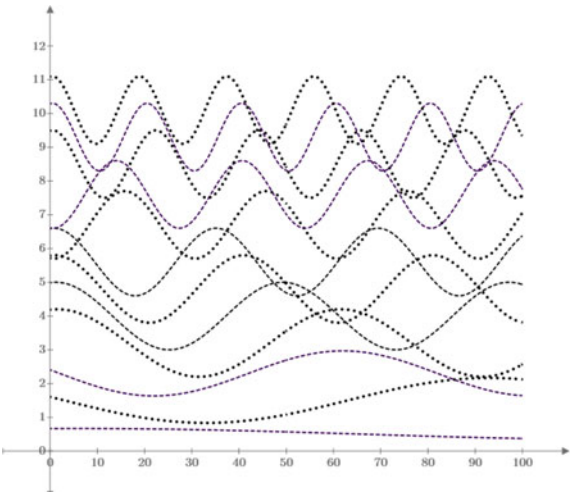


Fig. 4 Oscillatory variation of single vectors corresponding to a core matrix

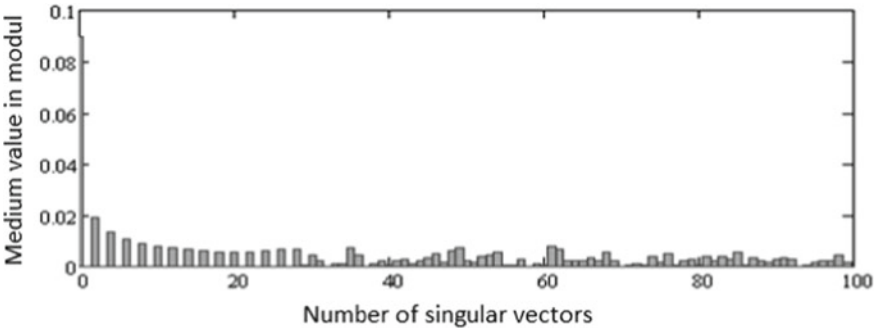


Fig. 5 Harmonic analysis of single vectors

superposition of the single vectors on the left, starting from the first in matrix V, up to the order of 10.

The results of limited experimental character avoid the assertion of a general sinusoidal oscillation property, for the single vectors of any matrix from an inverse problem even ill conditioned. The spectrum of oscillatory vectors is very atypical for each case of core matrix.

Given the conditions, the following common observations are issued to the problems studied:

- each singular vector has an oscillation around the Ox axis; the number of intersections the numerical function of the vector with the axis is equal to the order number of the vector in the matrix of the singular values to the left; this finding is generally valid for about 20% of the vectors constituting the matrix;
- **the average value in the module, of the sum of the elements of each singular vector,** decreases with the increase of the order of the vectors in the matrix; this aspect is illustrated in Fig. 5.

By an analogy with the harmonic analysis of a no sinusoidal signal, we estimate that each numerical function—a singular vector, can be assimilated to a harmonic that enters into the composition of the solution. To argue this statement, we repeat the expression (35):

$$z = \sum_{i=1}^{N_x} \frac{U^{<i>T} \cdot u}{\sigma_i} \cdot V^{<i>} = \beta_1 \cdot V^{<1>} + \beta_2 \cdot V^{<2>} + \cdots + \beta_{N_x} \cdot V^{<N_x>} \quad (35)$$

From this point of view, the solution of an inverse problem can be characterized by an overlapping of effects, **harmonics**, each amplified by the corresponding singular value. If we admit a hierarchy of the importance of the single vectors, in the appropriate reconstruction of the solution, according to the average value indicator in the way of the sum of the elements of the single vectors, then according to Fig. 5, as the

order of the "harmonics" (single vectors) increases, they must decrease their position when reconstructing the solution.

The assertion is supported by the several numerical simulations obtained both in this chapter and in the specific literature, by the TSVD methods.

We consider this approach as being an original one, by analogy with the superposition of effects to non-sinusoidal signals, only a starting point in the development of a new method of regularization.

8 Conclusions

The authors present a special category of engineering problems called inverse in their sense that the effect is known and what is to be found is the cause. Moreover, these problems are niched to the situation when the inverse issue is accompanied by a so called ill-posed definition. This is directly reflected in the stability and numerical precision of the solutions, as a small input perturbation in the effect or in the inverse problem definition it will cause a large variation of the initial result. Thus, starting from these problem definitions in the following sections of the chapter a physical and analytical description of the ill-posed inverse problems is provided, using Fredholm integral equations. Then, several procedures of numerical regularization solving methods are presented which can both apply on the Fredholm discrete model and on direct numerical systems of equations.

Beyond the availability of these numerical regularization procedures for ill-posed problems, the authors offer a step by step example implementation on a power flow case study, which can serve as a reference case.

The comparison between these numerical methods reflect both their advantages and shortcomings that can serve as a guide on what kind of numerical instrument to choose when dealing with different electrical engineering applications.

References

1. Sykulski Jan K (ed) (1995) Pawluk Krystyn, Computational magnetics (field synthesis). Chapman & Hall, London
2. Tikhonov AN, Glasko VB (1965) Use of the regularization method in non-linear problems. USSR Comput Math Math Phys 3:93–107
3. Hansen PC (2008) Regularization tools, a matlab package for analysis and solution of discrete ill posed problems, Denmark. www.imm.dtu.dk/~pch
4. Bronshtein IN et al (2004) Handbook of mathematics. Springer, New-York
5. Lavrentiev M et al (1986) Ill-posed problems of mathematical physics and analysis. American Mathematical Society
6. Sadiku Matthew NO (2000) Numerical techniques in electromagnetics, 2nd edn. CRC Press, New York
7. Palka R (1985) Synthesis of electrical fields by optimization of the shape of region boundaries. IEE Proc 132, Pt. A(1):28–32

8. Morozov VA (1966) Regularization of incorrectly posed problems and the choice of regularization parameter. USSR Comput Math Math Phys 1:242–251
9. Micu A, Micu Dan O (2002) Electromagnetic field synthesis. Ed. Dacia, Cluj-Napoca
10. Micu Dan O (1993) Contributions to the synthesis of the stationary electric and magnetic field. PhD thesis, Cluj-Napoca
11. Michael J (2004) Modular regularization algorithms. PhD thesis, Denmark
12. Ştefan K (2004) Metode numerice, Algoritme, Programe de calcul, Aplicaţii în Energetică, Ediţia a 3-a, Ed. Orizonturi Universitare, Timişoara
13. Hansen PC (1992) Analysis of discrete Ill-posed problems by means of the L-curve. SIAM Rev 34(4):561–580
14. Hansen PC (1992) Numerical tools for analysis and solution of Fredholm integral equations of the first kind. Inverse Probl 8:849–871
15. James Epperson (2002) An Introduction on Numerical Methods and Analysis. Wiley, New York
16. Chadebec O et al (2002) Recent improvements for solving inverse magnetostatic problem applied to thin shells. IEEE Trans Magn 38(3):1005–1009
17. Chadebec O et al (2003) How to well pose a magnetization identification problem. IEEE Trans Magn 39(3):1634–1637
18. Chadebec O et al (2004) Magnetization identification problem: illustration of an effective approach. COMPEL 23(2):518–530
19. Bărbulescu C, Ceclan A, Kilyeni Ş, Micu DD (2007) Power flow calculation for ill-conditioned systems. Simplified Newton method with SVD partial regularization, EUROCON, Varsovia
20. Tikhonov Andrev N (1981) Solution of incorrectly formulated problems. Mir, Wiley, New-York
21. Lamm Patricia K, Scofield Thomas L (2000) Sequential predictor-corrector methods for the variable regularization of Volterra inverse problems. Inverse Probl 16:373–399
22. Lavrentiev M (1967) Some improperly posed problems of mathematical physics. Springer, New-York
23. Gottvald A (1997) A survey of inverse methodologies, meta-evolutionary optimization and Bayesian statistics: applications to in vivo MRS. Int J Appl Electromagnet Mech 1:17–44
24. Strohmer T, Vershynin R (2006) A randomized solver for linear systems with exponential convergence. Random-Approx 4110:499–507

Advanced Energy Systems

Advanced Energy Systems Based on Energy Hub Concept

Hossein Shayeghi and Amir Mohammad Moghaddam

Abstract Energy systems in the past alone were providing consumer's needs, and their management did independently. By industrialization of the lifestyle and high dependence of human societies into energy and on the other hand, a lack of fossil resources, system exploitation seeks to use multi-carrier energy to meet the needs of their consumer's needs. Therefore, the use of multi-carrier systems with the advancement of technology became possible to increase the reliability of the system in the presence of different energy sources. Of course, the critical and challenging issue in it is the way of managing these systems on a large scale. The concept of an energy hub, which is like a virtual box including production, conversion, storage, and consumption of various carriers, has been introduced and the high potential for the operation of multi-carrier systems and optimal management of them possess. This chapter aims to introduce a comprehensive energy hub and its application in the management of multi-carrier systems and express its advantages in preserving the existing resources for posterity alongside the increasing demand for consumers with the lowest cost of operation without limitation its size. Energy hub is used in various sectors such as commercial, industrial, and even agricultural, and can integrate with the integration of these small network hubs from Micro Hub, which are called Macro hub. Moreover, the management of these hubs has dealt with detail.

Keywords Multi-energy systems · Energy hub · Advanced energy system · Micro hub · Macro hub

H. Shayeghi (✉) · A. M. Moghaddam
Energy Management Research Center, University of Mohaghegh Ardabili, Ardabil, Iran
e-mail: hshayeghi@gmail.com

A. M. Moghaddam
e-mail: ammoghaddam1389@gmail.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_10

Nomenclature

Acronyms

BEMS	Building energy management system
BTU	British thermal units
CHG	Closed green house
CHP	Combine heat and power
DER	Distributed energy resources
DG	Distributed generation
DSM	Demand-side management
EH	Energy hub
ESS	Electric storage system
HEMS	Home energy management system
HVAC	High voltage alternative current
IEMS	Industrial energy management systems
MES	Multi-energy systems
PV	Photovoltaic
OEF	Optimal energy flow
OPF	Optimal power flow
TOU	Time of use
V2G	Vehicle to grid

1 Introduction

The vital energy of the current industrial societies is an essential factor in the development *of countries.*

Besides the advancement of technology and machine life style, this needs to be more intuitive and more important in the day-to-day. Consequently, producing clean, safe, and reliable energy sources has been one of the severe difficulties of every country.

Energy is the most fundamental requirement of the current industrial societies and an essential factor in the development of countries. The most fundamental sources of energy supply during recent years were fossil fuels. The conversion of fossil fuels produces electrical energy in thermal power plants, and with low efficiencies and relatively high losses are transported to consumers by transmission lines. In these systems, intelligent protection, control, and data gathering tools are installed and used locally.

However, now, the power grid uses smart grid technology to be smart and operate automatically. In addition to traditional network capabilities (production, transmission, and distribution), the smart grid can store, interact, and make decisions.

The smart grid increased efficiency in the supply of demand and economic efficiency of existing equipment and no need to improve systems. Also, the smart grid has caused the integrity of renewable energy resources and mainly distributed energy resources on demand-side.

With the advancement of technology and increasing the efficiency of prime movers as well as improving market conditions for the use of high-performance technologies such as micro-turbines and internal combustion engines, the benefits of the distributed energy resource (DER) is used to form CO-Tri-generation. These compounds produced the hopes of making multiple clean energy systems (MES) with high affections. With the help of this equipment and DER, management and operating of such systems can optimize for economic, technological, and environmental benefits.

Using the DER on demand-side, in addition to reducing losses and costs due to energy transfer, increases the security and reliability of supply of local loads and also contributes to the active participation of consumers in demand-side management.

Demand-side management (DSM) is a comprehensive theory that combines load improvement plans, energy conversion, energy efficiency, and demand response.

Renewable energy source (RES), which is considered the most essential DER, is the ability to supply consumer demand during peak energy hours that can be reduced consumer dependence on the leading network.

Transfers of the consumer load of the peak-load into low-load reduce their energy costs and reduce network demand size. On the other hand, reducing unnecessary loads in a certain period, such as the peak demand period or the system's use demand, can be stored more energy and decreased costs.

Although to achieve the benefits of the said concept, which do with the help of DSM, the equipment used in the network and smart grid requires an integrated management model. The energy hub defined as a structure that generates, converts, store, and consume various carriers in it. Energy hub has a high ability to manage and control systems with several centralized inputs.

An extensive study has been conducted in the concept of energy hub and has proven that the use of several energy carriers to provide consumers needs better performance than systems that independently supply consumer's needs and control.

The energy hub has no size limit and can include many cases such as a residential house, spacious residential complex, commercial buildings, shopping malls, hospitals, industrial units, greenhouses, and even the whole city.

So, after introducing the overall energy hub plan, in part 2, the comprehensive energy hub with the introduction of its elements, the advantages of the energy hub for the whole system is introduced. In Sect. 3, the various types of energy hub classification have discussed, and it has described the residential, commercial, industrial, and agricultural hubs, and in the last section, the overall classification of the topics has been studied.

2 Concept of Energy Hub

For the first time, the energy hub theory introduced in the connection of a project called A Vision of Future Energy Networks (VOFEN) [1]. This task intended to describe future energy systems in the long term (20–30 years).

Many conceptual designs have published concerning the production, transfer, and distribution of integrated energy by using DER, which created the concept of an integrated energy hub. An integrated word means the use of various energy carriers in the system. This concept has shown in Fig. 1.

$$\begin{bmatrix} I_1 \\ I_2 \\ \ominus \\ I_3 \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} & \ominus & C_{1n} \\ C_{21} & C_{22} & \ominus & C2n \\ C31 & C32 & \ominus & C3n \\ C41 & C42 & \ominus & C4n \end{bmatrix} = \begin{bmatrix} L_1 \\ L_2 \\ \ominus \\ L_n \end{bmatrix} \tag{1}$$

As you can see in Fig. 1, in the matrix model of the hub, the input matrix is defined as a column matrix, each element of this matrix corresponds to an independent input. By multiplying the input matrix in the coupling matrix, which is a square matrix and each element is related to the elements used in the hub, the output matrix will be achieved (Eq. (1)) [2].

2.1 Basic Concept

The basic definition of energy hub is an integrated platform of energy producers, consumers and transmission equipment. Each system that has the following characteristics can call energy hub [3].

- Input and output
- Conversion
- Storage

The hybrid energy carrier system considers as a system of energy hubs that interconnected by several networks. In the realm of an energy hub, the energy transmitted by network and stored or converted into the energy hub to meet the demands. A more perspicuous definition of an energy hub presents in [1].

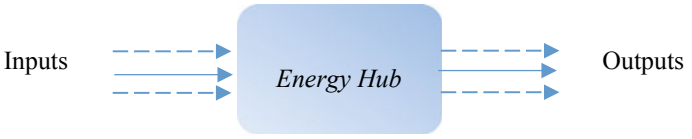


Fig. 1 Energy Hub concept

2.2 Elements of Hub

The main characteristic of the energy hub has three main elements:

- Direct connections
- Converters
- Storage

Direct connection is used to transmit a direct incoming carrier to the hub output without any conversion. (e.g., electricity voltage).

The upstream grid is an example of this type of component. Besides that, converter relates to devices that change the sort of energy carriers. For example, in the boiler, gas is as input and produces heat in its output.

The third element used in the hub refers to energy storage equipment, which can have various technologies. For example, electrical storage has a direct storage technology (e.g., supercaps, superconducting devices) or indirect (e.g., batteries). [3].

In Fig. 2, the complete model of the energy hub is shown with electrical and gas as inputs; and electrical, cooling, and heat as outputs. This model uses converters and storage, such as transformers, boilers, electrical/thermal storage inside the hub, for meeting the demand in output.

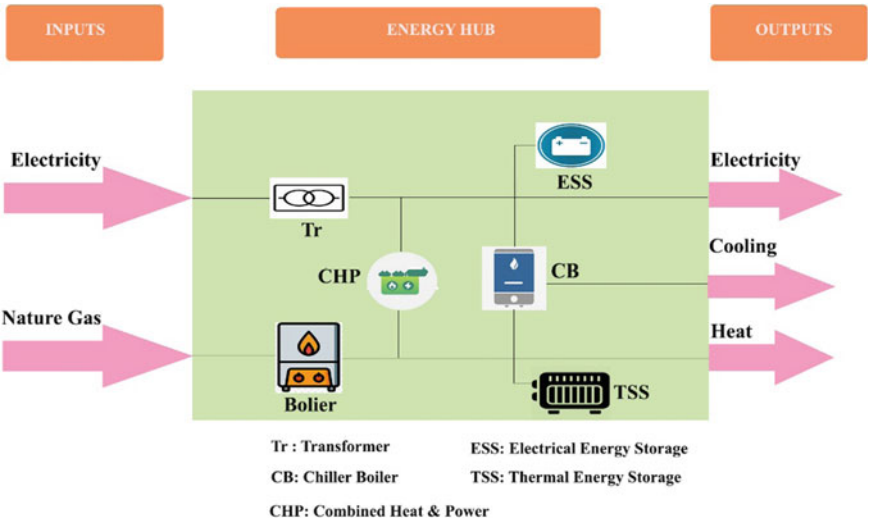


Fig. 2 The energy hub with different elements

2.3 Potential Benefits

According to definitions of an energy hub, a simple structure represented in Fig. 1. In this structure, each energy demand can meet by multiple energy carriers. In simple words, it means the energy hub is not only used one carrier to demand supply but also a combination of two or more carriers is used to supply demand, and this characteristic improves the degree of freedom and also increase reliability and security of consumer's demand. Using different carriers and various combinations of them can be used to optimize consumer's demands optimally.

Implementation of the energy hub can reduce operating costs because energy hubs choose the most energy-efficient carriers, or a combination of them to increase efficiency.

Other potential benefits of the energy hub can be improved load flexibility. Redundant ways inside the hub offer a specific degree of freedom in supplying the loads.

The hub can replace for an unattractive energy carrier, for example, high-tariff electricity. So, from a system view, the load looks to be more flexible in terms of its price and demand performance; also, the actual load on the hub output remains constant [3]. On the other hand, the use of various energy carriers to provide different demands can help expandtion the application of the energy hub and improve system performance.

For example, the use of energy that wasted in various carriers due to the structure of converters and other equipment can increase the system productivity. Therefore, one of the main advantages of an energy hub is the optimal use of multiple carriers to provide optimal demand.

3 Different Types of Energy Hubs

The energy hub has a high potential for increasing system productivity by taking different carriers and using various technologies. The energy hub divided into two parts of micro hubs and Macro hubs in terms of control. Micro hubs are located within the Macro hubs and controlled by them.

The micro hubs consist of four sections, tailored to the consumption pattern and the amount of demand, including residential, commercial, industrial, and agricultural. [4]. The energy hub categories are shown in Fig. 3.

Integrated management of these sections can have many benefits for both their own and the whole system. For example, the consumption pattern is different in the residential and industrial sectors. The residential peak demand occurs during the early night hours, whereas the industrial sector consumption is almost constant.

The use of DSM, especially the demand response (DR) programs and the RES, is of the great potential that can provide to the residential sector in addition to providing the need for their needs. The relation between these two parts usually established by

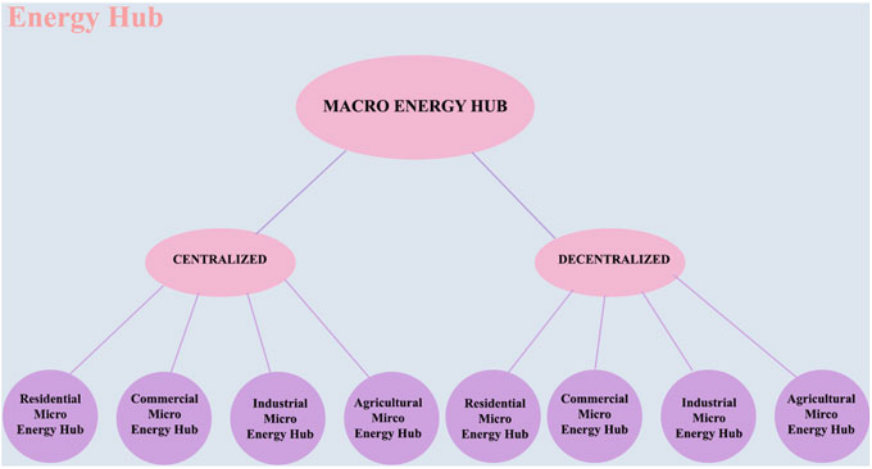


Fig. 3 The energy hub categories

electrical and gas. In the residential sector, it can also help consumers to meet the demands of the wind turbine and can be stored in the storage device and transferred to the industrial sector with the use of the RES, such as air turbines and PV throughout the day.

There are many methods for integrated management of multiple systems, but using the concept of energy hub is one of the best methods. With this method, the overall system optimization rate increases significantly [4].

3.1 Micro Energy Hubs

As mentioned earlier, the micro-hub consisted of four public groups. In this section, we will introduce and study the features and challenges facing these four groups. We also discuss the impact of energy hubs on solving these challenges.

(A) **Residential**

According to statistics published in Energy Information Administration (EIA) in America in the year 2018, 40% of the energy production has consumed in the residential sector [5].

For lots of reasons, including large transmission networks, extensive distribution, and poor management of power consumption by consumers, low efficient appliances, the energy in the residential sector are matched with its demand and consumption. The losses of electricity-related are 47% of this consumption [5]. Therefore, the residential energy hub is forced to use various carriers and efficient suppliers to meet the needs of consumers and reduce losses and operation costs.

In recent years, the use of multi-energy systems in the residential hub has increased dramatically, in which at least two different carriers are being used to meet consumer demand. Generally, two carriers used for the supply of demand that extensively utilized in the articles are electricity and gas. Energy hub is a suitable option for energy management in the residential sector [4].

There are lots of researches that show using an energy hub concept in them. An example of the energy hub model for smart homes with uncertainty in market pricing can find in [6]. In this study, a smart apartment building considered with ten homes, and each home has 12 appliances. The objective function considered as minimizing operation costs, and the results provided an optimum dispatch of equipment and timing of appliances have shown.

In [7] a multicarrier hub has been investigated to reduce the economic cost and CO2 emissions in a residential building with the influence of renewable resources. The system has been used electrical grid and natural gas network, a gas boiler, a heat pump, a photovoltaic plant, and a photovoltaic/thermal (PV/T) system, and has also been used an electrical storage to increase the efficiency.

According to the resource constraints, like scarcity of fossil fuel resources, uncontrolled increase in demand, expansion of environmental concerns and market deregulation, energy price in the residential sector is increasing.

In the traditional methods, demand responding is made directly by adjusting production schedules, But the organization and development of giant thermal power plants to meet the peak demands, besides the high costs of operation, environmental pollution problems, and deep affection, is not a reasonable approach anymore [10].

In the residential sector implementing the demand response consider as an option for energy management in buildings. Demand response has gained some penetration within the electricity market and is represented as "intended electricity consumption pattern modifications by end-use customers that are intended to alter the timing, level of instant demand, or total electricity consumption" according to it [8].

Implementation of demand-side management programs in the system is facing problems such as the lack of proper infrastructure of electrical network transmission equipment [9] and incentives for active participation of consumers in these programs, which has caused no variability or a profound change in consumer demand. Since billing based on a monthly average rate aggregated over a large number of customers [10]. A reflection of the price fluctuations of the power pools as long as the bills determined by the average monthly rate of a large number of consumers, it cannot be a benefit for residential consumers to change their energy demand in peak hours.

Several solutions like time of use, real-time pricing incremental block rates, critical peak pricing, and demand charges, have been introduced to solve the pricing issue. Of course, energy prices are not the only effective factor in the demand-side management program; i.e., consumers in the residential sector do not feel comfortable to spend time to analyze consumption decisions and managing household devices to save money [11]. According to the definition, as shown in Fig. 4, the analytical model of the residential energy hub is presented in two parts.

Residential consumers do not want to spend time to analyze consumption decisions. Therefore, the existence of an integrated management system that makes these

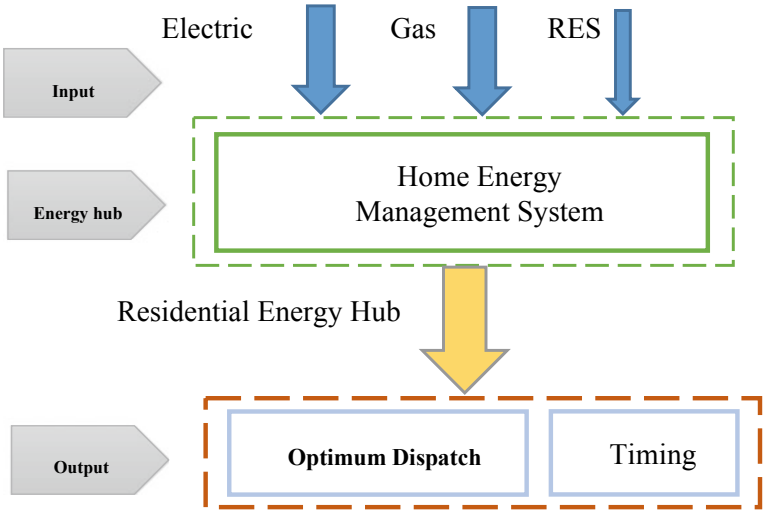


Fig. 4 Analytical model of residential energy hub

tasks feel felt. Recently, a system introduced as a home energy management system (HEMS) that performs all necessary measures. The HEMS uses a variety of parameters to prepare an optimized and scheduled energy consumption plan. These parameters include the price of energy in the market, climatic conditions, temperature. With using HEMS, it is possible to import various objective functions such as cost, emissions, load shaping, etc., which provided to the system. In [11] challenges, the methods of modeling and optimization of various types of HEMS are presented.

Using HEMS makes it easy to combine with multiple carrier systems and smart grids and improves performance and efficiency in the home energy system. Also, using HEMS facilities, home participation is increased in the DR program. Thus it can be a smart home as a small model of smart systems that are looking for optimal management of energy consumption and active participation in demand-side management (DSM) programs [12].

An intelligent decision-making system based on mathematical models used by a part of the Home energy management system for optimal management of energy consumption in buildings, and also makes successful participation in the program Of DSM [13]. In addition to determining the final level of the comfort and priority of consumers, this model has also used environmental information such as energy prices and weather forecasts. The interesting thing about this model is its actions to a real home in Ontario Canada, which has been energy saving as much as 45% and a 15% reduction in consumer peak demand. Instead of engaging local loads in demand-side management programs, it can be possible to control residential sector loads.

As shown in Fig. 5, residential loads are divided into two categories: responsive loads and non-responsive loads. Their responsive loads divided into four subgroups,

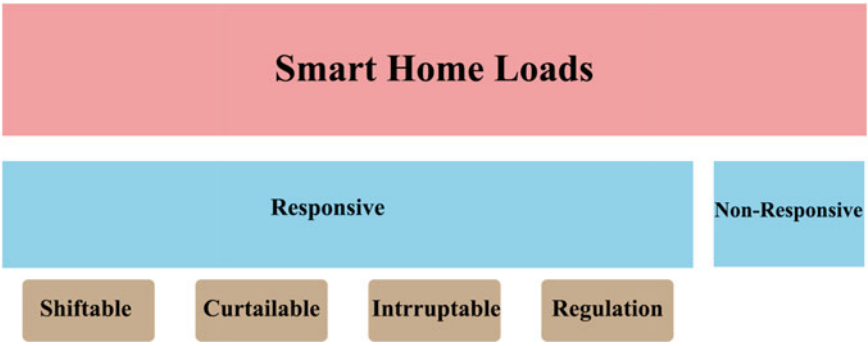


Fig. 5 Classification of the smart home loads

which include: shiftable loads, curtailable loads, interruptible loads, regulating loads. Shiftable loads are items where their energy consumption consumed in a specific timeframe and the opportunity of transferring this consumption at another time. (such as washing machines and dishwashers). Curtailable loads include a load that is possible to reduce the amount of consumption during peak hours. (such as dimming the lights) Interruptible loads are a bunch of loads that can be interrupted and have postponed their use to other times. (e.g. electric cars charge).

The regulating loads refer to the loads that allow programming at operating intervals with a slight change in the defined level. (such as heating and cooling demands). The effect of DR programs on both responsive and non-responsive loads has studied on [14]. In this paper, a residential building that includes CHP and PHEV represented as an energy hub. The energy hub model has investigated for optimal operation with the objective function of minimizing the cost of consumer's payment by taking TOU (Time of Use).

In the high energy price periods, a part of the demand is supplied by CHP with lower operating costs, while in the low energy price periods, the electricity supplied from the primary grid and PEV charged during this period.

Instead of providing the specified production schedule, it is done by changing the amount of shiftable load from peak demand to low-demand at the cost of economic savings consumers.

To take advantage of the benefits of the RES at smart home, the use of PEV and ESS for this purpose has a high potential. The PEV can also play the role of the ESS that helps reduce the cost of system operation, can also improve the energy exchange between PEV and the grid (V2G) by optimal management of the charging time and discharge of PEV. Integration of RES in a smart home requires a coordinated and comprehensive control system to be associated with the rest of the equipment especially PEV, ESS and responsive loads.

(B) **Commercial**

With the rise of commercial and residential buildings, the demand for electrical energy increased, which leads to an increase in the price of this patrol energy. On the other hand, greenhouse gas emissions increased with the activities of these centers. In these conditions, the existence of optimization and energy consumption management system profoundly felt and caused extensive research in this field.

Commercial buildings are a symbol of economic progress. Large commercial towers are competing in developed and developing countries. It should note that these buildings are great energy consumers and they should pay special attention to the energy efficiency issue in designing, building and operating them.

In traditional energy systems, electrical demand generally supplied by the high-power network of electrical energy and the thermal need provided by a gas boiler. However, with the increasing in the price of electric power and the emergence of new technologies in the field of distributed production resources, it was attempted to provide the electrical demand in place and by DG.

Energy production in the place of commercial buildings rose. As mentioned in the previous section, multi-energy systems (MEH) and RES systems are the perfect options for distributed generation (DG) in commercial buildings.

The main point in such an optimization algorithm is the analysis of system behavior and its evolution with more precision in a distance that leads to the supply of demand with optimal production and earning maximum profit.

Increasing the efficiency of commercial buildings is the key to energy-efficient consumption and reduction of environmental impact.

Many parameters can have a return on the energy consumption of buildings. Also, many methods help improvement of them. The full category of these different methods can found at [15].

In this paper, various methods are presented to improve the efficiency of commercial buildings, which are divided into three groups: technological effects, the culture and behavioral training of the occupier, and establishment of building energy management system (BEMS). In [16], the bi-level risk aversion function is considered for a commercial macro hub, which minimizes the load of the network and the economic cost and gas emissions of the greenhouse in a fixed amount.

In commercial buildings, the need for a comprehensive energy management system to determine the optimal timing, active participation in DSM programs, and the benefits of a smart grid feel.

A BEMS introduces for optimal operation of a commercial building to minimize the cost of energy, increase efficiency, and to reduce emission in [17]. According to the obtained results, it has shown that combining different technologies and integrated energy management systems of this building, the building's peak demand and energy costs reduced while the lower GHG emission obtained.

Therefore, BEMS can manage by information such as forecasting weather, production rate, and pattern of consumption to cause optimal commercial building performance. The BEMS used in the analytical model of commercial energy hub has shown in Fig. 6.

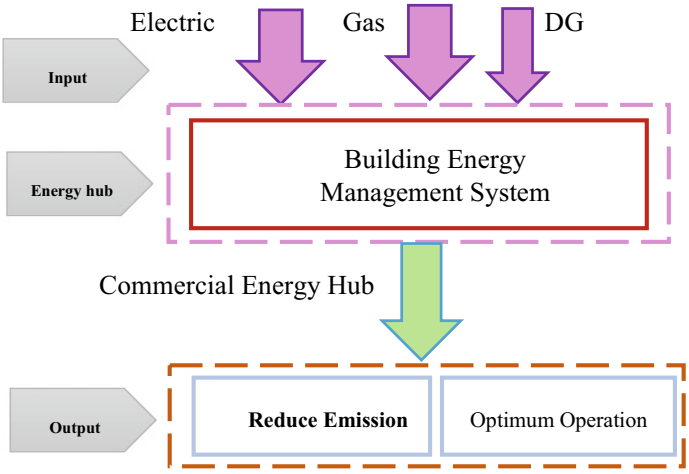


Fig. 6 Analytical model of commercial energy hub

The result is integrated management, cost reduction, peak shaving, reduction of environmental impact, and participation in DR programs.

BEMS correct operation performed with the impact of three items:

- Forecasting the load and weather
- Using different energy carrier
- Use of RESS and ESS

The most important advantage of commercial loads is the predictability of their consumption curve.

In residential loads, the consumption curve changed hourly and quickly, but changes in the consumption curve of commercial loads replaced daily weekly or even seasonal.

For example, in an office building, the consumption curve is divided into two parts of weekdays and weekends, which are followed by a fixed shape in each group. This feature of commercial loads has a high potential for managing demand and participation in DR programs. dependence on climate prediction can make vision changes during seasons changes.

In the commercial sector, the most impact on the load rate is thermal loads because most heating equipment will generally consume more than other energy equipment [18, 19]. The maximum demand for HVAC has been conclusively associated with climatic conditions and has also shown evidence that is responsible for the increasing demand for HVAC [20].

With the influence of distributed energy resources, it is necessary to obtain weather information for optimization algorithms due to the continuous output of this equipment to weather conditions.

In summary, due to the high cost of energy, environmental safety precautions, and the need to use high energy efficiency equipment in commercial buildings, different

methods for reducing costs and emission in this sector Have been used. Among these methods, the use of an integrated energy management system such as BEMS has taken more attention. Finally, the modeling of a BEMS will be available in a commercial section, taking into account the electrical energy Storage System (ESS), MES, RES, DSM, and uncertainty in the form of an energy hub.

(C) Industrial

The biggest energy consumer in the world is industrial sector. By changing the lifestyle and tendency to industrialization, increasing industrial countries, and adding energy consumption in developing countries, the amount of energy demand in the industrial sector has significantly increased.

Increasing demand in the industrial sector is high because of growing the population, increasing the demand for consumer goods, and also increased the use of electric vehicles.

World industrial sector energy usage increases by more than 30% from 2018 to 2050, reaching about 315 quadrillion British thermal units (Btu) by 2050 [5].

The primary solution for increasing energy efficiency and reducing emission in the industrial sector is the use of heat generated in the form of waste losses and energy auditing in different parts of consumption. Energy auditing is in the form of series power consumption. By specifying the consumption and allocation of the energy index to each of these sectors, energy auditing can plan for optimal consumption and energy storage opportunities.

For example, technology and methods used to increase energy efficiency and reduce emissions in the Chinese power plant. Alternatively, reducing the GHG release at the Thai ceramic production plant, exploration of energy-saving potential in the China power industry has been studied.

In addition to the mentioned methods, there are many methods to improve the efficiency of the industry sector. Methods such as monitoring consumption, proper and incentive policies, using high-efficiency equipment, improving the environmental conditions of factories and such as these.

Basically, the efficiency improvement methods for the industrial sector are divided into three categories: technical and technological conditions, effective policies and behavioral training, industrial energy management systems (IEMS). The first case includes the energy audit, the use of techniques and methods such as variable speed control for electric motors, waste heat reuse, Co-Tri generation, use of high-efficiency equipment. The next case includes incentives for the governments for industry owners to improve the efficiency, laws, and preventive measures such as obtaining taxes for carbon emissions, encouraging the use of the RES in industries, recruitment Interns, and improving the productive culture of the workplace. For more details on this subject, refer to [21]. The analytical model of this method is depicted in Fig. 7.

In [22], the concept of energy hub has been used to supply electrical, thermal, and water demand, a cement factory. The energy hub is optimized by the genetic algorithm (GA to provide demand and reduce cost and pollution, increasing efficiency. Results

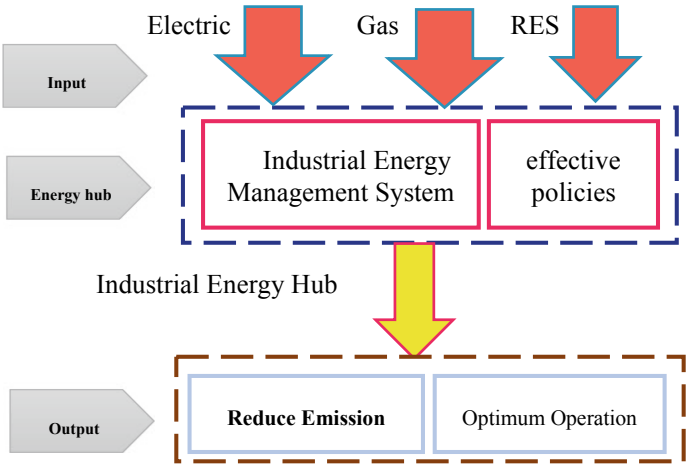


Fig. 7 Analytical model of industrial energy hub

show that the total annual cost of about 36,000 $/year decreased and exergy efficiency increased by a rate of 36%, as well as CO2 emissions decreased to 21,620 tons/year.

Industrial Energy Management System (IEMS) is integrated management to supply the amount of consumption. The primary purpose of this system is to reduce the cost of operation, reduce emissions in a way that does not enter into the system performance. Many technologies have been used to deploy in IEMS. A general overview of the management of various energy systems and their positive impact on energy storage in different residential and industrial sectors in [23] performed.

According to the findings obtained from previous research, most methods of reducing the cost of operation and energy storage related to techniques such as MES, DSM, and the influence of the RES. Therefore, in the industry, to benefit from DER, DSM, MES, and getting more energy efficiency, it requires IEMS to manage production, transfer, storage, and consumption.

The most prominent characteristic of industrial loads is its predictability, which is considered a decisive point in terms of planning and management. Besides, the possibility of using the RES in this section is more comfortable than the other parts.

Also, the cost of setting up and investing the RES in the industrial sector is higher due to the need for greater power.

Generally, the preferences of industry owners have cost to environmental safety precautions. Therefore, government policies to encourage industry owners to adhere to environmental safety precautions and on the considering of penalties for emission can have a significant impact on the use of renewable resources.

Accordingly, the effect of protectionist policies for the optimal combination of power system for an industrial unit in Italy in [24] has studied. The results show that the use of PV system without government supportive and incentive policies due to its high cost is almost impossible. Also, the use of wasted heat for thermal and even

electrical demand by CHP and other RES in the form of DSM and especially the DR programs can bring many advantages to the entire system.

The use of DR programs in the industrial sector is facing challenges. The reason for this is the difficulty of using various electrical, gas, water, raw materials in the industrial area. Moreover, the balance issue is critical and requires special attention. Therefore, this section requires an optimal management system to make technology and different energy requirements in the most efficient way possible.

In [25] a MILP model has been studied for the management of industry demand to participate in the DR program in spite of responsive loads and participation in day-ahead electricity markets. Considering various objective functions such as minimizing cost, increasing profits, reducing losses, and environmental impact can be an optimal operation of industrial units.

Unlike the high potential of the industrial sector for optimal operation and management in the form of industrial energy hubs, very little research has conducted in this area. In [26] a comprehensive model of the optimal management of industrial energy hubs in the form of smart grid was presented. Mathematical modeling of this system involves minimizing the cost of energy. The model has applied in flour mill, and a water pumping facility, which showed that the model presented in combination with control tools and communication better performance in the form of industrial energy hubs in connection with the Smart network.

Due to the need for energy-saving and improving the efficiency of equipment in the industrial sector, an IEMS is needed to benefit from the different methods of improving energy efficiency.

The combination of smart grid with IEMS, considering the demand schedule and uncertainties, the use of DER, and many other technologies can be provided in the platform of an industrial energy hub model.

(D) Agricultural

The agriculture sector is the most crucial productive part of any country. The agricultural sector is in direct contact with people's food and must pay special attention to this section for food security.

The use of primary energy in the agriculture sector is less than in other sectors. Energy consumption in the agriculture sector has increased due to the need for food, lack of mechanized system, the lack of technology improvement.

To achieve sustainable agriculture, the challenges of energy consumption and efficiency, production cost, environmental impact will resolve.

Many methods to address these problems have been introduced to make them easier to check in two categories that include improving the traditional system performance and using alternative systems such as Greenhouses.

In each of the two categories, we can improve the performance of the agriculture sector by using different methods. Many studies show high losses in the energy consumption efficiency of the agricultural sector.

The energy carriers used in agriculture are mainly electricity and gases that use in both direct and indirect forms. Direct consumption of energy carriers mainly uses

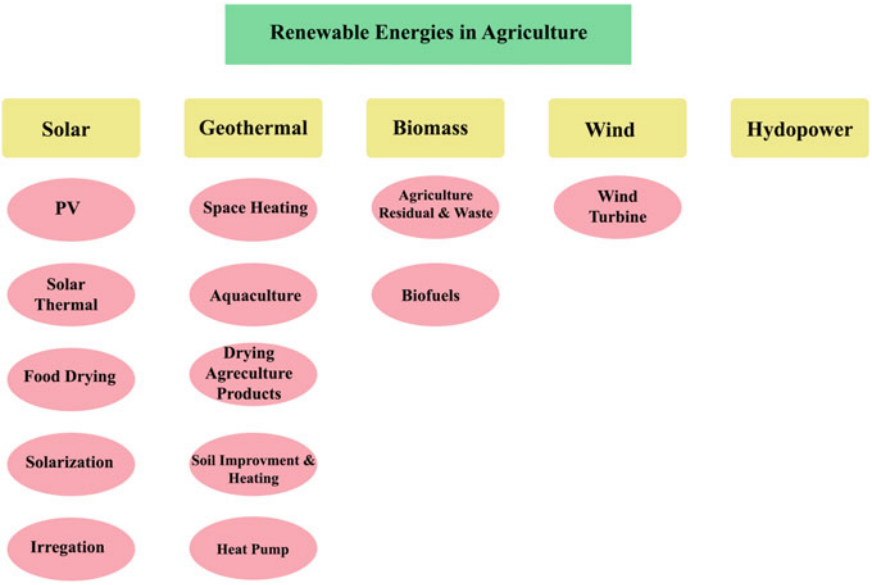


Fig. 8 Application of RES in the agricultural sector

for the energy supply of pumps and fuel for agricultural machinery. Besides these, it makes uninterrupted use of these carriers for heating and ventilation of greenhouses, livestock, and chickens. Indirect consumption of energy carriers also includes the use of carriers in the process of production of consumer and intermediary institutions required by the agriculture sector.

Energy efficiency and product affordability are dependent on these inputs. Therefore, the optimization of quantity and combination of these methods has a significant impact on improving productivity. Solar, wind, biomass, geothermal, which are renewable sources used in the agricultural sector is shown in Fig. 8 along with the application of each part of the sector.

In [27], the use of the particle swarm optimization (PSO) algorithm has been investigated to achieve an optimal combination of energy hub components and optimal scheduling of natural gas, wood chips, biomass, and electrical energy. Its objective function is to reduce the operation cost and emission of CO2. The results showed that the PSO algorithm has efficiency for optimum scheduling. Also, the gas turbine is superior to the biomass production unit in reducing the cost of operation.

According to most of the studies in the field of agricultural energy hub, the best use of fossil fuels to supply electrical needs with the possibility of using renewable energies and fuels Renewable, such as biogas, bioethanol and biodiesel, which can be used in a CHP to supply electrical demand, is consumed in this section.

Also, considering that the agricultural lands are in remote areas, using the RES can be reduced on the transfer and distribution of electric energy costs to these areas.

The largest application of renewable resources are in agricultural units, which will cause many benefits, such as reducing energy costs, reducing environmental impact, and thereby producing green.

Solar energy in the agricultural sector can be used to produce electricity, heat, cold, and drying of agricultural products. Biomass can apply for the energy supply of CHP or CCHP.

The geothermal energy can be used in agricultural works, including heating greenhouses, adjusting the temperature of agricultural products, drying agricultural products, and also in heat pumps for thermal and cooling demand. In addition to traditional windmills, the wind turbine can use for providing electrical demand without the cost of transmission and distribution. For more information on the application of the RES in agriculture, see [28] for more details.

In addition to the use of the RES recently, smart methods have attracted much attention. These quick methods include smart metering, intelligent networks with improvements in monitoring, control, and measurement, improves efficiency, sustainability, reliability, and optimal distribution of energy resources. Smart technologies are not limited to smart networks but also helps create a new structure of Smart Farms and Smart agricultures. Therefore, information gathering and communication technologies can be used in the agricultural sector to monitor and control the physical information of Farms and products.

The various models of wireless communication technologies and their applications in different agricultural sectors can find in [29]. According to the findings, there are no comprehensive models for optimal management of Smart Farms Energy management and Energy exchange between adjacent systems.

Based on the ability of energy hub in integrating various energy systems, with the presence of various energy sources, there is a great potential for the use of various energy carriers in smart farms, despite the ability to integrate these resources into the energy hub, which makes it possible to use smart technologies in the form of smart agricultural energy hubs.

Greenhouses that are known as small-scale Farms have good potential in the production of agricultural products by controlling the conditions and approaching it to the optimal state. On the other hand, the greenhouse has many advantages, such as the possibility of multi-product harvesting during the year, less impact of climatic conditions, integrated management, and better cost control. Based on high costs allocated to this section, poor management of infrastructure and operation of greenhouses has reduced the efficiency of resources used.

For the management and control of greenhouses, various objectives such as reducing costs, reducing energy consumption, reducing environmental impact can be considered. The methods of improving efficiency in greenhouses have been proposed in [30]. With closed greenhouse system the maximum improvement of greenhouse performance achieves.

Keeping the CO2 and humidity at optimal levels is challenging and has a high energy loss. One of the best solutions for this problem is the use of the Closed Green House system, which has no connection with the outside. CGH has a high potential for the use of renewable resources and is also independent of climatic conditions.

The use of the CGH concept in the application of RES and integrated management can improve sustainable agriculture anywhere in the world, and in addition to the production of green products, to provide the demand for nearby buildings. In this study, the CGH shows to increase production and improve in addition to efficient use of energy, as well as the use of RES to reduce environmental impact and green manufacturing. The concept of this method has shown in Fig. 9.

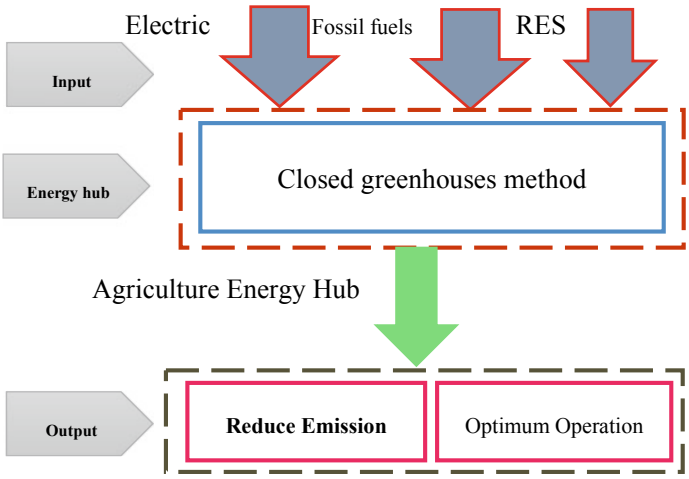


Fig. 9 Analytical model of agriculture energy hub

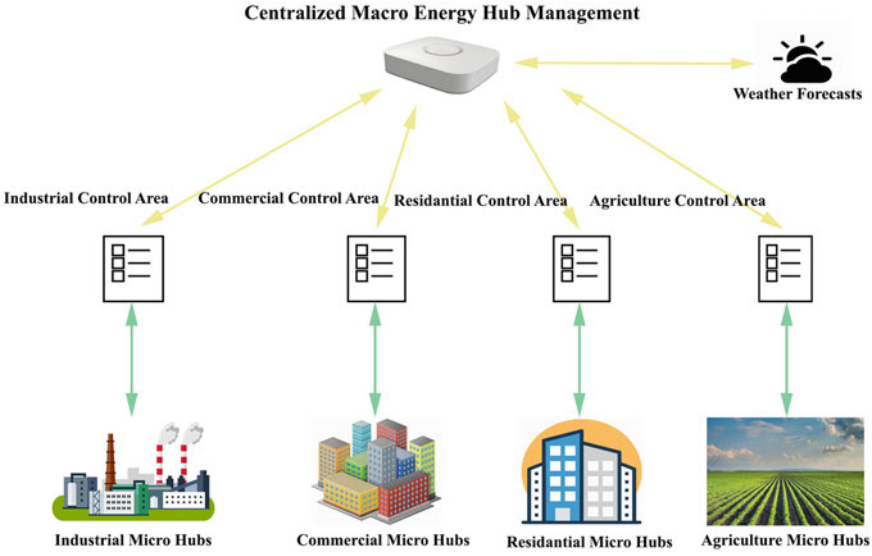


Fig. 10 Centralized macro energy management structure

As it can see, the agricultural sector has a great ability to use RES and increase its efficiency through local multi-carriers, smart technologies, and integrated management, especially in CGH. A few types of research have conducted in the field of energy hub in agriculture.

Considering the concept of energy hub for agriculture sector in the form of agricultural hubs, it is possible to obtain many advantages from the integrated management of the agricultural hubs system, such as optimum operating, reducing environmental impact, and the integration of RES can reduce energy cost, after all, produce an environmentally friendly product.

3.2 Macro Energy Hubs

Using the concept of energy hub, it can be modeled without limitations on the size of different systems. The entire energy hub has a supportive concept and divided into two categories of Micro Energy & Macro Energy. With the increase in the number of micro-hubs, the need for an integrated and coordinated management system will be at a higher level to form a continuous network of energy hubs. This network is called Macro Hub.

Macro hubs are a set of energy hubs that manage and control them in a coordinated manner. By the concept of macro hubs, huge energy systems are modeled, like residential complexes, industrial towns, and even the entire city. Macro hubs are integrated model of micro-hubs, which their management can be achieved a lot of advantages for the entire system, as well as any of the micro-hubs.

These sector's integrated management increases the efficiency and reduction of losses, fossil fuel consumption, environmental pollution, and total system costs. There are many ways to integrate energy carriers, but one of the best methods is to use the concept of an energy hub.

Optimization opportunities for the entire system are provided by modeling these systems in the form of macro hub. The energy hub considered a supernode that connects various energy carriers.

The advantage of modeling using the energy hub can consider the priority of each consumer, such as the total system constraints in the model and the optimization.

In any financial systems, the cost of consumers considered, the energy system is not exception by this rule and reducing the bills cost and energy consumed is the attention of consumers while the institutions of the supplier this energy, such as power companies, in addition to minimizing cost and increasing profits, should consider the power quality, peak shaving and the modification of the consume curve.

In order to optimize macro hubs, it must be considered the priorities of the entire system, which optimization problem is a double-layer hierarchical problem. Therefore, in addition to optimizing the micro-hub in the macro hub, we must also consider optimizing the micro-hub. Their consumer preferences and constraints considered at the micro hob level, and the issue of the optimal power flow (OPF) and in the overall state of the optimal energy flow (OEF) issue at the macro hub level concedes.

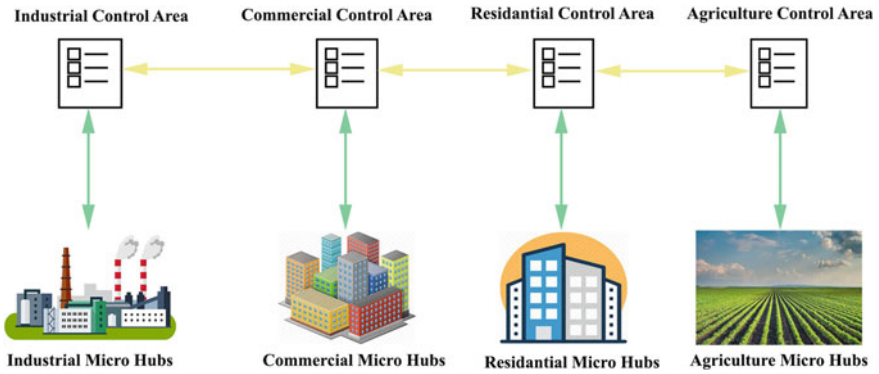


Fig. 11 Decentralized macro energy hub structure

Macro hub management and control do in two ways: centralized and decentralized. In the centralized model, that has shown in Fig. 10, the optimization issue for optimal macro hub performance and optimal functional conditions of each component define and responsible for data collection and signal processing, and sending control information is Central management unit. This unit adopts all decisions related to the optimum operation of the macro hub.

With the enlargement of the system scale and increasing the number of components and energy carriers, increasing the number of variables in the optimization problem and data needs to be processed, volume, cost and time calculation of them increased, and in some cases, the optimal response will be hard at a reasonable time, and the ability to control and monitor online is lost.

On the other hand, some representatives of active participation in central control programs are not allowed to access all the system representatives, information, and control of the equipment. In such circumstances, a decentralized or distributed system discussion occurs while shown in Fig. 11.

In this decentralized system, the nature of the optimization problem is not changed and will only divide into several sub-issues, which solve each sub-sections in parallel to reduce the calculation time and processing in a large scale system. On the other hand, this type of system and its division to sub-issue can increase the reliability of the decentralized system, as the disruption in the control area will only affect the same area and it will have a slight impact on the entire system performance whilst in a centralized system, the shape of the data collection and processing can have an effect on the overall system performance.

4 Conclusion

At this search, reviewing the concept of energy hubs was comprehensively investigated. Advanced Energy systems have been able to improve performance and efficiency in these systems by using the concept of an energy hub. More on the introduction of residential, commercial, industrial, and agricultural micro hubs and the challenges and characteristics of each of these sectors of energy consumption discussed. In the next stage, we discussed the management of these micro-energy hubs, which led to the formation of a network that is called Macro hub.

In the residential sector, due to long transmission and distribution lines, poor energy consumption management, low efficiency of energy consumption equipment is significant and requires integrated and optimal control. However, in recent years, with the renewable influence, energy storage systems, increasing the efficiency of equipment has improved the performance of these systems.

In the commercial sector because of a fixed consumption pattern that changes on a weekly, monthly, and even seasonal basis, energy management is more convenient. than the residential area and has a high potential for the use of the RES in this sector. Also, the commercial sector has large thermal loads such as HVAC, which requires optimal control of resources and cost reduction and emission. The industrial sector has a more predictable consumption pattern than the rest of the areas. On the other hand, due to the hardness of production scheduling, the existence of various energy carriers as inputs, the use of DR programs in this section is more complicated. Also, due to the need for high power, the implementation of the RES systems has a high cost. Therefore, the need for supportive policies on behalf of the Government for the industrial sector and increasing the efficiency of equipment to reduce the impact of the RES and increase the hope of participation in DR programs.

Energy plays a fundamental role in agriculture sector. Fossil fuels are the second source of energy sources due to the high potential of using renewable resources in this sector. in the agriculture sector, using the CGH system, high efficiency by contributing to the RES can be achieved. Three critical characteristics in all the studied sections are the same, including the improvement of equipment and techniques, policies and behavioral training, integrated energy management systems.

Energy management systems have more impact than other characteristics of the energy system. So, there is a high potential for the energy hub model in the integrated management of energy systems to benefit from DSM and especially DR programs and renewable resources and storage.

In addition to integrated management of various carriers in the form of the micro hub, integrated management of different micro hubs in the way of the Macro hub has technical, economic, and environmental advantages.

References

1. Favre-Perrod P (2005) A vision of future energy networks. In: IEEE power engineering society inaugural conference and exposition. Africa, pp13–17
2. Geidl M, Andersson G (2005) A modeling and optimization approach for multiple energy carrier power flow. In: IEEE Russia power tech conference, pp 1–7
3. Geidl M (2007) Integrated modeling and optimization of multi-carrier energy systems
4. Mohammadi M, Noorollahi Y, Mohammadi-Ivatloo B (2018) An introduction to smart energy systems and definition of smart energy hubs. In: Operation, planning, and analysis of energy storage systems in smart energy Hubs, pp 1–21. Springer International Publishing, Cham
5. https://www.eia.gov/energyexplained/use-of-energy/
6. Afshin N-G, Zare K (2018) Heating and power hub models for robust performance of smart building using information gap decision theory. Elec Power Energy Syst 98:23–35
7. Yuan Y, Bayod-Rújula AA, Chen H, Martínez-Gracia A, Wang J, Pinnarelli A (2019) An advanced multicarrier residential energy hub system based on mixed integer linear programming. Int J Photoenergy. https://doi.org/10.1155/2019/1384985
8. Albadi MH, El-Saadany EF (2008) A summary of demand response in electricity markets. Elec Power Syst Res 78:1989–1996
9. Pedrasa MAA, Spooner TD, MacGill IF (2010) The value of accurate forecasts and a probabilistic method for robust scheduling of residential distributed energy resources. In: IEEE 11th international conference on probabilistic methods applied to power systems, pp 587–592
10. Pedrasa ST, MacGill I (2010) An energy service decision-support tool for optimal energy services acquisition. Technical report. University of New South Wales
11. Beaudin M, Zareipour H (2015) Home energy management systems: a review of modelling and complexity. Renew Sustain Energy Rev 45:318–335
12. Sharma I, Cañizares C, Bhattacharya K (2015) Residential micro-hub load model using neural network. In: North American Power Symposium, pp 1–6
13. Bozchalui MC, Hashmi SA, Hassen H, Canizares CA, Bhattacharya K (2012) Optimal operation of residential energy hubs in smart grids. IEEE Trans Smart Grid 3:1755–1766
14. Rastegar M, Fotuhi-Firuzabad M, Lehtonen M (2015) Home load management in a residential energy hub. Elec Power Syst Res 119:322–328
15. Brahman F, Honarmand M, Jadid S (2015) Optimal electrical and thermal energy management of a residential energy hub, integrating demand response and energy storage system. Energy Build 90:65–75
16. Raza A, Khan MFN, Malik TN (2019) Risk constrained energy consumption control for commercial buildings. In: 2nd international conference on computing, mathematics and engineering technologies, pp 1–6
17. Moghaddam IG, Saniei M, Mashhour E (2016) A comprehensive model for self-scheduling an energy hub to supply cooling, heating and electrical demands of a building. Energy 94:157–170
18. Pérez-Lombard L, Ortiz J, Pout C (2008) A review on buildings energy consumption information. Energy Build 40:394–398
19. Lu N, Taylor T, Jiang W, Correia J, Leung LR, Wong PC (2009) The temperature sensitivity of the residential load and commercial building load. In: IEEE power & energy society general meeting, pp. 1–7
20. Steinfeld J, Bruce A, Watt M (2011) Peak load characteristics of Sydney office buildings and policy recommendations for peak load reduction. Energy Build 43:2179–2187
21. Abdelaziz EA, Saidur R, Mekhilef S (2011) A review on energy saving strategies in industrial sector. Renew Sustain Energy Rev 15:150–168
22. Mostafavi SM, Noorpoor A, Shafie-Pour MM (2019) Optimal model development of energy hub to supply water, heating and electrical demands of a cement factory. Energy 177:574–592
23. Lee D, Cheng C-C (2016) Energy savings by energy management systems: a review. Renew Sustain Energy Rev 56:760–777

24. Casisi M, Nardi AD, Pinamonti P, Reini M (2015) Effect of different economic support policies on the optimal synthesis and operation of a distributed energy supply system with renewable energy sources for an industrial area. Energy Conv Manag 95:131–139
25. Ding YM, Hong SH, Li XH (2014) A demand response energy management scheme for industrial facilities in smart grid. IEEE Trans Ind Inform 10:2257–2269
26. Paudyal S, Cañizares CA, Bhattacharya K (2015) Optimal operation of industrial energy hubs in smart grids . IEEE Trans Smart Grid 6:684–694
27. Farah A, Hassan H, Kawabe K, Nanahara T (2019) Optimal planning of multi-carrier energy hub system using particle swarm optimization. In: IEEE innovative smart grid technologies—Asia, pp 3820–3825
28. Bayrakcı AG, Koçar G (2012) Utilization of renewable energies in Turkey's agriculture. Renew Sustain Energy Rev 6:618–633
29. Aqeelur R, Abbasi AZ, Islam N, Shaikh ZA (2014) A review of wireless sensors and networks' applications in agriculture. Comput Stand Interfaces 36:263–270
30. Vadiee A, Martin V (2014) Energy management strategies for commercial greenhouses. Appl Energy 114:880–888

Sustainable Energy Systems Based on the Multi-energy Sources

Mehmet Zile

Abstract One of the most important and indispensable needs of people is energy. Today, with the increasing population and developing technology, the energy demand has increased considerably. Currently, one of the most important problems in the world is finding new energies and ensuring its continuity. Development in the economy is possible by having cheap, sufficient, quality and reliable energy sources. Energy, an important element of international power struggles, has gained an international dimension. Energy will continue to be at the forefront of the economic and social development of the countries in the coming centuries. This chapter provides detailed information about non-renewable and renewable energy sources. Solar and wind energy systems are emphasized. The working principles of hydroelectric and geothermal energy systems are explained. Mas-biomass energy systems, water wave energy systems are given extensive information on. Applications on these energy systems are emphasized. Examples of sustainable energy systems with different energy sources are given. Renewable energy sources can be used directly or converted to another form of energy. Examples of direct use are solar powered appliances, geothermal heating and water or wind mills. Examples of the most direct use are wind turbines or photovoltaic batteries for electricity generation. The development of renewable energy is related to human use of renewable energy sources. The interest in the development of renewable energy is directly related to the waste gases that fossil fuels give to the environment and the risks of fossil fuels and the use of nuclear energy.

Keywords Energy systems · Multi energy sources · Sustainable energy

M. Zile (✉)
Mersin University, Mersin, Turkey
e-mail: mehmetzile@mersin.edu.tr

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_11

Nomenclatures

A. Acronyms

AC	Alternative Current
DC	Direct Current
PIC	Peripheral Interface Controller
LCD	Liquid Crystal Display
LPG	Liquified Petroleum Gas
PV	Photovoltaic

Symbols/Parameters

V_{off}	open circuit voltage
P_{max}	maximum power
R_s	solar radiation
A_p	solar panel area
π_c	conversion efficiency of solar energy
δ	density
A	sweep area
v	speed
P	power
h	head difference in elevation drop
η	efficiency percent energy transferred
F	flow
g_a	acceleration of gravity
E	amount of energy available
B_F	biomass feed amount
B_{FCR}	biomass fuel conversion rate
F_{TV}	thermal value of fuel
E_{pw}	potential energy of wave
E_{kw}	kinetic energy of wave
d	density of water
a_w	wave amplitude
A_w	wave area

1 Introduction

Energy is an indicator of community development and a factor that directly affects development. Energy is an indispensable need at every point of human life and the continuation of life depends on the existence of energy. Energy is used in industry, transportation, current life and every stage of life. Therefore, humanity is constantly focused on finding and developing new sources of energy. Energy sources that can be

obtained by using various methods and techniques for economic purposes are called energy sources. There is a wide variety of energy sources both underground and above ground. Therefore, energy sources can be classified in many different ways. These; the substance may be classified according to its storage, conversion, renewability and use. Energy sources can be classified according to underground and surface sources. Underground sources are coal, oil, natural gas, uranium, thorium and geothermal. Aboveground resources are wood, biomass, wastes, sun, wind and biogas. Energy sources as origin; it is classified as inorganic and organic energy sources. Inorganic ones; uranium and thorium are sources of energy. Organic ones; coal, oil, wood, biogas and biomass are energy sources. Energy sources are classified as solid fuels, liquid fuels and gaseous fuels by substance. Solid fuels are coal, anthracite, hard coal, lignite, peat, asphaltite, wood and nuclear ores, coke and briquettes. Liquid fuels are gasoline, diesel, fuel oil, benzone, tar and derivatives, artificial gasoline, artificial diesel, alcohols, biomotorine and alternative oil. Gas fuels are gases such as natural gas, methane, butane, propane, synthesis gas, city gas, LPG, biogas, biohydrogen, wood gas. Energy sources are classified as primary and secondary energy sources according to their conversion. Primary Energy Sources are coal, oil, natural gas, biomass, solar, wind, water power and nuclear energy. Secondary Energy Sources are electrical, thermal and electro-magnetic energy sources. Energy sources are classified as conventional energy sources and alternative energy sources according to their availability. Traditional energy sources are coal, oil, natural gas and nuclear energy. Alternative energy sources are sun, wind, water power and biomass. Energy resources are classified as renewable and renewable energy sources according to their renewability. Non-renewable energy sources are coal, oil, natural gas and nuclear energy. Renewable energy sources are sun, wind, biomass and water power. The development of renewable energy is related to human use of renewable energy sources. The interest in the development of renewable energy is directly related to the waste gases that fossil fuels give to the environment and the risks of fossil fuels and the use of nuclear energy. The future outlook for energy supply and demand suggests that we will have to resort to other renewable energy sources more often. Today, everyone has a duty to protect and consume natural resources together with predicting climatic changes. Energy consumption continues to increase continuously and is becoming increasingly dependent on fossil fuel imports. This could also endanger security in the provision of energy. Moreover, the increase in the use of fossil fuels has negative effects on the environment. We have to reduce carbon dioxide production and slow down global warming. Energy supply and demand will have to apply to other renewable energy sources more frequently in the future. We have to give priority to effective methods in the field of energy. In order to ensure economic and social growth, as well as to improve the environment and public health, we are obliged to prevent the consumption of natural resources and to anticipate climatic changes, to contribute to the energy sector and provide alternatives. Today, as well as renewable energy fields such as wind energy and solar energy, heating, cooling, conversion of heat energy into electrical energy and industrial development in the fields of effective technologies continues to develop. Thanks to the positive improvements such as efficiency increase and cost reduction, systems become more accessible and applicable.

Solar energy, biofuels, efficient use of energy in buildings and high-efficiency city heating methods or technologies using renewable energy create new markets for themselves. This chapter provides a quick overview of energy resources and explains how sustainable energy can be obtained. This section can be grouped under the following groups.

- Introduction
- Non-Sustainable Energy Sources
- Sustainable Energy Sources
  - Solar Energy Systems
  - Wind Energy Systems
  - Hydroelectric Energy Systems
  - Geothermal Energy Systems
  - Biomass EnergySystems
  - Water Wave Energy Systems
  - Hydrogen Energy Systems
- Sustainable Energy Systems Created on the Multi-energy Sources in Silifke District of Turkey
- Renewable Power Plant Created in SilifkeDistrict of Turkey
- Sustainable Energy Management Model
- Conclusion

Practical examples of sustainable energy systems based on hybrid energy sources are given.

2 Non-sustainable Energy Sources

Non-renewable energy sources are energy sources that cannot be regenerated when consumed. Fossil energy sources such as coal, oil and natural gas and uranium and thorium which are the raw materials of nuclear energy are non-renewable energy sources. Non-renewable energy sources consist of fossil energy sources and nuclear energy sources. Fossil Energy Resources are coal, oil and natural gas. The raw material sources of nuclear energy are uranium and thorium ores with the highest radioactivity. Coal; These are rocks of organic origin which contain between 60% and 90% carbon, which can burn directly with the oxygen of the air. Petroleum is a source of energy that occurs over millions of years as a result of the accumulation and decay of earth layers over organic materials such as animal and plant remains. The main element in the chemical composition of crude oil is carbon (C). Natural gas is an energy source formed by the transformation of organic materials in the lower layers of the earth under pressure and heat as a result of the natural transformation that lasted for millions of years. Natural gas contains methane and lower amounts of ethane, propane, nitrogen, oxygen, carbon dioxide. Among the nuclear energy sources, uranium contains compounds such as uranitite, peblende, carnotite, otunite,

torbenite and tuyamunite. In uranium, a source of nuclear energy, thorium is almost like uranium. It is found in every kind of rock. It is concentrated in pegmatites of acid depth rocks, pneumatological and hydrothermal formations. It is more difficult and more expensive to produce nuclear energy than thorium. 4 kwh of 1 kg coal, 5 kwh of 1 kg oil, 60,000 kwh of 1 kg uranium are produced [1].

3 Sustainable Energy Sources

Renewable energy sources are independently renewable energy sources. These include hydroelectric energy, geothermal energy, biomass energy, wind energy, solar energy, tidal energy and hydrogen energy. Renewable energy sources are inexhaustible and the least harmful to the environment from conventional energy sources. Each of the various applications of renewable energy sources has special advantages. None of these resources creates pollution, neither liquid nor gas, throughout their operations. Two systems, passive and active, are utilized in the use of renewable energy sources. Examples of direct use are geothermal heating, water or wind mills. In these examples, energy savings can be achieved by significantly reducing heating and cooling costs. The indirect system is provided with mechanical equipment integrated to the structure in order to obtain electricity and heat energy from renewable energy sources such as solar and wind. Examples are wind turbines or photovoltaic batteries used in electricity generation. It is possible to obtain cheap and clean energy from renewable energy sources [2–4].

3.1 Solar Energy Systems

Fossil fuel sources are running out. Due to the increasing costs of energy and environmental pollution, intensive studies are being carried out on renewable energy sources. Solar energy is important in terms of the wide usage area and the possibility of utilization. Solar Power Plants are formed by combining a large number of solar panels. Solar power plants that produce renewable electricity provide the energy of the settlements. The units called photovoltaic, solar panel, solar cell, pv, solar module, solar cell are transformed into electricity from the sun. The first energy obtained from the sun is dc direct current electrical energy. This dc direct current electrical energy is converted into ac alternating current energy in switchyard and used in homes and factories. Solar cells are semiconductor elements that convert sunlight into electrical energy. The solar cells are square, rectangular and circular. The area of the solar cells is around 100 cm^2. The thickness of the solar cells is around 0,2–0,4 mm. When light falls on the solar cells, electrical voltage is generated at the ends [5]. A large number of solar cells are connected in parallel or in series to increase power output. This structure is called a solar cell module or a photovoltaic module. The production of electrical energy with PV solar panels is given in Fig. 1. The conversion efficiency of

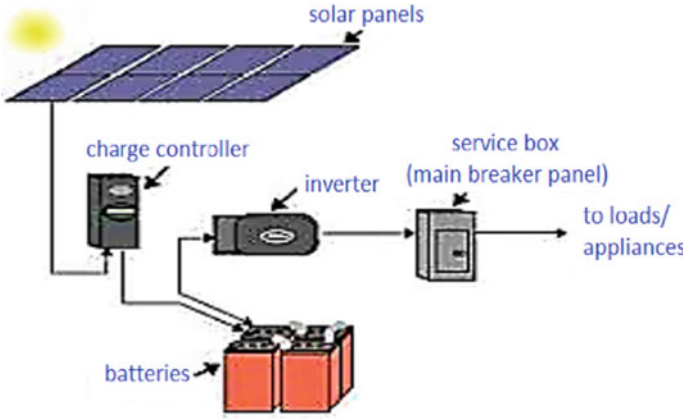


Fig. 1 The production of electrical energy with PV solar panels

a photovoltaic cell is indicative of how much of the incoming photons are converted into electrical energy [6]. In the photovoltaic cells, electrons are activated and current and voltages are formed.

The energy that can be obtained from a photovoltaic system depends on the amount of solar radiation, sunshine duration, panel efficiency and air temperature. Open circuit voltage Voff refers to the voltage observed in the circuit when the current is assumed to be zero. It refers to the observed current value when no load is present. How well the connections are made in the cell and how low the resistance caused by the connections is expressed by the filling factor. Maximum power obtained from panel is given in Eq. (1).

$$P_{\max} = R_s.A_p.\pi_c \tag{1}$$

$P_{\max}$ is maximum power obtained from panel, R_s is solar radiation, A_p is panel area and π_c is conversion efficiency of solar energy to electrical energy. In series of connected panels; the voltage of a single panel is the voltage of the array and the sum of the current values of all panels is the current value of the array. In the series of parallel connected panels; the current of a single panel is the current of the array and the total voltage of the panels is the voltage of the array. Among the many semiconducting materials, the most suitable ones for making solar cells are silicon, gallium arsenite, cadmium and tellurium. At the interface of the PN material, i.e. in the joint region, a negative charge is accumulated on the P side and a positive charge on the N side. This region is referred to as "transition region" or "load-free region". The electric field formed here is called the "structural electric field". For the semiconductor joint to work as a solar cell, photovoltaic transformation must be achieved in the joint region. This transformation occurs in two stages, first, by reducing the light to the joint region to form electron-hol pairs. Secondly, they are separated from each other by the electric field in the region. Semiconductors consist

of two energy bands separated by a forbidden energy range. When a photon with energy equal to or greater than this energy range is absorbed by the semiconductor, it transmits its energy to an electron in the valence band, allowing the electron to rise to the conductivity band to form an electron-hol pair. If this pn joint is formed at the interface of the solar cell, the electron-hall pairs are separated from each other by the electric field there. Thus, the solar cell works as a motor pump, which pushes electrons to the n-zone and hells to the p-zone. Separated electron-hall pairs generate energy by generating a power output at the ends of the solar cell. This process continues continuously in the same way as a photon strikes the surface of the battery to obtain direct current electrical energy. The obtained direct voltage is converted to electrical energy by using converters in switchyard and residential and factory. Solar cells can be formed using many different elements. These are crystal silicon, gallium arsenite, amorphous silicon, cadmium telluride. In concentrating solar power plants, solar energy is collected in collectors. Water, air, helium and sodium are used as heat carriers. The generation of electrical energy is obtained from steam or gas turbines using an organic liquid, such as toluene [7, 8]. The generation of electrical energy with a concentrating solar power plant is given in Fig. 2.

With reflectors made of heliostatic mirrors, solar radiation is reflected to the heat generator mounted on the tower and concentrated [9]. The fluid in the pipes in the receiver absorbs solar energy. This liquid, which receives high heat, turns into pressurized steam in the steam generator. This steam quickly hits the steam turbine and turns into motion energy. Motion energy is converted into electrical energy in

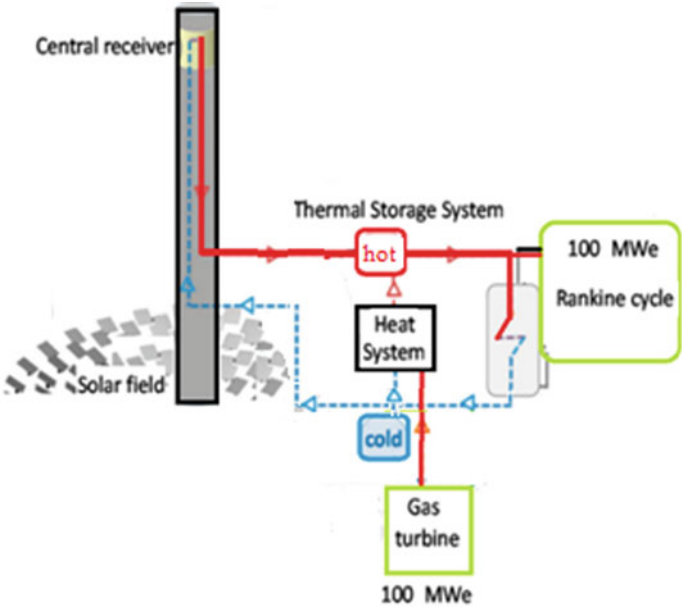


Fig. 2 The generation of electrical energy with a concentrating solar power plant

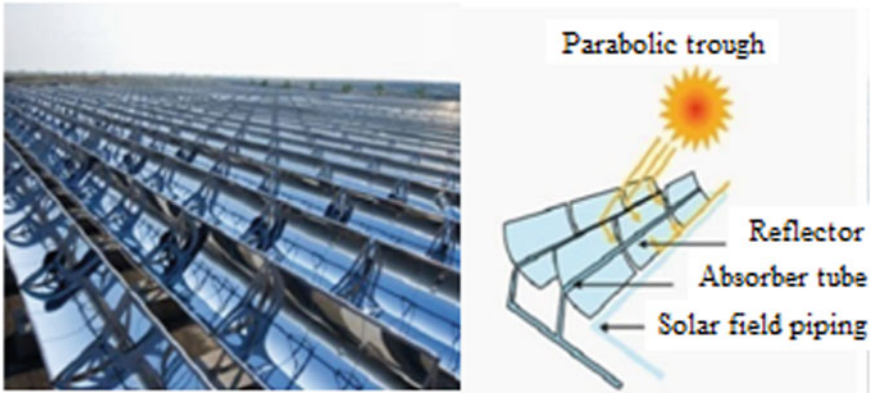


Fig. 3 The parabolic trough reflective solar power plant

the electric generator. The panels in the solar plant may be movable or immobile. The feet move horizontally and vertically, facing the sun. When the panels follow the sun, energy production is 25% higher. In solar energy systems with parabolic trough reflectors, solar radiation is reflected to the receiving pipe by parabolic bent panels. In these pipes, high heat energy is carried by flowing material which can carry high temperature. This high heat energy is converted into steam in the steam generator. The parabolic trough reflective solar power plant is given in Fig. 3.

Sufficient space is required for each module to install solar modules [10]. Since the amount of energy that can be obtained through each module is certain, the area covered must be increased in order to increase the production capacity. Solar Power Plants investment can be costed in two stages. The first stage is the costing of the solar modules and the second stage is the costing of the solar cells. The cost of solar modules and the cost of solar cells are on average equal to each other and their share in the investment is around 50%. It is understood that the total installation cost is covered after five years of operation of each solar power plant, including the depreciation costs of the solar panels and power control and converter units to be used in the power plant.

3.2 Wind Energy Systems

Wind is defined as the movements of the air caused by solar rays heating the highly variable surface of the Earth at different rates. Since the surface is not homogeneous, the rate of absorption of solar energy varies according to location, time and geographical characteristics. This creates differences in atmospheric pressure, temperature and density. As a result of the resulting forces, air travels from one place to another, causing wind flows. It is obtained by using turbine asynchronous motor as generator in wind power plants [11–13]. Electricity production from wind energy is shown in

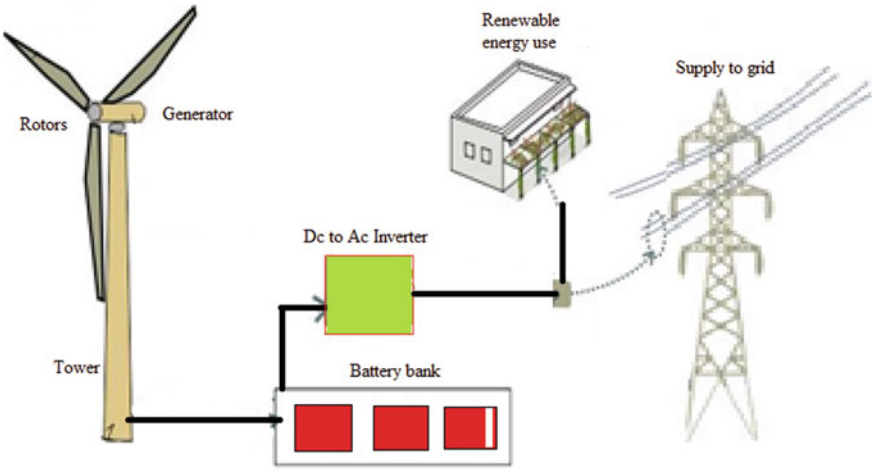


Fig. 4 Electricity production from wind energy

Fig. 4 [14]. The theoretical power to be obtained from the wind is obtained as in Eq. (2).

$$P = \frac{1}{2}.\delta.A_s.v^2 \tag{2}$$

Here, δ is air density, A is sweep area and vis speed. When the Eq. (2) is examined, the wind power changes in proportion to the square of the rotor sweep area and the wind speed. When wind turbines use wind speed, the power to be obtained cannot be the same as theoretical power as in Eq. (2). Wind turbines can use up to 65% of the power obtained from wind power. An important problem of wind power plants is the difficult control of power output. In these plants, safety must be ensured in high wind and braking of the system should prevent overproduction.

Power electronic control circuits are required to control the variable rotor speed. The turbine operates at 2–12 m/s wind speed. For wind turbine design, for turbine selection; cold-hot climate, low-medium-high wind regime, small-high density, network quality and environment are important. The power of a turbine in the wind is between 600 kW and 4 MW. In today's technical conditions, the economic wind potential that can be established at a speed of 6 m/h at a height of 10 meters with a usage time of 2800 h per year is around 10,000 MW or 28 billion kWh. This potential drops to 1400 h under the worst wind conditions [15–18].

3.3 Hydroelectric Energy Systems

Hydropower energy means the power provided by moving water. Hydroelectric power is generated by hydroelectric power plants which are installed in front of natural or artificial dam lakes and which are quite low according to the water fall level. Hydroelectric energy is the energy provided by converting the potential energy of water into kinetic energy. The amount of energy depends on the drop height and flow variables. Hydroelectric Energy System is given in Fig. 5 [19]. The power generated by a hydroelectric power plant is calculated as in Eq. (3).

$$P = \delta.h.\pi.F.g_a \tag{3}$$

P is power, δ is density, h is head difference in elevation drop across the turbine, η is efficiency percent energy transferred, F is flow and g_a is acceleration of gravity. High flowing water hits the turbine. The potential energy of the incoming water is converted into kinetic energy in the turbine.

In the generator connected to the turbine, this energy is converted into electrical energy. The drop height and flow rate of the water coming from the turbine determines the power to be produced. Flow rate is the amount of water flowing through a certain section over a period of time. Hydraulic power-based power plants can be constructed in certain geographical locations due to the characteristics of the streams and regimes of the rivers.

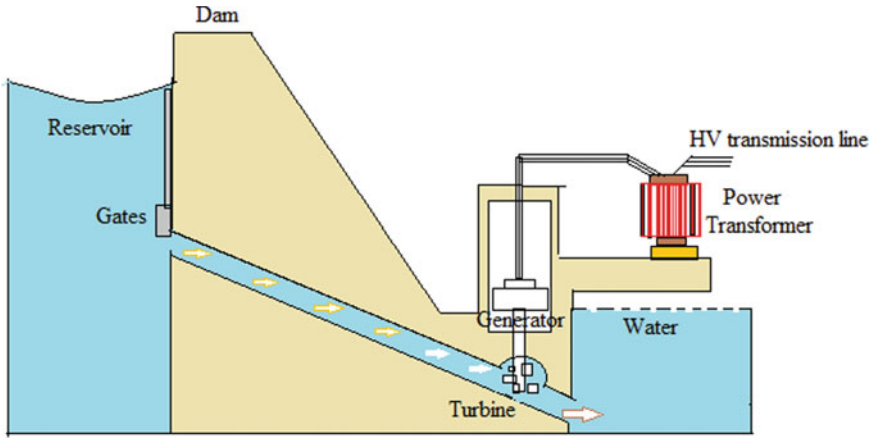


Fig. 5 Hydroelectric energy system

3.4 Geothermal Energy Systems

Geothermal Energy System is given in Fig. 6 [20–24].

Geothermal energy is a type of energy that occurs as a result of the transfer of the heat coming from the depths of the earth's crust to the underground waters and the warmed water reaches the earth. When magmatic activity in the Earth's crust ceases, the magma gradually cools. Cooling continues for thousands of years and occurs slowly. During cooling; together with some gases, natural steam is also produced. Gases and steam reach the surface in the form of hot spring water, geysers and natural steam through the cracks and crevices in the depths of the volcanic zones or forty lines in the earth's crust. Natural steam is drilled to the surface. When natural steam strikes the steam turbine, movement energy is generated. When this motion energy rotates the generator, electrical energy is generated.

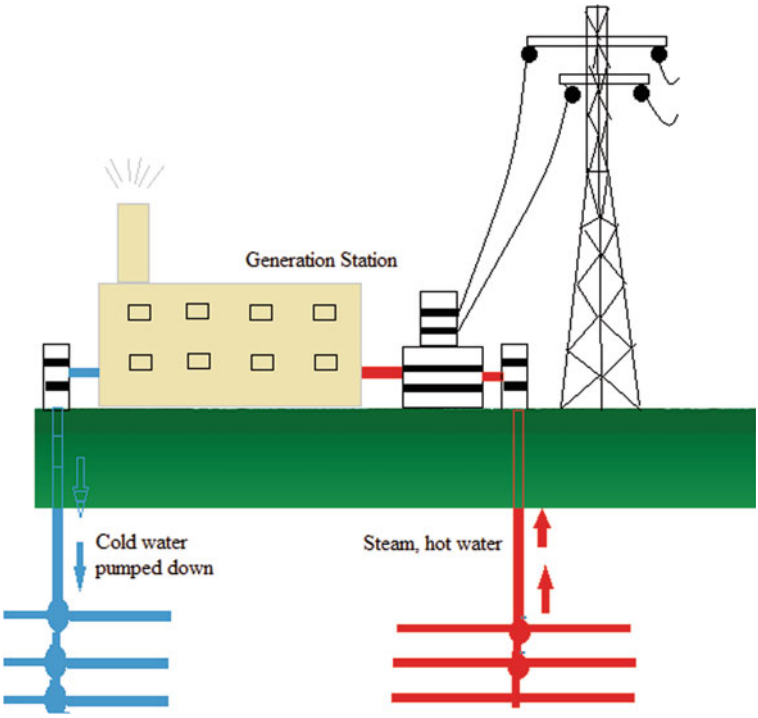


Fig. 6 Geothermal energy system

3.5 Biomass Energy Systems

Biomass energy is a type of energy gained by photosynthesis. The materials of biomass energy are plant and animal products. Biomass energy is derived from biomass fuel of organic origin. Wood waste, forest waste, plants such as corn and wheat, herbs, algae, algae in the sea, vegetables and fruits from households, organic waste of all kinds, animal droppings, fertilizer and food industry waste constitute biomass. The biomass is fermented in an oxygen-free environment to produce a combustible gas. This flammable gas is called biogas. This gas burns with a bright blue flame, is a gas of high calorific value. This flammable gas produces high-value steam. Biomass Energy Systems is shown in Fig. 7 [25]. Each of the biochemical and thermochemical processes has processes and calculations based on the analysis of chemical reactions. The energy of the resulting solid, liquid and gas fuels is calculated as in Eq. (4).

$$E = B_F.B_{FCR}.F_{TV} \tag{4}$$

E is amount of energy available (MJ), B_F is biomass feed amount (kg), B_{FCR} is biomass fuel conversion rate and F_{TV} is thermal value of fuel (MJ/kg). The generated steam is converted into motion energy in the steam turbine. This generated motion energy is converted into electrical energy in the generator. Garbage fuels from urban wastes are also important sources of biomass. In addition to energy generation, garbage thermal power plants are also important for waste disposal. It is not only

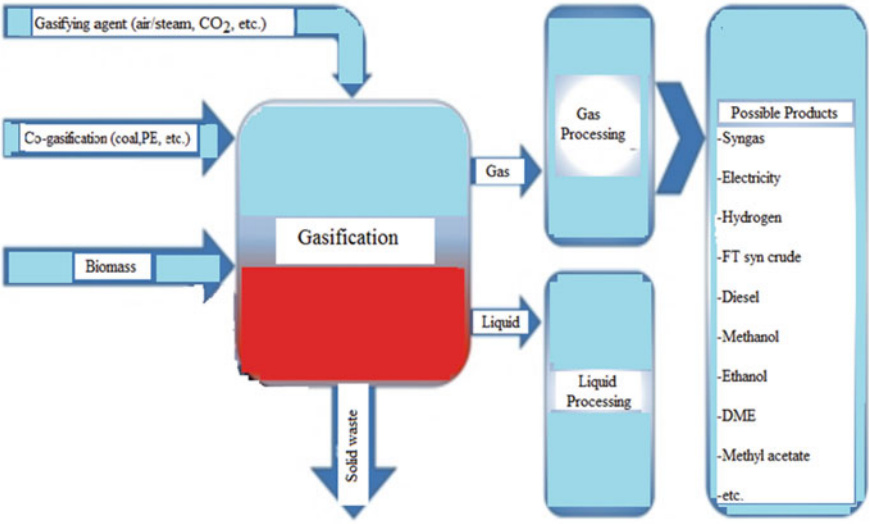


Fig. 7 Biomass energy system

possible to generate electricity from these plants, but it is also possible to build combined cycle plants where heat and electricity are produced together.

3.6 Water Wave Energy Systems

Water Wave Energy System is shown in Fig. 8 [26]. The attraction power of the Moon and Sun on Earth shows on the ocean's water bodies. Ocean waters swell twice within 24 h, with an interval of about 6 h and 20 min, and back twice to normalize. These movements are called tidal. The potential energy of the wave is calculated in Eq. (5) and the kinetic energy of the wave is calculated from the equation in Eq. (6).

$$E_{pw} = \frac{1}{4}.d.g.a_w^2.A_w \tag{5}$$

$$E_{kw} = \frac{1}{4}.d.g.a_w^2.A_w \tag{6}$$

E_{pw} is potential energy of wave (J), E_{kw} is kinetic energy of wave (J), d is density of water, g is gravity acceleration, a_w is wave amplitude and A_w is wave area. These kinetic movements occur in electrical energy by rotating the water turbine. It is possible to generate electrical energy by operating a water turbine, especially in tidal hydroelectric power plants to be constructed at certain locations along the shores of the oceans.

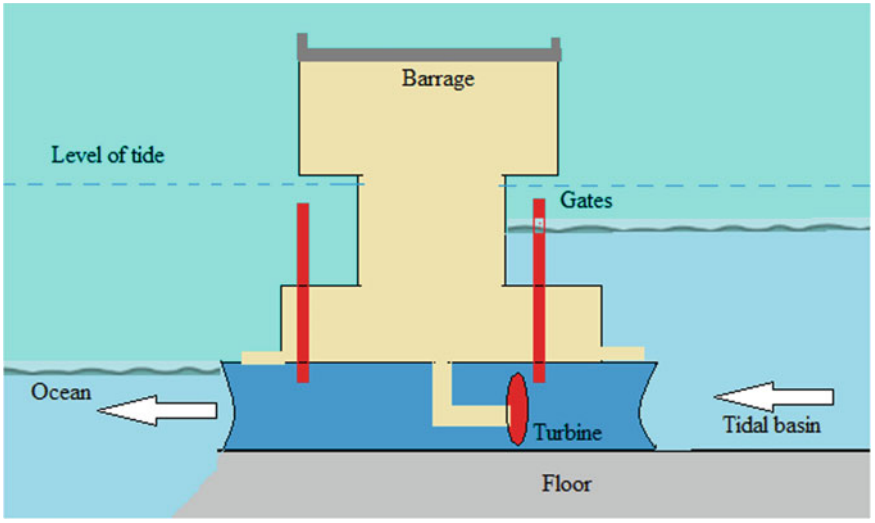


Fig. 8 Water wave energysystem

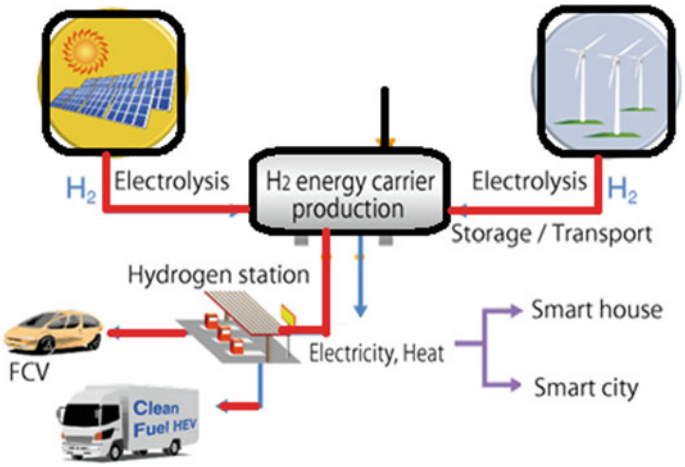


Fig. 9 Hydrogen energy system

3.7 Hydrogen Energy Systems

Hydrogen Energy System is shown in Fig. 9 [27].

The importance of hydrogen in new energy sources is increasing rapidly with each passing day. Hydrogen is the most abundant element in the world and is a colorless, odorless, approximately 15 times lighter than air and completely non-toxic. Due to its light weight, it has a very high spreading property. It can be liquefied at −250 °C. Hydrogen has the highest energy content per unit mass of all known fuels. The systems in which hydrogen is used as fuel and chemical energy is converted directly into electrical energy are called hücr fuel cells. The combustion products of hydrogen in these systems are only water and water vapors. The desired energy is obtained by introducing hydrogen directly or through any source that releases hydrogen. This energy from hydrogen is called Hydrogen Energy. Hydrogen gas is produced by water, solar energy, or wind, wave, and biomass, which are regarded as different methods, as well as by water.

4 Sustainable Energy Systems Created on the Multi-energy Sources in Silifke District of Turkey

The solar panels used have photovoltaic cells previously used in calculators and clocks. Photovoltaic cells are made of semiconductor material used in computer chips. When sunlight is absorbed by these substances, electrons are separated from the atoms in which they are released and an electric current is formed in the matter. The conversion of light to electricity is called photovoltaic. Accumulators are used to

Fig. 10 Environment and park lighting application

store this current. It is possible to obtain the desired 12, 24 and 48 V dc voltage with serial and parallel connections of the accumulators. Depending on the application, a new solar energy system has been created by using solar cell modules together with accumulators, inverters, battery charge control elements and various electronic circuits. By using huge mirrors directed to a central focus, very high radiation was obtained at the focal point, resulting in more electrical energy. With this system, it is difficult and expensive to carry fuel to a generator that is not far from residential areas, for example devices, indoor or outdoor lighting, agricultural irrigation, lighthouses. In the solar energy lighting system, lighting can be installed anywhere without the need for cable pulling, without any expense other than the cost of installing the system. Environment and park lighting applications are shown in Fig. 10.

In a city dwelling, a hybrid energy system was designed and implemented where the electrical energy of the entire dwelling was provided from the solar energy in the presence of sunlight and when the sunlight was insufficient or not, the energy supplied to the dwelling was provided from the solar energy batteries or in case the batteries were empty. Hybrid energy application is given in Fig. 11.

On the square surface of 5 m^2, 0,3 mm thick angle adjustable photovoltaic element (PV) is used to reflect the sun rays to the photovoltaic element, and two step motors are used for angle adjustment of these mirrors and photovoltaic element. Solar regulators are designed and used to provide battery charging under optimal conditions throughout the year. In order to convert the 24 V dc voltage stored in the batteries into full sine 220 V AC voltage required for feeding the instruments in the dwelling, a converter power electronics circuit was created using thyristors and diodes as a main element and a 220 V AC voltage and a power of around 2950 watts were obtained. At certain times of the day, the mirrors used to reflect the sunlight to the photovoltaic element at certain angles are automatically adjusted by the stepping motor with time relay. In the embodiment, the mirror angle is 100^o at 9:00, the mirror angle is 95^o at 10:00, the mirror angle is 90^o at 11:00 and the mirror angle is 85^o at 12:00. Two way switches are used in each electrical device in the residence to use solar energy if desired or mains electricity when desired. With this hybrid system, it is possible to

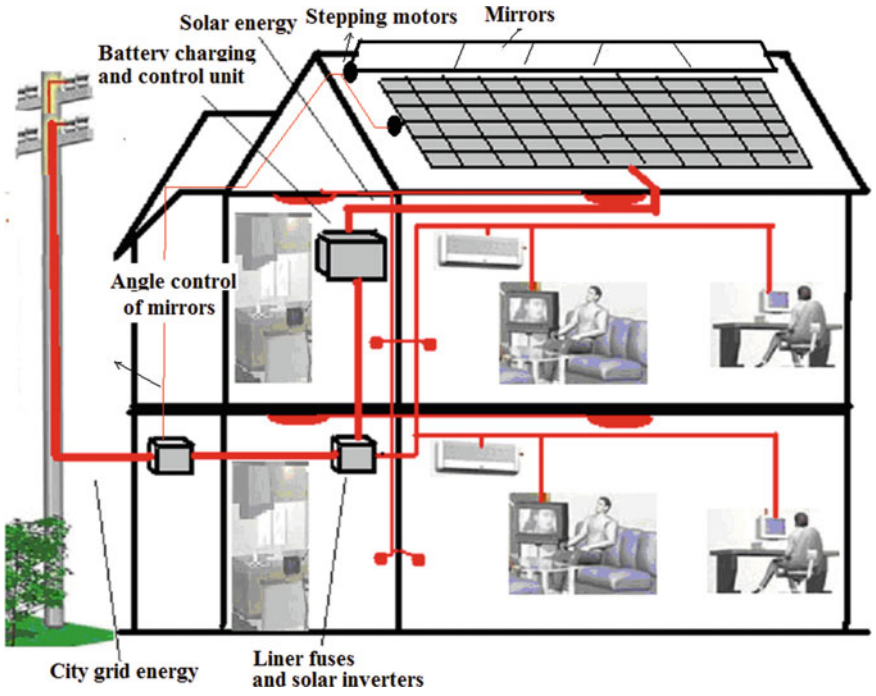


Fig. 11 Hybrid energy application

use energy at a very low price in the residence. Considering the fact that city mains electricity will be cut off in earthquake zones, solar stair lighting application using 24 V DC lamps without any harm to human health has been implemented in order to provide emergency exit safely to the building residents. Stair lighting application of buildings in earthquake risk areas is given in Fig. 12.

An energy system in which the chalet (plateau house), which cannot reach the city electricity grid, is supplied from solar energy, has been designed and implemented. On the square surface of 3 m^2, 0.3 mm thick photovoltaic element (PV) with angle adjustment, angle concentrator mirrors to reflect the sun rays to the photovoltaic element, and two step motors for angle adjustment of these mirrors and photovoltaic element. By using solar regulators, battery charging is provided in optimal conditions throughout the year. In order to supply the 24 V dc voltage stored in the batteries and convert it to the required full sinus 220 V AC voltage, a converter power electronics circuit was created using thyristors and diodes as a main element and a 220 V AC voltage and a power of 2650 watts were obtained. At certain times of the day, the mirrors used to reflect the sunlight to the photovoltaic element at certain angles are automatically adjusted by the stepping motor with time relay. With this system, the opportunity to use energy is provided in this mountain house (plateau house) where the city electricity network cannot reach. Mountain house (plateau house) application is given in Fig. 13.

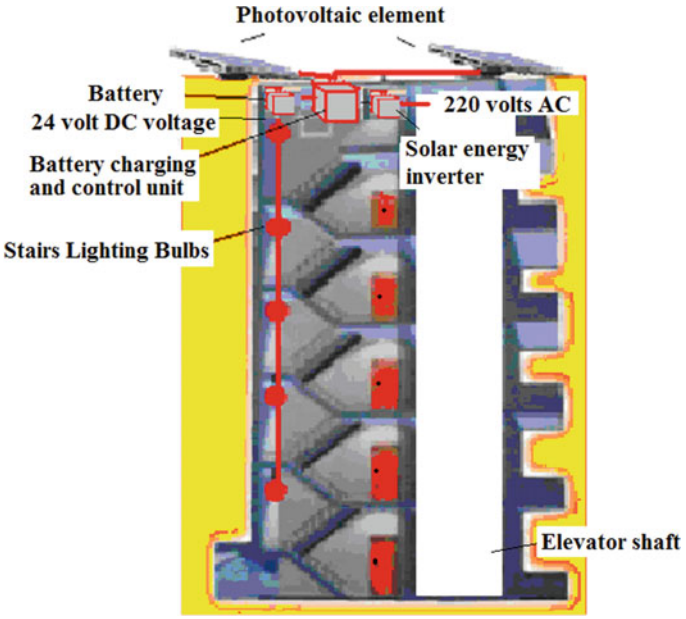


Fig. 12 Stair lighting application of buildings in earthquake risk areas

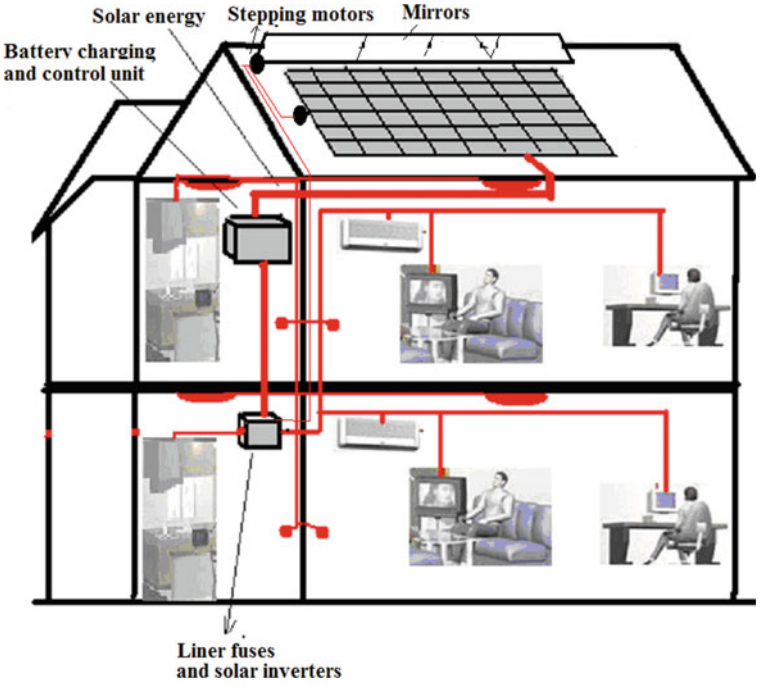


Fig. 13 Mountain house (plateau house) application

In the agricultural lands away from the network lines of the districts of Mersin, the most suitable method to meet the irrigation pumps with groundwater sources is the solar and wind energy based irrigation systems. Regular irrigation in agricultural lands is one of the most important issues. By means of drilling pumps mounted in the opened well, the water can be taken from underground or wells and transported to irrigation canals or highland houses. It is possible to draw water up to 350 tons per day from a depth of 60–70 m. In practice, systems that combine more than one energy source are called hybrid systems. The purpose of the hybrid systems is to increase the efficiency and to ensure that the energy needs of the system are met by the use of energy resources. Hybrid system components can be created by combining two or more sources. For example, solar energy, wind power, diesel or solar, fuel cell, wind or solar, wind, hydrogen energy sources such as hybrid energy systems are available. The most important factors that determine the number of sources and the type of the source in such applications are that there is sufficient level of resource in the region where energy will be produced and there is sufficient technology to bring the system together in some types of energy [28, 29]. Equipment used in solar and wind power agricultural applications: drilling pumps used in underground water extraction, solar panels converting solar energy into electrical energy and wind rose generators that convert wind energy into electrical energy. The inverter is the switching unit and the battery charging unit. In Fig. 14, it is shown the hybrid irrigation agriculture applied.

In Fig. 15, it is shown a diagram of the implementation of the hybrid energy of a village providing drinking and potable water.

Today, due to the finite fossil fuels and the environmental problems they have created and the high costs, the energy resources and production techniques have been revised. Environmental problems such as air pollution, global warming, soil

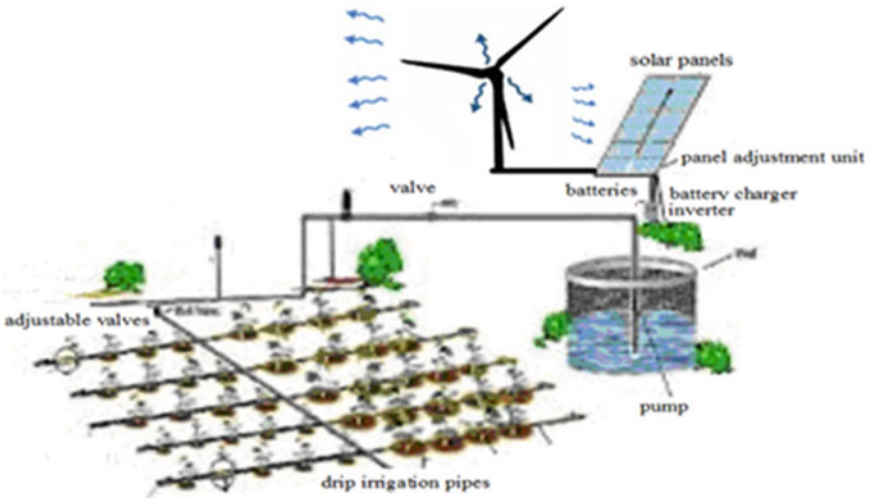


Fig. 14 The hybrid irrigation agriculture applied

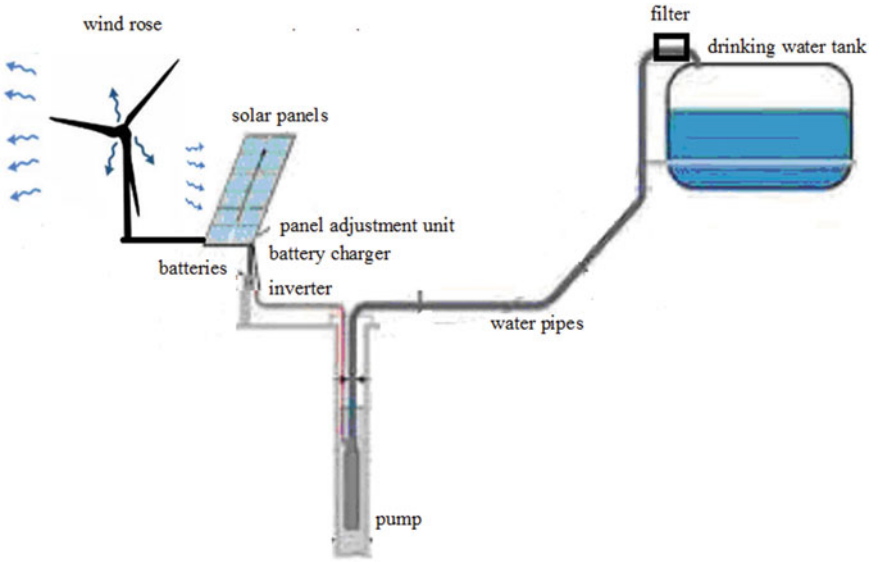


Fig. 15 The diagram of the implementation of the hybrid energy of a village providing drinking and potable water

and water pollution that are generated during the production of energy from fossil fuels are increasing day by day. In order to eliminate these problems and to reduce the increase in production and transmission costs, renewable energy sources should be utilized [30–32]. In case of an earthquake fire or central transformer failure on the intercity roads without mains electricity, the traffic lights due to the interruption of the city mains electricity did not work and as a result of this, electric accidents could occur and solar energy traffic lamps and solar traffic signs were designed and applied. Signaling applications are given in Fig. 16. It is seen that the traffic lights which are applied with solar energy and traffic signs which are illuminated with solar energy are more effective in concentrating the attention of drivers on these signs.

It has been found to be very cost-effective and efficient in terms of not requiring any cable installation along the streets and long roads, it can be easily installed at every intended point, and at the same time it provides self-consuming electrical energy. Accumulator and battery charge control circuits are used for continuous operation in the absence of sunlight. The presence of a lighthouse to guide the vehicles at sea and to indicate hazards is essential. The light of the lighthouse illuminates the entrance of the harbor to the sailors. Solar light systems should also be used in light houses where city mains electricity cannot be reached. Lighthouse application is given in Fig. 17.

The city electricity grid is designed as a hybrid energy system where the electricity is supplied from the solar energy in the presence of sunlight in an existing dwelling and the necessary energy is supplied to the dwelling from the city network when the

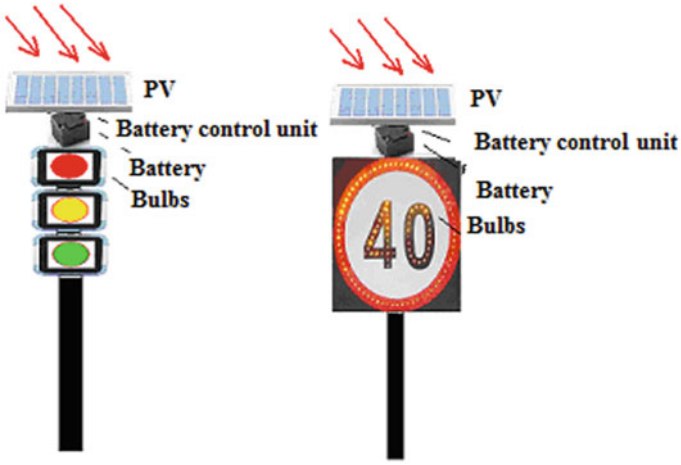


Fig. 16 Signaling applications

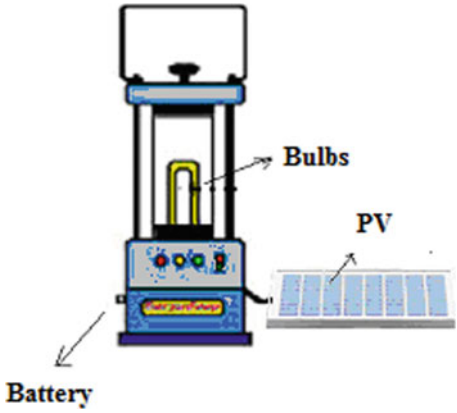


Fig. 17 Lighthouse application

sunlight is insufficient or not. energetic residential system was designed and implemented. In order to provide intense sunlight to the solar panels, the mirrors can be turned directly to the sun, thus providing more efficient energy and storage. Again, in the vineyards and gardens where the city's electricity network could not reach, the solar-powered agricultural irrigation system, where the electricity required for irrigation was supplied from solar energy, was designed and implemented. On intercity roads without mains electricity, in case of earthquake fire or central transformer failure, traffic lights due to interruption of city mains electricity are not working and as a result of this, electrical accidents may occur and solar energy traffic lights and drivers are designed and applied. Since solar applications do not require cable installation, it is easy to install and does not have any expense other than its cost. The primary purpose of our system is to make the most of the sun's rays and wind

as much as possible. The better the solar panel takes the sun rays and the better the position of the wind gust wind, the higher the energy production will be. One of the goals is to detect the location of the sun and to set the solar panels to the exact position where they need to be. The other is that the seven wind speed sensors detect the orientation of the wind and bring the wind mood to the exact position required for the horizontal and vertical axes. In our system, according to the fuzzy logic theory, each sensor is able to make the position adjustment control itself of the panels and the wind gauges with the results taken into consideration. Parts in the system; seven light sensors and seven wind speed sensors, stepper motors with solar panel and wind rose for power generation, solar panel and wind rose, PIC control card for fuzzy logic control of the system, LCD displays for system status information retrieval. In addition to the components of this hardware, we have software that is installed in our microcontroller and enables the system to function. The membership grades and values of each sensor are determined according to the information from the sensors. The specified membership values are recorded in the relevant part of the database. The location information and values are saved in the database where it is separated. These values are constant in each case. Program sections are prepared for each rule. These sections include the conditions set out in the rules. The minimum sensor value determined in each rule section is multiplied by the position constant and stored in the relevant memory area as the rule result value. Rule result values and rule minimum values that are found when all rules are completed are collected separately, and these sum values are stored in the relevant memory area. The total rule value found is divided by the total rule minimum value and the corresponding output port is activated according to the value obtained. Renewable energy system block diagram is shown in Fig. 18.

The energy system is shown in Fig. 19.

The system flow diagram is shown in Fig. 20.

The first step is ‘get information from the Sun and wind speed sensors’. The second step is ‘convert analog data to digital’. The third second step is ‘calculate fuzzy membership values’. The fourth step is ‘apply the fuzzy rule chart’. The fifth step is ‘make the corresponding output active’. As a result ‘bring the panellies and the windy smile to the desired position’.

5 Renewable Power Plant Created in Silifke District of Turkey

Renewable power plant created in Silifke-Turkey is shown in Fig. 21.

Modeling of battery-powered wind-solar hybrid power generation system components, energy flow control with the switched controller and fuzzy logic controller of the modeled system has been realized. Hybrid power generation systems are power generation systems in which two or more classical and renewable power generation systems are connected in parallel to one another. Hybrid power generation systems are

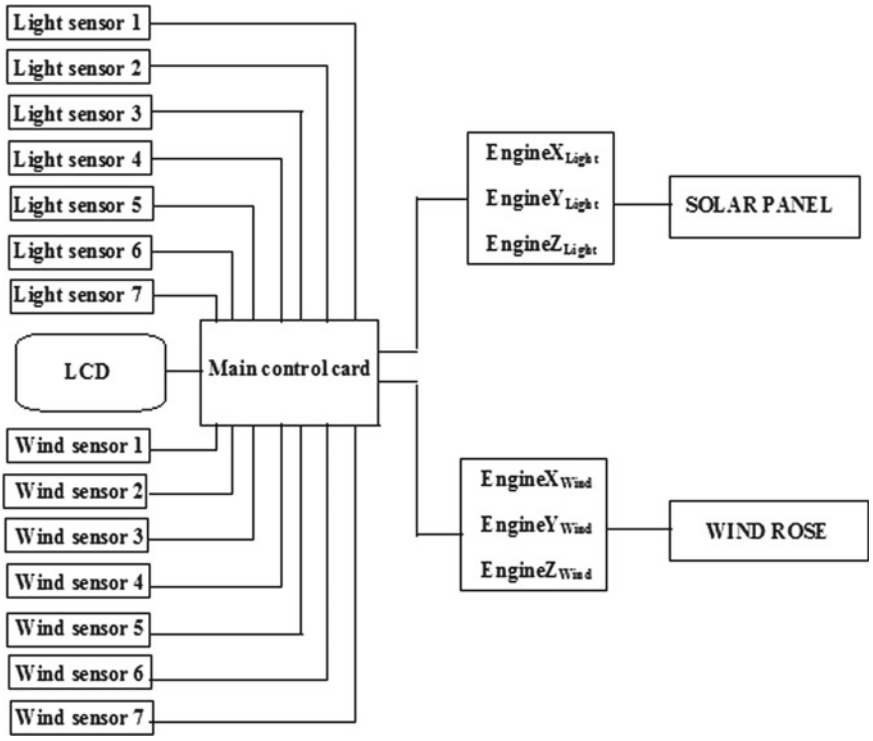


Fig. 18 Renewable energy system block diagram

one of the best ways to meet the electricity needs of remote small networks and small settlements from power generation and distribution centers. The most common is the wind-solar hybrid power generation system. The battery-powered wind-solar hybrid power generation system was installed. The established hybrid power generation system consists of a 5000 W wind turbine, a 300 W 3 solar panel, a battery group, a 1500 W hybrid charge control unit that allows these units to operate in harmony, and a 6.5 kW full sine wave inverter for alternating current consumers. In addition, the data of the energy produced in the hybrid power generation system can be monitored over the charge control unit at intervals of 10 min and recorded as data file by following the interface program and computer. Dynamic modeling was performed using SimPowerSystems, which is included in the Matlab/Simulink program, in accordance with the actual behavior of the hybrid power generation system. The output voltage of the power generation system is regulated at DC 24 volts with AC-DC and DC-DC power converters. A hybrid power generation system was established by parallel connectionincluding battery group connected to DC bar. The measurement station is located at 36.160996 latitude and 32.804022 longitude. According to the measurements made, the total amount of solar energy measured in Silifke District is 800 (kWh/m^2). The average amount of solar energy is 525 (kWh/m^2). The duration of

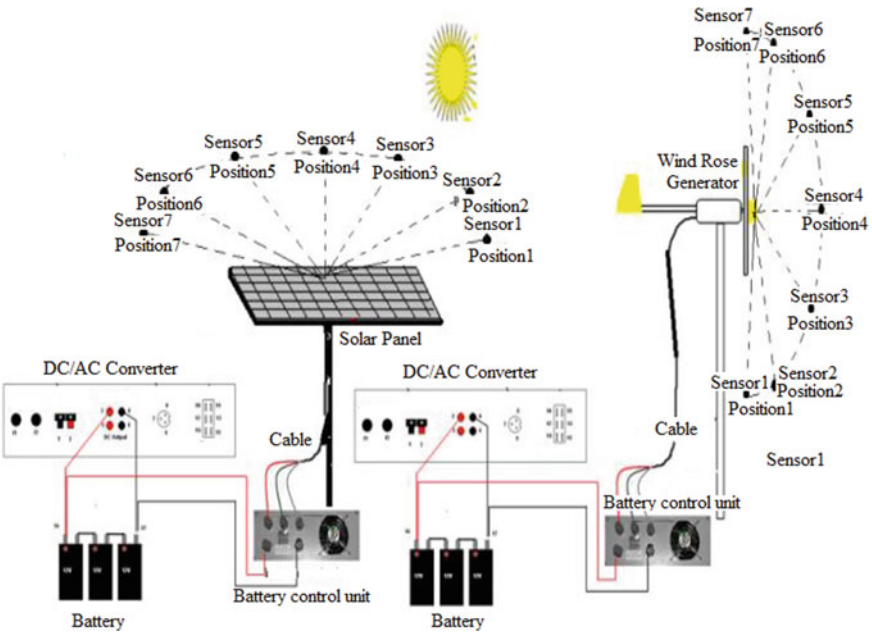


Fig. 19 Renewable energy system

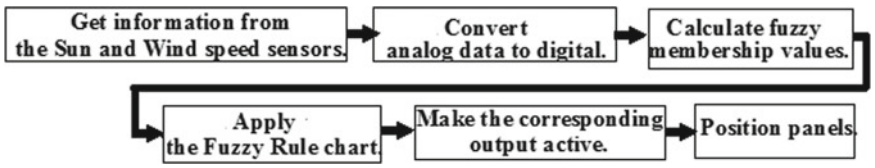


Fig. 20 System flow diagram

sunshine per year is 2950 (hours/year). Annual wind speed average value (km/hour) measured in Silifke District is 29 km/hour. The temperature (°C) measured in Silifke District during the year is shown in Fig. 22.

The maximum temperature measured is 43 °C. The amount of solar radition (kwh/m^2) measured in Silifke District during the year is shown in Fig. 23.

The wind speed values (km/hour) measured in Silifke District during the year is shown in Fig. 24. Established measurement station, solar radiation, wind speed and direction data were obtained for 12 months between June 2017 and June 2018.

Criteria such as the high area to be measured and the absence of high obstacles nearer than 10 times the height of the obstacle have been evaluated when the measurement station is selected. The positioning of sun/wind measurement stations and wind turbines is crucial for obtaining accurate measurements and for efficient energy production. The project planning and economics of the sun/wind power plant

Fig. 21 Renewable power plant created in Silifke-Turkey

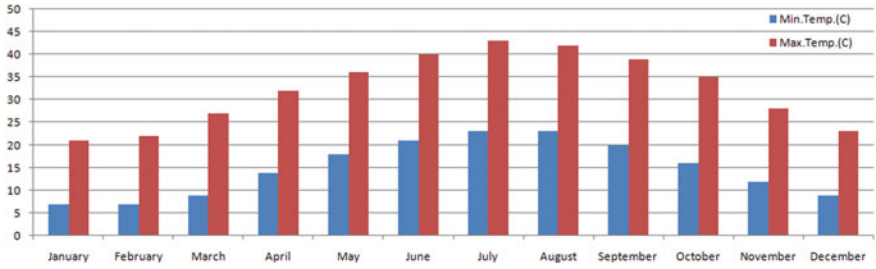


Fig. 22 Temperature (^{o}C) in Silifke District during the year

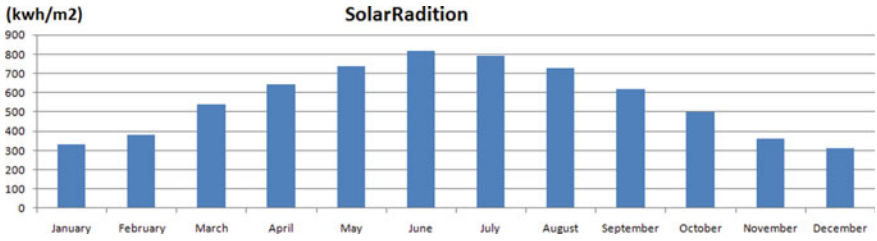


Fig. 23 The amount of solar radition (kwh/m^2) in Silifke District during the year

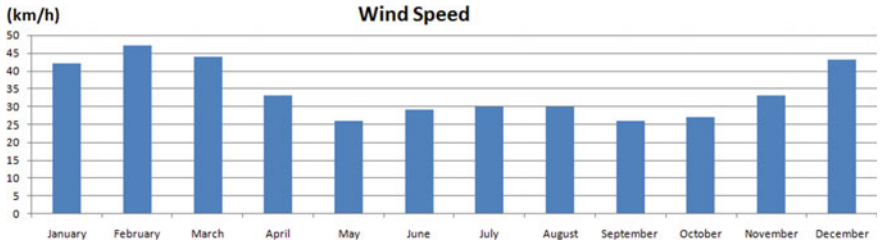


Fig. 24 The wind speed values (km/hour) measuredin Silifke District during the year

depends mainly on the amount of energy that can be generated from the selected plant site. For this reason, the appropriate spot sun/wind observation station should be established considering the tomographic structure of the land. The area where the station will be installed should be selected in areas where there is no obstacle to influence the dominant wind direction. Otherwise, the wind speed may be too far away from the true value due to turbulences. Measurements such as wind speed and wind direction made in accordance with the standards with measuring instruments placed on the station measurement pole will be used to evaluate the data that are later stages of the sun/wind power plant project, to determine the amount of energy production and to select the wind turbine. In particular, sensitivity to wind speed and direction for energy purposes is required to be high. The sun/wind energy potential of a zone. The wind speed is proportional to the cuboid. The measurement error, which can be made in wind speed measurements, is very influential on the accuracy of determining the wind energy potential. The sun/wind measurement data should be continuous, intermittent and incomplete as much as possible. The missing data in the measurement series can be up to 10% of the total data. The losses above this ratio do not allow a sensitive evaluation. So, the measurement values obtained should be canceled. Since station measurement equipment operates under completely natural atmospheric conditions have a shorter economic life than other equipment operating in closed areas. Due to factors such as rusting, corrosion, corrosion, abrasion and icing, wind measurement values of wind equipments are normally removed rapidly. Because of this, wind measurement equipment and recording systems and electrical circuits must be frequently checked, maintained and calibrated, absolutely in accordance with the standards and in the technical manuals of the instruments. The hybrid power plant, in which wind and solar energy are used jointly, was designed and applied. The block diagram of the applied system is given in Fig. 25.

In the Silifke District, 6.5 kW wind-solar hybrid power generation system with battery support was installed to feed dwellers. The realtime work to be done on the installed system carries great risks. Therefore, block diagrams of the system components have been created using Sim-Power-Systems in the Matlab/Simulink program. Simulink graphical design, simulation and analysis of dynamic systems have been realized with the Matlab program. This program supports either linear or nonlinear systems, including digital or both point signals. Easy to use, it is a worldwide accepted program for reasons such as testing the accuracy of the designed

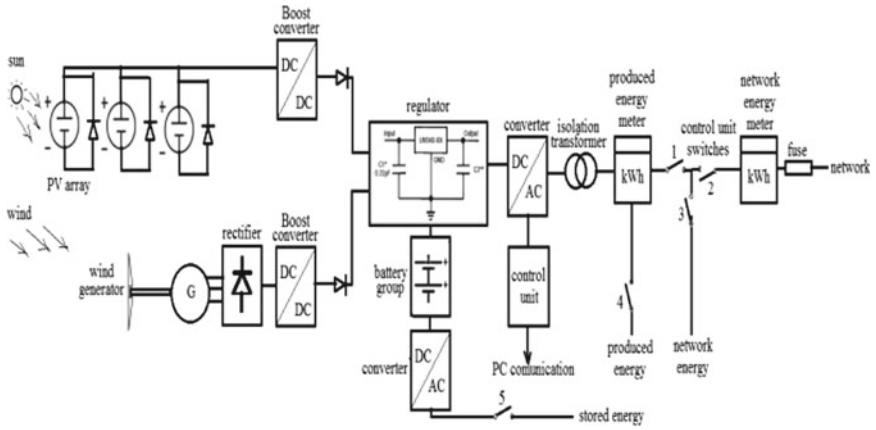


Fig. 25 Block diagram of the designed and control panel applied hybrid energy system

projects before they are put into practice, supporting the experimental results to be done. These block diagrams are combined to create an individual simulation model of the installed system. In addition, in order to increase the system efficiency, modeling is done by designing the system with energy flow control, switched controller and fuzzy logic controller. The solar energy system module is given in Fig. 26.

The voltage obtained from the wind system is also irregular. This can be shown as the greatest factor of the wind's constant speed. The AC voltage generated in the wind system is converted to DC voltage by an AC/DC converter. DC/DC converters are then used to reduce or increase the voltage according to the load requirements or the tension from the solar system. In addition, as DC is more advantageous to collect and transmit signals, both DC and DC signals are converted to DC signals in the wind system. The PV used in the system does not stay constant because of the electrical energy generated by the changing weather conditions, the system may overheat over time. Therefore, it is first necessary to rectify the energy obtained from the solar system with the help of a regulator. Some of this rectified voltage is stored with a

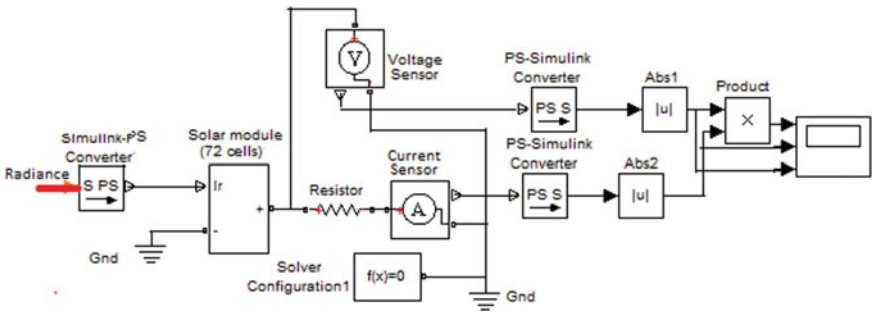


Fig. 26 Solar power generation system simulation block diagram

battery to meet the night energy need, if possible, if the power of the solar system can not meet the load requirement. The output voltage from the regulator is pulled to the desired level with the DC/DC converter. Thus, it is possible to adjust the voltage according to the need and the energy from the wind system. The generated DC signals are collected by means of a bar and given DC/AC invert to alternate current. The alternating voltage from the inverter is set via a busbar after it is set by the transformer. The fuzzy logic controller design is designed with 3 inputs and 5 outputs. The inputs are battery capacity, load power and hybrid power. These values are taken and applied to fuzzy inputs. Input values are processed according to the written rule base, outputs are determined, and keys are controlled by fuzzy. Membership functions have been determined primarily. Once we have defined entry membership functions, we can get the outputs we want by describing how the controller should work with 'and, or not' commands. In this study, the desired results were obtained by using 5 rules. These rules can be increased if desired. But too many rules slow down the controller speed. The fuzzy logic rule table is shown in Table 1.

The fuzzy logic software for solar/wind hybrid energy applications is given in Fig. 27.

It was determined that SilifkeDistrict is suitable for solar/wind hybrid energy applications in terms of solar radiation and wind potential as a result of efforts

Table 1 The fuzzy logic rule table

Number	Key control rule	S1	S2	S3	S4	S5	System states
1	$P_{Load} < P_{Produced\ energy}$	0	0	0	1	0	Production adequate, load on, battery charging
2	$P_{Network\ energy} = 0$ and $P_{Load} > P_{Produced\ energy}$	0	0	0	1	1	Production inadequate, load on, battery discharge
3	$P_{Network\ energy} > 0$ and $P_{Load} > P_{Produced\ energy}$	0	1	1	1	0	Production inadequate, mains and load on, battery discharge
4	$P_{Produced\ energy} = 0$ and $P_{Load} > P_{Stored\ energy}$	1	1	1	0	1	No production, mains and load on, battery discharge
5	$P_{Produced\ energy} = 0$ and $P_{Load} < P_{Stored\ energy}$	1	0	0	0	1	No production, load on, battery discharge

- If PL is L and PP is H then S1 is off and S2 is off and S3 is off and S4 is on and S5 is off else
- If PN is NA and PL is H and PP is L then S1 is off and S2 is off and S3 is off and S4 is on and S5 is on else
- If PN is A and PL is H and PP is L then S1 is off and S2 is on and S3 is on and S4 is on and S5 is off else
- If PP is NA and PL is H and PS is L then S1 is on and S2 is on and S3 is on and S4 is off and S5 is on else
- If PP is NA and PL is L and PS is H then S1 is on and S2 is off and S3 is off and S4 is off and S5 is on else
- ……

Fig. 27 The fuzzy logic software for solar/wind hybrid energy applications

to determine solar and wind energy potential. In Silifke District, the values to be obtained in wind and solar electricity production are examined. The measuring station is set up. At the measuring station, solar radiation intensity measurements were made at wind speeds of 5, 10 and 15 meters. With these measurements, firstly the potential determination work has been done. It is reached that it is a suitable zone for producing wind/solar hybrid energy. The Silifke District proved to be in a very advantageous position in terms of wind and solar energy potential. It is understood that solar and wind potential measurements are higher than solar ray and wind speed measurements made. It has been observed that even in winter the battery-powered solar-hybrid power generation system, which has been installed, can operate lamps and electric appliances in the room. Currently, small power generator systems are installed to minimize the effects of power failures. Thanks to this system, it has been determined that the interior lighting needs can be achieved by hybrid power generation systems, at least when there is a power failure in the mountain houses where the electricity lines do not reach and in the dwelling houses. Experiments on real systems have great risks for both system components and human life. Thus, an individual model of the battery-powered wind-solar hybrid power generation system that has been installed is done in the MatLab program. When the curves obtained from the simulation results of the modeling operation of the hybrid power generation system are taken into consideration, it is seen that there is not a great difference in the electrical and mechanical magnitudes parallel to the dynamic behavior of the hybrid power generation system installed. Generally generated power is sent directly to the battery group and the receivers are fed through the battery group in a hydropower hybrid power generation system. In this system, the power required by the receiver is supplied directly to the receiver and supplied to the receiver. If there is more power to produce, the battery is stored in the battery according to the occupancy rate of the battery. In this system, design of switching controller and fuzzy logic controller is done. The controllers designed for energy flow control seem to work very well. The battery group enters and exits the circuit depending on the behavior of the system. Thus, the life of the battery is prevented from decreasing. In fuzzy logic controller, desired control can be obtained by several rules. Therefore, it is found that it is more convenient than the switching control in terms of usage. My future work is to create different hybrid systems by adding additional power generation systems to the solar radiation and wind speed hybrid power generation system.

6 Sustainable Energy Management Model

Sustainability is defined as meeting the growing needs of the current generation without compromising the ability of future generations to meet their own needs. Sustainability; It is used to define the balance between economic growth, environmental protection and energy. It is possible to recover the existing energy potential with effective management. Energy consumption is directly proportional to economic, social and environmental sustainability. Today, it is very important to make

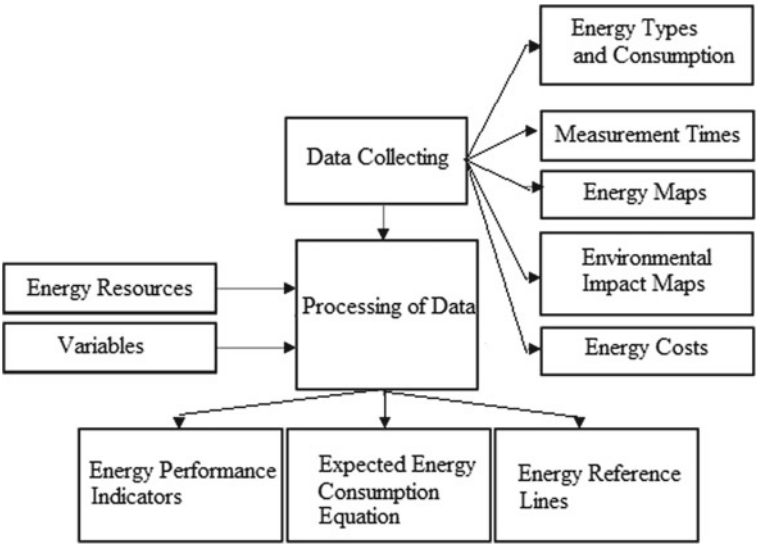


Fig. 28 Sustainable energy management model

energy management sustainable. Energy management should not only be considered economically. Preserving and maintaining the clean environment in energy investments has become a necessity today. Sustainable Energy Management Model is given in Fig. 28.

All data on sustainable energy planning should be collected and analyzed until results are obtained. The aim of sustainable energy planning is; energy reference line, energy performance indicators, objectives, objectives and action planning outputs. All previous and current data on energy use and consumption should be collected. The energy resources used and the equipment and processes in which these resources are used should be determined. In particular, identifying high and intensive energy consumption points is important in terms of defining the saving potential and improvement opportunities at these points. By analyzing energy data, firstly energy maps should be created and important energy users should be identified. When determining important energy users, the users with the highest share in total energy consumption should be selected. Since energy consumption requires financial transactions, it is not possible to continue energy consumption if this is not converted into profit in some way. One of the most important indicators of energy efficiency is the energy density value. Energy density refers to the energy spent per unit of production or service. This value is calculated by dividing the total energy consumption by the gross domestic product. A small value means that energy is used efficiently. Therefore, in our country, which is dependent on foreign energy, it is necessary to reduce the energy density without compromising on production quality, quantity and comfort. The environmental dimension of sustainability aims to reduce the negative effects of energy consumption on ecosystems and human

health. Greenhouse gas emissions, especially CO_2 from fossil fuel, are an important factor in global warming. Today, the issue of climate change, greenhouse gas emissions from fossil fuels is one of the most serious environmental problems. Energy consumption in the world is estimated to increase by 65% by 2050. This will create serious problems related to the depletion of energy resources and global warming issues. If the necessary measures are not taken to reduce CO_2 and other greenhouse gas emissions, the world's surface temperature will rise. This increase in temperature will harm the environment, human health and natural ecosystem.

7 Conclusion

Technological developments and population increase in the world are constantly increasing the need for energy. Each country aims to increase efficiency in energy production and use and to reduce dependence on foreign sources. Sustainable energy sources can also be considered as renewable energy sources. Renewable energy sources can be thought of as wind energy, solar energy, geothermal energy, which we are all familiar with. Energy efficiency plays a major role in the development of sustainable energy resources. Carbon dioxide emission has recently caused environmental and climatic damages. Therefore, a method has been sought for both to meet the increasing energy demand and to minimize these damages. The popularity of sustainable energy sources is also increasing through social formations. With the development of new technologies, it has even entered the private sector. Non-renewable energy sources carry risks in terms of sustainability. The use of renewable energy sources has become more widespread in recent years. The dependence on non-renewable energy sources will not be abandoned in the short and medium term. The dependency on non-renewable energy sources is still high among all energy sources worldwide. Electricity is generated from renewable energy sources. However, non-renewable energy sources can be utilized in different fields besides electricity generation. It is used in the production of electricity from natural gas, as well as in the heating and transportation sectors. With the developing technology, electric vehicles started to be used in transportation sector. The most widely used non-renewable energy source is oil. Nuclear energy used in electricity generation is another non-renewable energy source that is the most controversial issue in the world. The use of this energy source is risky, since it has caused plant accidents and deaths in the past and the waste generated after use cannot be completely destroyed. Huge amounts of electrical energy can be generated from very small inputs from nuclear energy. In addition, the amount of gases emitted to the environment is the lowest in all other energy sources. The majority of the energy used in the world from non-renewable sources can be prevented from transferring the resources to future generations in a sustainable manner. Instead of generating renewable energy sources from all over the world, efforts are being made to produce the same energy with alternative sources. With the development of technology, it has become possible to obtain energy production from renewable sources, and energy production has turned towards renewable

areas in recent years. Renewable energy sources are ready to meet our energy needs, which is one of the most basic needs of today and our future; These are open-source energy sources that may prevent the degradation of nature, such as the issue of carbon dioxide emissions, global warming and climate change.

References

1. Zile M (2018) Energy sources of the mediterranean region and current energy production, 1. Int Mediterr Symp 5:395–412
2. Park S, Miura Y, Ise T (2012) A maximum power control scheme based on multi-agent system for distributed flexible network photovoltaic system. In: 2012 International conference on renewable energy research and applications (ICRERA), pp 1–6
3. Leong Hai K, Yen Kheng T, Peng WKJT (2009) Renewable energy integration into smart grids: problems and solutions—Singapore experience. IEEE/ASME Int Conf Adv Intell Mechatron AIM 9(12):1–7
4. Bird L, Milligan M (2013) Integrating variable renewable energy: challenges and solutions
5. Zile M (2016) Mersin-Silifke district of solar power plants places could assembly the evaluation of installation and operating cost. Solar Conf Exhib 615–622
6. Chowdhury AH, Chowdhury N, Longo M, Yaïci W (2019) System and cost analysis of stand-alone solar home system applied to a developing country. Sustainability 11(5):1403
7. Zile M (2008) Generating electricity from the sun, projecting and improving different practics. Mersin Symp
8. Zile M (2005) A new energy system with solar and wind sources. III. Renew Energy Res Symp, Yeksem
9. Corona B, Ruiz D, San Miguel G (2016) Life cycle assessment of a HYSOL concentrated solar power plant: analyzing the effect of geographic location. Energies, MDPI
10. Amir F, Maysam M, Zaki A, Intesar A (2013). Aluminium alloys in solar power. https://doi.org/10.5772/54721
11. Zile M (2018) Implementation of solar and wind energy by renewable energy resources with fuzzy logic. Int J Tech Phys Probl Eng (IJTPE) 10(34):46–51
12. Zile M (2018) Hybrid energy agricultural irrigation in Mersin region, 1. Int Mediterr Symp 5:389–394
13. Zile M (2018) Determination of the hybrid energy source potential with measurements in Mersin and its region, 1. Int Mediterr Symp 5:373–378
14. Bray JW, Fair R, Haran K (2014) Wind and ocean power generators. IEEE Trans Appl Supercond. https://doi.org/10.1109/TASC.2013.2282698
15. Zile M (2018) Determination and analysis of wind speed and solar radiation potential for energy production in Aydincik District of Mersin Province. In: Cilicia international symposium on engineering and technology, Ciset, vol 52, pp 184–189
16. Zile M (2018) Wind and solar energy hybrid power generating system installation in anamur district and fuzzy logic based energy flow control. Cilicia international symposium on engineering and technology CISET, vol 52, pp 98–102
17. Zile M (2013) Integration of solar and wind power plants into smart grids for tarsus district. In: Smart grid symposium on the future of electricity networks
18. Dixon C, Mahajan R, Agarwal S, Brush AJ, Lee B, Saroiu S, Bah, P (2012) An operating system for the home. In: Proceedings of the 9th USENIX conference on networked systems design and implementation, pp 25–25
19. https://psapublishing.com/Hydroelectric.pdf. Accessed 20 Aug 2019
20. Fulzele JB, Dutt S (2002) Optimium planning of hybrid renewable energy system using homer. Int J Electr Comput Eng 2(1):68–74

21. Wei Z, Chengzhi L, Zhongshi L, Lin LHY (2014) Current status of research on optimum sizing of stand-alone hybrid solar-wind power generation systems. Res J Appl Sci Eng Technol 8(14):1684–1690
22. Kabbara S, Moubayed N (2010) Hybrid wind/PV system: The lebanese case. Ann Univ Craiova Electr Eng Ser 34:241–244
23. Hidayatullah NA, Paracha ZJ, Kalam A (2011) Impact of distributed generation on smart grid transient stability. Smart Grid Renew Energy 02(02):99–109
24. Suzuki D (2016) Geothermal energy better than site C, The Rossland Telegraph
25. Farzad S, Mandegari MA, Görgens JF (2016) A critical review on biomass gasification, co-gasification, and their environmental assessments. Biofuel Res J (BRJ). https://doi.org/10.18331/brj2016.3.4.3
26. https://www.greencitytimes.com/hydroelectricity. Accessed 20 Aug 2019
27. https://www.aist.go.jp/fukushima/en/unit/HyCaTe.html. Accessed 20 Aug 2019
28. Maegaard P (2009) Wind energy development and application prospects of non-grid-connected wind power. IEEE Xplore. https://doi.org/10.1109/WNWEC.2009.5335847
29. Wong JKW, Zhou J (2015) Enhancing environmental sustainability over building life cycles through green BIM: A review. Autom Constr 57:156–165
30. Schachinger D, Kastner W, Gaida S (2016) Ontology-based abstraction layer for smart grid interaction in building energy management systems. Institute of Computer Aided Automation TU Wien, Vienna, Austria
31. Gasho EG (2016) Methodical recommendations on calculation effects from the implementation of measures for energy conservation and energy efficiency, State guidelines, Analytical Center under the Government of the Russian Federation, Russian Federation
32. Gasho EG, Guzhov SW, Krolin AA (2016) Methodical recommendations on calculation effects from the implementation of measures for energy conservation and energy efficiency, State normative document, Ministry of Construction and Housing and Communal Services of the Russian Federation, Russian Federation

Modeling of Energy Systems for Smart Homes

Hossein Shayeghi and Masoud Alilou

Abstract In the last years, smart homes have been introduced for improving the lifestyle of people. The energy system of a smart home is similar to the power network because both consumer and producer types of devices exist in the smart home. This complexity causes that manually managing the smart home becomes more difficult than a traditional home. Knowing the energy system of the smart home and automatically managing of its devices are important in increasing the efficiency of the smart home and the welfare of consumers. For this reason, analytical and mathematical modeling of devices of the smart home is investigated in this chapter. Modeling of the producer and consumer devices of the smart home's energy system is presented analytically and mathematically. Moreover, an optimization algorithm is also presented to ponder the proposed model of the power system and select the optimal schedule of devices for having the highest operation of the energy system of smart homes. A technical-economic objective function is considered for finding the best schedule of devices. Ultimately, the proposed method is simulated in a sample smart home for evaluating the model of devices and energy system. Ultimately, the efficiency of the proposed energy system of the smart home is pondered based on the simulation results.

Keywords Analytical and mathematical modeling · Energy system · Renewable distributed generation · Smart appliances · Smart home

Nomenclature

I_t	Mode of appliance i at period t
n_p	Number of the time interval during the day

H. Shayeghi (✉) · M. Alilou
Energy Management Research Center, University of Mohaghegh Ardabili, Ardabil, Iran
e-mail: hshayeghi@gmail.com

M. Alilou
e-mail: masoud.alilou@yahoo.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_12

p_s	Starting time interval of appliance i
p_e	Ending time interval of appliance i
OP_i	Proper number of operating time intervals of appliance i
K	The shape coefficient of Weibull
C	The scale coefficient of Weibull
V	Wind speed
V_{ci}	Cut-in speed of wind
V_{co}	Cut-out speed of wind
V_r	Rated speed of wind
$P_{w-rated}$	Rated output power of the WT
Si	The solar irradiance
η	Efficiency of PV
A	Total area of PV
I_{EP}	Electricity bill of the smart home
I_{PD}	Peak demand of the smart home
C_{PN}	Cost of purchased energy from the power network
n_a	Number of appliances
pa_i	Consumption of each appliance at each period
P_{Ap_t}	Total consumption of appliances at period t
n_{type}	Technology of the DG unit
n_{DG}	Number of considered type of DGs
P_{DG_t}	Total produced power of renewable DGs at period t
Tr_t	Electricity tariff at period t
pr_t	Electricity price of sold back to the grid at period t
pr_{DG_i}	Cost of producing 1 KWh energy by renewable DG unit
P_{App_t}	Total consumption of all appliances of SH at period t
P_t^{APP}	Demand of appliances
P_t^{SH2PN}	Sold energy to the power network
P_t^{PN2SH}	Purchased energy from the distribution system
P_t^{RDG}	Produced power of renewable DG units
P_{RDG}^{min}	Minimum level of producing the power of DG
P_{RDG}^{max}	Maximum level of producing the power of DG

1 Introduction

Nowadays, the imagination of the world without electrical energy is approximately impossible. This dependence causes the power network becomes more complex over time in order to respond to the demand of consumers. On the other side, the progress of technology improves the performance of old devices of the energy system and also adds new devices to this network. All the mentioned issues increase the complexity of energy systems. Analytical and mathematical modeling of devices of the system is one of the useful methods for proper recognizing the devices and better managing the

energy system. Modern homes, which are known as smart homes, are a power network in the small size. For this reason, the energy model of the smart home (SH) is so similar to the model of the energy system according to the consumption and production sides. A smart home has the consumer and producer devices the same as an energy network. In the smart homes, appliances like refrigerator and lights consume electricity while domestic renewable energy sources like wind turbine and photovoltaic panel produce green energy. Renewable units produce electricity based on the variation of their initial energy such as wind speed and solar irradiance; therefore, it is better that a stochastic formulation is considered for modeling renewable distributed generation units due to the stochastic behavior of their initial energy [1–3].

The concept of a smart home is a fundamental combination of different subsystems related to automatically control of home devices through advanced technologies. It can both share generation units and transfer within the home, and the operator can exchange data with your SH external network through your home smart gateway. Its base purpose is to supply people with an efficient, easy, secure, useful and eco-friendly living environment integrating system, service and management. Devices of a smart home are connected with each other and also are controllable through one central point such as home energy management system (HEMS). HEMS of the smart home observes and manages all appliances in real-time, regards to predetermined constraints and user's preferences, for decreasing electricity bill, decrease the dependence of the SH the power network and improve energy utilization efficiency. The sample diagram of a SH is demonstrated in Fig. 1 [4, 5].

As mentioned above, the management of a smart home has an effect on the performance of that. In this regard, some researchers have studied energy sources, appliances and operational schedule of smart homes. For example, In Ref. [6], the alternating current and direct current grid arrangement of smart homes has been proposed for improving the used penetration of distributed generation units and reducing the dependence of the Microgrid on the upper network. In this paper, devices of SHs in a micro grid have been managed. Moreover, both grid-connected and off-grid modes of Microgrid have been evaluated in this study. An energy management framework has been proposed for optimizing the operational schedule of home devices in Ref. [7]. This optimization has been done in two steps. In the first step, each consumer minimizes his pay charge. In the second step, smart distribution company minimizes deviation of distribution system load and the cost of modifying the desired scheduling of consumers. In another study, the demand side management has been done in a SH with photovoltaic panel and battery [8]. In this chapter, management charging and discharging of storage system is the main purpose which is solved by a methodology based on artificial neutral network. The availability of solar irradiance and the cost of purchasing power from the power network have been considered for better managing the operational time of battery. In another study, authors have studied the residential load commitment framework for achieving the minimum electricity bill for consumers [9]. The operating schedule of appliances and also the charging time and the discharging time of battery and electric vehicles has been scheduled in this paper.

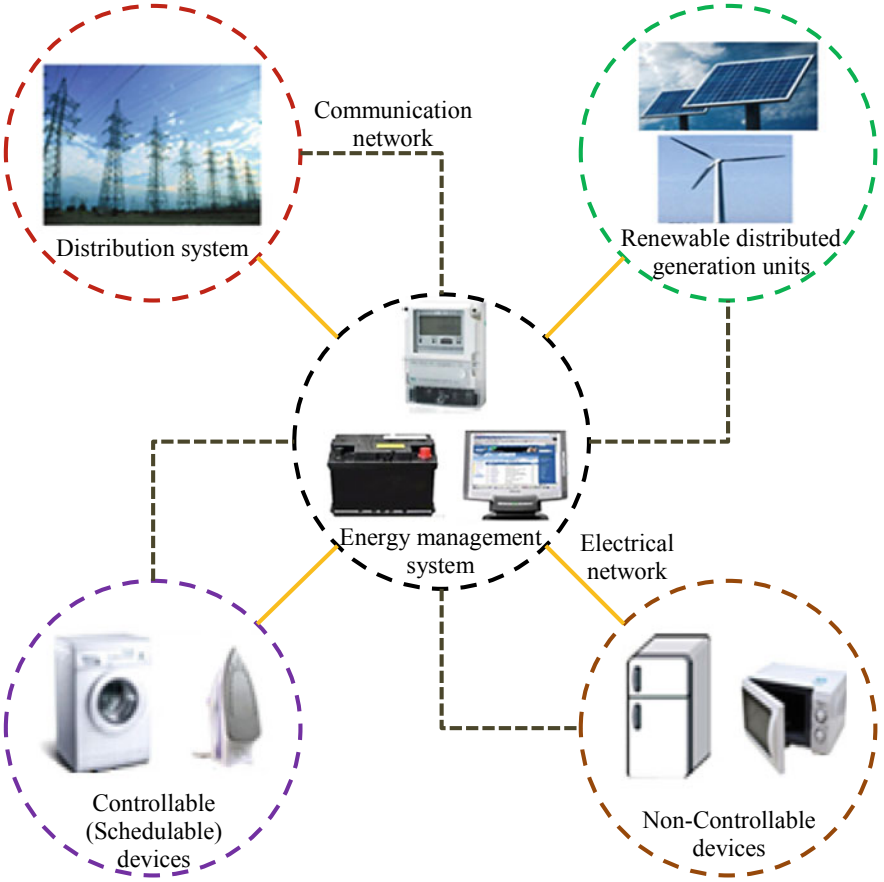


Fig. 1 Overall diagram of a smart home

Therefore, the advantages of properly managing smart homes have been shown in some papers. For completing the previous studies, this chapter focuses on the analytical and mathematical modeling of the energy system of a smart home. An optimization algorithm is also utilized to ponder the proposed model of the energy system and find the highest efficiency of devices for having the best operation of the energy system of SH. Firstly, the total introduction of energy systems is presented and then modern smart homes are introduced and their energy system is described. Later, modeling of the producer and consumer devices of the smart home's energy system is presented analytically and mathematically. Then, a new method based on intelligent algorithms is introduced for optimizing the hourly performance of home appliances and renewable energy sources with regard to stochastic behavior of renewable units for decreasing the electrical cost of the SH and improve the efficiency of the energy system. Ultimately, the proposed method is simulated in a sample smart home for

evaluating the model of devices and energy system. Ultimately, the efficiency of the proposed energy system of the smart home is pondered based on the simulation results.

2 Energy Systems

A system is a group of interrelated or interacting entities which work together for getting common purposes. This feature can be seen clearly in energy systems. An energy system, which consists of many parts and devices, is utilized to deliver the electricity to the consumers. The energy system of electrical energy is the combination of generation, transmission and distribution systems so that each of these systems also consists of some subsystems for better delivering the electrical power to the costumers of the energy system.

2.1 Smart Homes

A smart home has the energy system in a small size. Of course, the transmission system doesn't exist in the energy system of smart homes. In other words, both types of producer and consumer devices are in a smart home while only consumer appliances were in traditional homes. In smart homes, the demand of controllable and non-controllable appliances is supplied using the produced power of renewable energy sources and the bought power from the upstream power system.

The optimal management of the operating schedule of appliances and renewable energy sources of smart homes has high effect on the technical and economic parameters of homes. Moreover, the environmental indices are also improved by reducing the dependence of the smart home on the produced power of non-renewable sources.

The HEMS has the responsibility to optimal manage the devices of smart home based on the predetermined constraints and user's preferences. The HEMS finds the best schedule of energy sources and appliances of the smart home so that for improving the efficiency, economics, reliability, and energy conservation of both the smart home and the distribution system.

The HEMS optimizes the best schedule of devices using the predetermined parameters, the condition of market price and the situation of the produced power of wind turbine and photovoltaic panel. The modeling of domestic energy sources and home appliances of the smart home are defined in the HEMS. Moreover, main purposes, constraints and the optimization method are also predetermined in the HEMS based on the user's preferences. The required modeling, objective functions, constraints and optimization algorithm for better managing the energy system of the smart home are explained completely in the next sections.

3 Modeling of Smart Home Energy Devices

In this section, the modeling of home appliances and renewable energy sources are explained analytically and mathematically.

3.1 Home Appliances

One of the main differences between a smart home and a traditional home is the ability to control the appliances of the smart home in order to decrease the consumer's costs and improve the performance of the home's energy system. Appliances are connected to the HEMS of the smart home using the internal technology or another external device.

Smart appliances can be divided into controllable and non-controllable devices. The controllable appliances are the devices which can participate in the demand-side management program and can be scheduled in 24-hour based on the situation of market price and domestic energy sources. On the other side, non-controllable appliances are non-schedulable devices so that their operational time is constantly based on the user's predetermined schedule.

- **Non-controllable appliances**

In the smart home, some appliances cannot participate in the energy management program because their operating schedule is predetermined by the user's preference and lifestyle. This type of devices is called non-controllable or non-schedulable appliances. Samples of this type of appliances can be shown in Fig. 2.

- **Controllable appliances**

Unlike non-controllable devices, some appliances of the smart home participate in the home energy management program for improving the performance of electricity and decrease the electricity bill of the consumer. This type of appliances is called controllable or schedulable devices. The operating time of manageable devices is managed based on the situation of the electricity cost, weather and other predetermined constraints. Figure 3 shows samples of controllable appliances in smart homes.

- **Modeling**

In this chapter, the time-varying model of appliances is considered. In other words, the starting time of appliances is only optimized by the proposed method while the consumption power of devices is constant at each time interval of the operating period. It means that the demand profile of appliances should not be changed by the HEMS; otherwise the considered device operates incorrectly. For instance, if the power of the oven at period t is reduced by the energy management program, the

Fig. 2 Samples of non-schedulable appliances

temperature of the oven doesn't reach the proper degree for baking a delicious Apple pie. Hence, only the starting time of the appliances is optimized by the HEMS.

Therefore, the home energy management system selects the best schedule for the operation time vector of appliances. Equation (1) shows the daily state vector of each appliance in the smart home.

$$I_i = \left[I_1, I_2, \ldots, I_t, \ldots I_{T-1}, I_T\right] \tag{1}$$

Here, I_t is the mode of appliance i at time-interval t; the amount of this parameter is 1 or 0 that it is equal to on/off of appliance i at time-interval t. All devices should be executed during the day; this constraint is formulated in Eq. (2).

$$\sum_{t=1}^{n_p} I_{it} \geq 1 \tag{2}$$

Moreover, each appliance has the number of operating periods. In other words, each device needs to run for consecutive periods without interruption when it is started. Equation (3) demonstrates this constraint.

Fig. 3 Samples of schedulable appliances

$$\sum_{t=p_s}^{p_e} I_{it} = OP_i \tag{3}$$

As regards to this constraint, the starting time of devices should be selected based on Eq. (4).

$$p_s \leq n_p - OP_i + 1 \tag{4}$$

Therefore, the operating schedule of appliances should be based on Eqs. (1)–(4). Although the starting time of non-controllable appliances is predetermined, the operating time intervals and the ending time of devices are selected using these equations.

3.2 Home Renewable Energy Units

Wind turbines and photovoltaic panels are the most useful technologies of distributed generation units that are used in smart homes. Although renewable energy sources are Eco-friendly, the intermittent nature of the initial energy of those should be considered

for calculating the produced power of renewable DG units. So, the uncertain speeds of wind and illumination intensity of the sun are the main parameters that have effect on the output power of WT and PV, respectively.

In this chapter, a stochastic method is used to compute the amount of uncertain parameters of wind turbine and photovoltaic panel. The hybrid method of Latin hypercube sampling algorithm and K-means clustering algorithm is utilized to find the conserved stochastic data [10].

In the following, the analytical and mathematical modeling of home renewable energy sources is presented [11, 12].

- **Wind turbine unit**

In most countries which are the pioneer in the clean energy, wind turbine (WT) is one of the common and useful renewable DG technologies. WT converts wind energy into electrical energy. In smart homes, the small size of wind turbines is utilized to supply the demand of the home. Figure 4 shows the wind turbines which is used in the residential areas.

Totally, the wind turbine has a stochastic performance. So it is better that a probabilistic model is simulated according to available historical data. The meteorological

Fig. 4 Wind turbines of residential areas

data is the best reference for estimating the wind energy potential of a special site. Regards to wind speed behavior and a large number of experiments, Rayleigh probability density function is utilized to model wind speed. Rayleigh distribution is a particular form of Weibull probability density function in which the shape index is equal to 2 [11].

Mathematically, the Weibull distribution of wind speed is presented by Eq. (5).

$$f(v) = \frac{K}{C^K} v^{(K-1)} e^{-\left(\frac{v}{C}\right)^K} 0 \leq v \leq \infty \tag{5}$$

Thus, the equation between the electrical power of a WT and the wind speed can be formulated by Eq. (6).

$$P_w = \begin{cases} 0 & 0 \leq V \leq V_{ci}, V_{co} \leq V \\ P_{w-rated} \times \frac{V - V_{ci}}{V_r - V_{ci}} & V_{ci} \leq V \leq V_r \\ P_{w-rated} & V_r \leq V \leq V_{co} \end{cases} \tag{6}$$

According to Eq. (6) can be said that a wind turbine can generate energy when the wind speed is more than the minimum wind speed. The wind turbine produces electrical power until the wind speed is lower than the maximum limit.

- **Photovoltaic panel**

Photovoltaic (PV) is one of the popular and useful technologies of renewable DGs. PV panels can be located on the roof, in the yard and in the backyard of smart homes. Samples of utilized PV panels in smart homes are demonstrated in Fig. 5.

Although the area and efficiency of the solar panel have an effect on the produced power of PV panel, the electrical power of a PV unit is predominantly affected by the illumination intensity. So, the PV panel has also a stochastic behavior due to uncertain characteristic of solar power [12].

According to the last studies and experiments, it is shown that the stochastic behavior of solar power can be simulated by the Beta probability distribution function. Therefore, the equation between solar irradiance and the electrical power of a PV unit can be described by Eq. (7) [12].

$$f(Si) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times Si^{(\alpha-1)} \times (1 - Si)^{(\beta-1)} & if \;\; 0 \leq Si \leq 1, \alpha \geq 0, \beta \geq 0 \\ 0 & otherwise \end{cases} \tag{7}$$

Here,

$$\beta = (1 - \mu) \times \left(\frac{\mu \times (1 + \mu)}{\sigma^2} - 1 \right) \tag{8}$$

$$\alpha = \frac{\mu \times \beta}{1 - \mu} \tag{9}$$

Fig. 5 Samples of PV panels in smart homes

where μ and σ are the mean and standard deviation, respectively; they are extracted from the random variable. After calculating the probability distribution function of the solar irradiance, the electrical power of a PV is calculated by Eq. (10).

$$P_S(Si) = \eta \times A \times Si \tag{10}$$

4 Optimizing Method of Smart Home Devices

In this section, objective functions, constraints and the optimization algorithm for selecting the optimal schedule of smart home's energy sources and appliances are described completely.

4.1 Objective Function

Minimizing the electricity bill and the peak demand of the SH is considered as the main purpose of the home energy management program. Therefore, optimizing the operational schedule of devices is a multi-objective problem. Mathematically, the main purpose of the optimization is presented by Eq. (11).

$$Objective_{function} = min\{I_{EB}, I_{PD}\} \tag{11}$$

- **The Electricity bill of the smart home**

The HEMS provides the electricity power of smart home by using of purchased energy from the power network and home energy sources. Renewable distributed generation units are home energy sources. The produced power of the WT and the PV is variable based on weather and hour.

Therefore, the HEMS of the smart home schedules the hourly performance of appliances and home energy sources so that the total consumer's electricity bill becomes minimal.

The total electricity bill of the smart home is formed from the following items:

1. Cost of bought electricity from the power network
2. Cost of the produced electricity of wind turbine and photovoltaic panel
3. Income from selling the electricity to the power network

Consequently, the electricity bill of the smart home is formulated based on the costs and incomes in Eq. (12).

$$I_{EB} = C_{PN} + C_{RDG} - I_{PN} \tag{12}$$

C_{PN} is the cost of purchased energy from the power network. The consumer purchases his demand from the distribution-company with a variable price. As regards that there are renewable distributed generation units the smart home, a part of electricity demand is supplied by home energy sources and so the extra power is purchased from the distribution system. So, firstly, the total demand of all appliances at time-interval t is calculated by Eq. (13).

$$P_{Ap_t} = \sum_{i=1}^{n_a} pa_i \tag{13}$$

Then, Eq. (14) is utilized for computing the hourly production power of renewable distributed generation units.

$$P_{RDG_t} = \sum_{i=1}^{n_{type}} \sum_{j=1}^{n_{DG}} pd_{ij} \tag{14}$$

Consequently, Eq. (15) is utilized for calculating the hourly extra demand of SH that it is supplied by the power network.

$$P_{PN_t} = P_{Ap_t} - P_{RDG_t} \tag{15}$$

Finally, the daily cost of energy purchased from the distribution system is expressed in Eq. (16).

$$C_{PN} = \sum_{t=1}^{n_p} P_{PN_t} \times Tr_t \quad if \quad P_{PN_t} > 0 \tag{16}$$

In this equation, Tr_t is the electricity tariff at period t which is predetermined by the independent system operator (ISO). It is worth mentioning that the length of each period is considered one hour; therefore, the total number of time-intervals in the day is 24.

As can be seen in Eq. (16), the smart home purchases energy from the distribution system when the consumption of appliances is more than the generation of home energy sources at the considered period. On the other side, the smart home sells power to the network at some periods when the generation of home energy sources is more than the consumption of appliances. Therefore, the income from selling the electricity to the power network is formulated in Eq. (17).

$$I_{PN} = \sum_{t=1}^{n_p} -1 \times P_{PN_t} \times pr_t \quad if \quad P_{PN_t} < 0 \tag{17}$$

Here, the amount of pr_t is predetermined between ISO and consumer for various time-intervals.

The cost of the produced energy of renewable distributed generation units is presented in Eq. (18).

$$C_{RDG} = \sum_{t=1}^{n_p} \sum_{i=1}^{n_{type}} \sum_{j=1}^{n_{DG}} pd_{ijt} \times pr_{DG_i} \tag{18}$$

where, pr_{DG_i} is the cost of the produced 1 KWh energy by renewable DG unit; the costs of investment, maintenance and operation of DGs are merged in this parameter.

- **The peak demand of the smart home**

This index shows the peak consumption of the SH. This index, which is a technical index of the smart home, demonstrates the maximum hourly purchasing power of the smart home from the power network during the day. Equation (19) is utilized to calculate the peak demand of the smart home.

$$I_{PD} = \max_t \left\{ P_{App_t} \right\} \tag{19}$$

4.2 Constraints

In this chapter, the power balance constraint and renewable distributed generation units' constraint are considered in optimizing of the operational schedule of devices of the smart home.

- **Power balance constraint**

This constraint is presented at period t by Eq. (20).

$$P_t^{PN2SH} + P_t^{RDG} = P_t^{APP} + P_t^{SH2PN} \tag{20}$$

This equation states that the demand of appliances (P_t^{APP}) and the sold energy to the power network (P_t^{SH2PN}) are satisfied by the purchased energy from the distribution system (P_t^{PN2SH}) and the produced power of renewable DG units (P_t^{RDG}).

- **Renewable DG constraint**

Renewable energy resources including wind turbine and photovoltaic can produce electrical energy in the following range:

$$P_{RDG}^{min} \leq P_{RDG} \leq P_{RDG}^{max} \tag{21}$$

4.3 Intelligent Algorithm

In this research, the combination of multi-objective grey wolf optimization (MOGWO) and fuzzy method is utilized to multi-objective optimize the operating

scheme of smart home's appliances and energy sources. For getting the best result, firstly, the MOGWO is utilized to optimize the multi-objective function and create the optimal Pareto-front. After applying the MOGWO algorithm, the fuzzy method is used to find the optimal particle from the non-dominated particles.

- **Multi-objective grey wolf optimization algorithm**

Intelligent algorithms are usually inspired from existing natural behaviours of nature. The optimization method of MOGWO is inspired from the grey wolves. Grey wolves which belong to Canidae family are considered as apex predators. This means that they are at the top of the food chain. Grey wolves mostly prefer to live in a pack with a group size [13].

The MOGWO is based on social behavior of Wolves. In this meta-heuristic algorithm, the best solution is considered as α wolf. β and δ wolves are the second and third solutions, respectively. The rest of particles are assumed as ω wolves. The optimization method of the MOGWO is managed by alpha, beta and delta so that the omega wolves follow α, β and δ wolves for reaching the best result.

So in this algorithm, the position of each particle is updated by [13]:

$$X(t+1) = X_P(t) - A.D \tag{22}$$

where,

$$D = |C.X_P(t) - X(t)| \tag{23}$$

$$A = 2a.r_1 - a \tag{24}$$

$$C = 2r_2 \tag{25}$$

Here, X_P and X show the position vector of the prey and the position vector of a grey wolf, respectively. Vectors of A and C are the coefficients. Moreover, a linearly decrease from 2 to 0 over the iterations while vectors of r_1 and r_2 are random in [0 and 1].

The extra detail of the MOGWO algorithm is available in Ref. [13].

- **Fuzzy method**

After optimizing the technical-economic issue of energy management of the smart home, the fuzzy decision-making method run to choose the optimal compromise solution which represents the optimal amount of technical and economical indices equal to the best schedule of smart home energy sources and appliances.

In the fuzzy decision-making method, the best result is selected in two steps. Firstly, membership values of each particle based on various objective functions are calculated by Eq. (26) [14].

$$\mu_i^k = \begin{cases} 1 & F_i^k \le F_i^{min} \\ \frac{F_i^{max} - F_i^k}{F_i^{max} - F_i^{min}} & F_i^{min} < F_i^k < F_i^{max} \\ 0 & F_i^{max} \le F_i^k \end{cases} \tag{26}$$

$$\mu^k = \frac{\sum_{i=1}^{NO} \mu_i^k}{\sum_{k=1}^{NK} \sum_{i=1}^{NO} \mu_i^k} \tag{27}$$

Then, Eq. (27) is utilized to calculate the total membership value of each non-dominated particle which is in the Pareto-front. In these equations, F_i^{min} and F_i^{max} demonstrate the lower and upper bound of index i, respectively. F_i^k shows the amount of particle k based on the objective function i.

Finally, each particle which has the highest amount of total membership value is selected as the best compromise solution [14].

Consequently, the complete method for managing the daily energy of the smart home is demonstrated in Fig. 6. This figure presents that firstly, details of home appliances, wind turbine unit, photovoltaic panel and economic parameters are inputted into the HEMS. Then, the HEMS produces initial random particles for utilizing the intelligent algorithm and calculating the stochastic amount of renewable DG units. Particles are improved by applying MOGWO with considering objective functions and constraints. Finally, HEMS utilize the fuzzy decision-making method for selecting the best particle equal to the optimal daily schedule of home appliances, WT and PV.

5 Numerical Results

In the previous sections, the modeling of home's appliances and energy sources are analytically and mathematically described. In this part, the proposed method for energy managing the smart home is pondered using the evaluation of the considered technical and economic objective functions.

It is considered that the smart home purchases energy from the grid with a variable tariff based on the time-of-use demand response program which is defined by the independent system operator. Figure 7 shows the hourly price of the electrical energy. Moreover, the distribution company buys the extra produced energy of the smart home with the tariff of 0.129 $/KWh. The price of the produced power of the wind turbine and photovoltaic are also determined at 0.075 and 0.063 $/KWh, respectively. The considered cost of renewable energy units is an overall cost which is the combination of the costs of installation, maintenance and operation of these units.

As mentioned above, the smart home energy management can utilize the wind turbine and photovoltaic for providing the demand of the smart home. The installed renewable sources are the PV of 335 W SolarPower X21 and the WT of 1500 W WINDMILL. The rated illumination intensity of the PV is 1000 W m^{-2}, and the shape and scale indices are 1.8 and 5.5, respectively. Moreover, the cut-in, normal

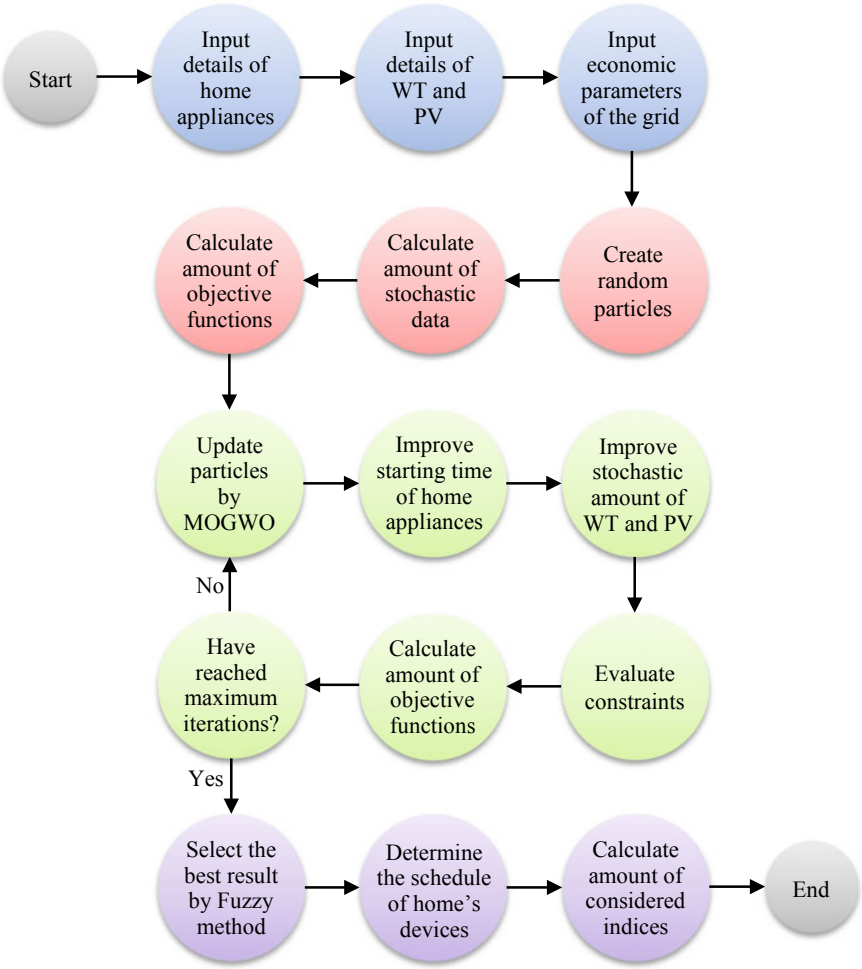


Fig. 6 Flowchart of the energy management of the smart home

and cut-out speeds of the WT are 4, 13 and 20 m/s. The shape and scale indices are 2 and 6.5, respectively. The maximum daily wind speed and solar irradiance in the area, which the sample smart home is located, are 8.5 m/s and 1 kW/m^2, respectively. The hourly variations of stochastic parameters in Pu are presented in Fig. 8 [15, 16].

The appliances of the smart home are divided into controllable and non-controllable devices. Washing machine, dishwasher, boiler, vacuum cleaner and Iron are considered as controllable appliances while refrigerator, purifier, lights, microwave oven, oven and TV are non-controllable devices of the Home.

The number of time-intervals of the day is considered equal to 24. Figure 9 and Fig. 10 show the demand of controllable and non-controllable appliances in their operational period, respectively [17]. It is worth mentioning that each appliance

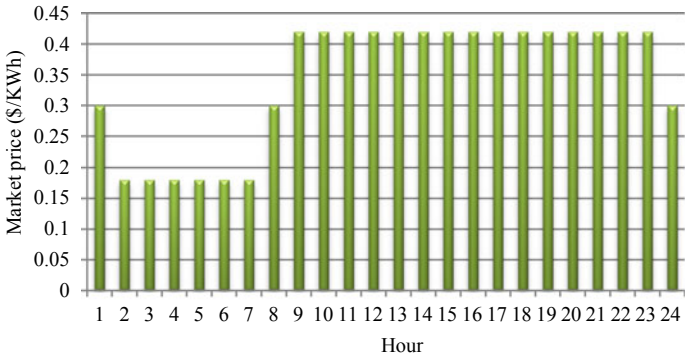


Fig. 7 Hourly market price of the grid

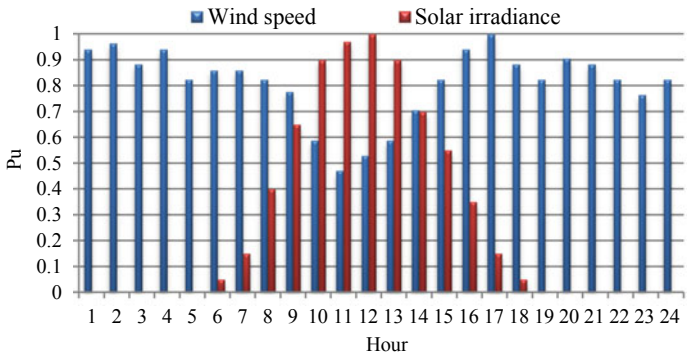


Fig. 8 Hourly variation of wind speed and solar irradiance

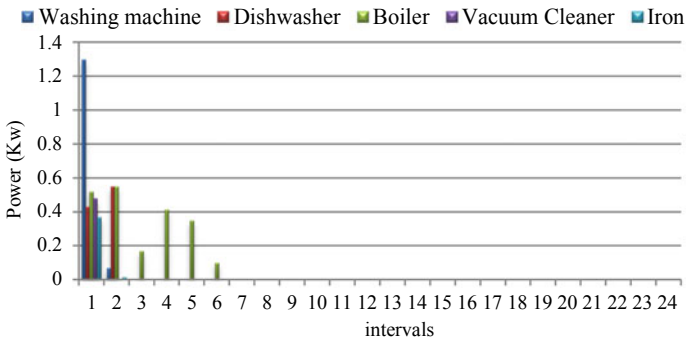


Fig. 9 The demand of controllable appliance in their operational period

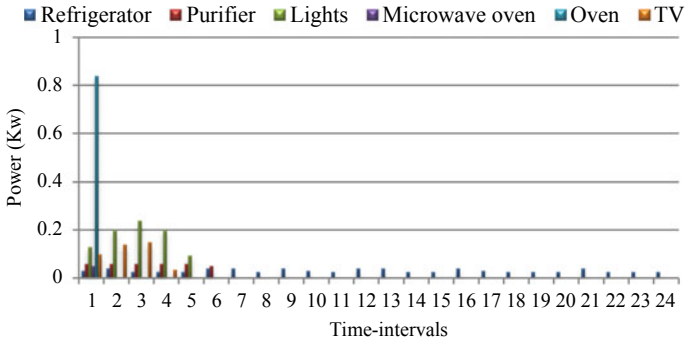


Fig. 10 The demand of non-controllable appliances in their operational period

needs a constant number of time-intervals for properly doing its duty. The number of time-intervals of each device is shown in Figs. 9 and 10.

So, the proposed management method is used to find the optimal schedule of home's devices and renewable sources. Firstly, the MOGWO algorithm is used to multi-objective optimized the techno-economical objective function. The variations of the economic and technical indices during iterations of MOGWO and also the obtained optimal Pareto-front after applying the meta-heuristic algorithm are demonstrated in Fig. 11.

After finding the optimal Pareto-front, the fuzzy method is applied to non-dominate solutions for selecting the best result. The optimal starting time of appliances is presented in Table 1.

The starting time of appliances is selected so that the load density reduces and also the most part of the demand is provided by the smart home renewable energy sources. This management causes to improve the technical and economic indices of the home.

Figure 12 demonstrates hourly power of all appliances and also the stochastic power of renewable units of the smart home. As can be seen in this figure, wind turbine and photovoltaic can provide the most part of the home's demand so that these sources supply approximately 97% of daily demand of the smart home. Moreover, the owner of the smart home can sell power to the grid at time-intervals that the produced power of DG units more than the demand of the home. On the other hands, appliances of the smart home are operated at times when home's energy sources are available or the cost of the market price is low. Totally, the proposed energy management causes to reduce the purchasing power of the smart home from the distribution system. This issue is clearly evident in Table 2.

Table 2 shows the hourly transferred power between the SH and the grid. In this table, the positive number shows that the smart home purchases energy from the grid while the negative number presents that the home sells energy to the grid at time-interval.

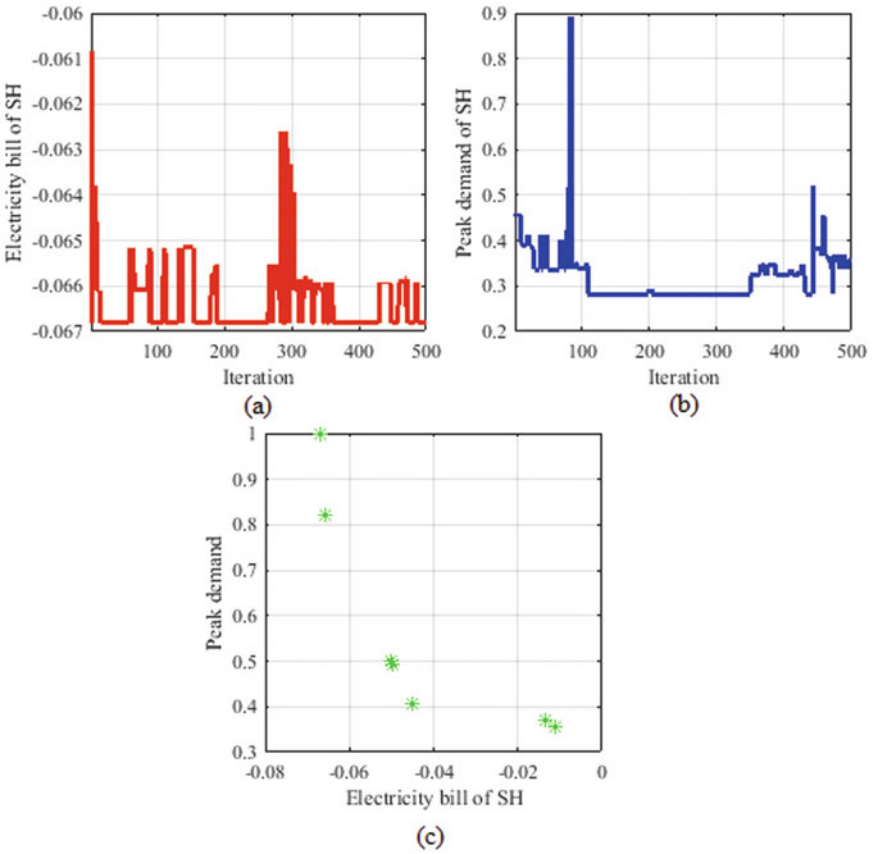


Fig. 11 **a** The variation of the economic index at each iteration, **b** the variation of the technical index at each iteration and **c** the obtained optimal Pareto-front after applying the MOGWO algorithm

Table 1 The optimal starting time of smart home appliances

Controllable	The starting time	Non-controllable	The starting time
Washing machine	13	Refrigerator	1
Dishwasher	3	Purifier	12
Boiler	4	Lights	20
Vacuum cleaner	22	Microwave oven	16
Iron	14	Oven	15
		TV	20

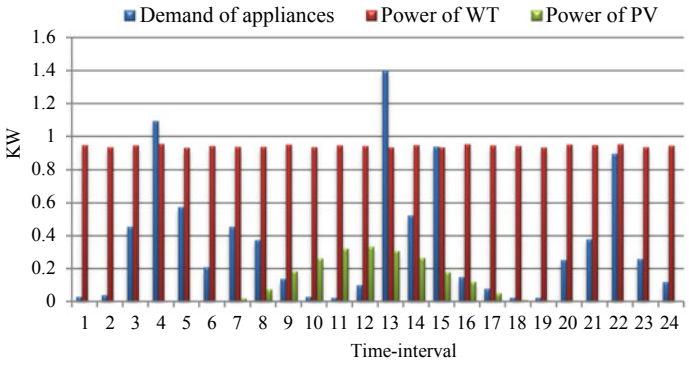


Fig. 12 Hourly demand of appliances and stochastic power of renewable sources

Table 2 Hourly transferred electricity between the smart home and the network

Time-interval	Amount of power	Time-interval	Amount of power
1	−0.9181	13	0.1562
2	−0.8957	14	−0.6904
3	−0.4915	15	−0.1736
4	0.1398	16	−0.9261
5	−0.3579	17	−0.9193
6	−0.7340	18	−0.9302
7	−0.5027	19	−0.9151
8	−0.6381	20	−0.6975
9	−0.9958	21	−0.5688
10	−1.1684	22	−0.0580
11	−1.2441	23	−0.6759
12	−1.1770	24	−0.8246

Therefore, the peak demand of the smart home is improved 88.85% after applying the proposed method so that the peak demand of the home is 0.1562 kW after managing the home's appliance and energy sources using the proposed method while the initial amount of this index is 1.4 kW.

The hourly economic details of the smart home including of cost of purchasing power from the system, the profit of sold power to the system, the cost of produced power of WT and PV and also the hourly electricity bill are presented in Table 3. According to this table, the proper coordination between the operational time of home's appliances and the produced power of renewable sources causes that the owner of the smart home has to buy energy from the grid on at 4th and 13th time-intervals. At other time-intervals, the customer earns money by selling energy to the distribution company. The initial electricity bill of the smart home is 2.8635 $

Table 3 Hourly economic details of the smart home

Time interval	Cost of purchased power	Profit of sold power	Cost of renewable energies	Electricity bill
1	0.0000	−0.1184	0.0711	−0.0473
2	0.0000	−0.1155	0.0702	−0.0454
3	0.0000	−0.0634	0.0710	0.0076
4	0.0252	0.0000	0.0716	0.0968
5	0.0000	−0.0462	0.0700	0.0238
6	0.0000	−0.0947	0.0708	−0.0239
7	0.0000	−0.0649	0.0716	0.0067
8	0.0000	−0.0823	0.0751	−0.0072
9	0.0000	−0.1285	0.0830	−0.0455
10	0.0000	−0.1507	0.0867	−0.0640
11	0.0000	−0.1605	0.0913	−0.0692
12	0.0000	−0.1518	0.0918	−0.0601
13	0.0656	0.0000	0.0896	0.1552
14	0.0000	−0.0891	0.0880	−0.0011
15	0.0000	−0.0224	0.0814	0.0590
16	0.0000	−0.1195	0.0792	−0.0402
17	0.0000	−0.1186	0.0743	−0.0443
18	0.0000	−0.1200	0.0715	−0.0485
19	0.0000	−0.1180	0.0704	−0.0476
20	0.0000	−0.0900	0.0714	−0.0185
21	0.0000	−0.0734	0.0712	−0.0022
22	0.0000	−0.0075	0.0715	0.0640
23	0.0000	−0.0872	0.0702	−0.0170
24	0.0000	−0.1064	0.0708	−0.0355

while the daily electricity bill of the home is −0.2046 after applying the proposed management method. It means that the owner of the smart home earns money after operating the appliances and renewable sources of the smart home by the HEMS.

6 Conclusion

The proper management of appliance and energy sources of a smart home has a high effect on the performance of the smart home. Manually controlling the home's devices is difficult for the consumer due to the number of devices and their load profile, the variation of market price and the situation of renewable units. For this reason, automatically managing the smart home's appliances and renewable energy resources

was investigated in this chapter. Firstly, analytical and mathematical modeling of devices was presented and then the optimal hourly schedule of appliances, WT and PV was selected by applying the combination of MOWOA and fuzzy method to the techno-economic objective function.

The numerical results demonstrate that the home energy management system can properly schedule the operational time of home's devices in order to improve the considered indices. Appliances are operated at times when the market price is low or the produced electricity of the WT and the photovoltaic panel is high. On the other hands, renewable sources reduce the dependence of the smart home on the power of the grid so that sources can supply about 97% of the home's demand. The effect of the WT is higher than the PV because the power of WT is available approximately all hours of the day while the operational time of the PV panel is limit based on solar irradiance. The considered technical and economic indices of the smart home are improved about 89% and 107% after operating the home's appliances and renewable sources using the proposed energy management method, respectively. Therefore, it can be said that the proposed method improves the efficiency of electrical power and the comfort of the consumer.

References

1. Alilou M, Nazarpour D, Shayeghi H (2018) Multi-objective optimization of demand side management and multi dg in the distribution system with demand response. J Oper Autom Power Eng 6:230–242
2. Killian M, Zauner M, Kozek M (2018) Comprehensive smart home energy management system using mixedinteger quadratic-programming. Appl Energy 222:662–672
3. Zhou B, Li W, Chan K, Cao Y, Kuang Y, Liu X, Wang X (2016) Smart home energy management systems: concept, configurations, and scheduling strategies. Renew Sustain Energy Rev 61:30–40
4. Shayeghi H, Alilou M, Tousi B (2018) Multi-objective optimization of demand side management in the presence of dg and demand response. Int J Ind Electron Electr Eng 6:1–7
5. Beaudin M, Zareipour H (2015) Home energy management systems: a review of modelling and complexity. Renew Sustain Energy Rev 45:318–335
6. Wang Y, Li Y, Cao Y, Tan Y, He L, Han J (2018) Hybrid AC/DC microgrid architecture with comprehensive control strategy for energy management of smart building. Electr Power Energy Syst 101:151–161
7. Rastegar M, Fotuhi M, MoeiniAghtaie M (2018) Developing a two-level framework for residential energy management. IEEE Trans Smart Grid 9:1707–1717
8. Santo K, Santo S, Monaro R, Saidel M (2018) Active DSM for households in SG using optimization and artificial intelligence. Measurement 115:152–161
9. Rastegar M, Fotuhi M, Aminifar F (2012) Load commitment in a smart home. Appl Energy 96:45–54
10. Mazidi M, Zakariazadeh A, Jadid Sh, Siano P (2014) Integrated scheduling of renewable generation and demand response programs in a microgrid. Energy Convers Manag 86:1118–1127
11. Aghajani GR, Shayanfar HA, Shayeghi H (2017) Demand side management in a smart microgrid in the presence of renewable generation and demand response. Energy 126:622–637

12. Sadati S, Moshtagh J, Shafie-khah M, Catalão J (2018) Smart distribution system operational scheduling considering electric vehicle parking lot and demand response programs. Electr Power Syst Res 160:404–418
13. Mirjalili S, Saremi S, Mirjalili M, Coelho L (2016) Multi-objective grey wolf optimizer: a novel algorithm for multi-criterion optimization. Expert Syst Appl 47:106–119
14. Alilou M, Talavat V, Shayeghi H (2018) Simultaneous placement of renewable DGS and protective devices for improving the loss, reliability and economic indices of distribution system with nonlinear load model. Int J Ambient Energy
15. Modi A, Bühler F, Andreasen J, Haglind F (2017) A review of solar energy based heat and power generation systems. Renew Sustain Energy Rev 67:1047–1064
16. Stathopoulos T, Alrawashdeh H, Al-Quraan A, Blocken B, Dilimulati A, Paraschivoiu M, Pilay P (2018) Urban wind energy: some views on potential and challenges. J Wind Eng Ind Aerodyn 179:146–157
17. Barbato B, Capone A, Chen L, Martignon F, Paris S (2015) A distributed demand-side management framework for the smart grid. Comput Commun 57:13–24

Finite Volume Method Used for Numerical Investigations of Electrochemical Devices

Elena Carcadea and Mihai Varlam

Abstract The chapter provides a general overview on the Finite Volume Method (FVM) and on Computational Fluid Dynamic (CFD). It introduces the FVM by using a general scalar transport equation and it describes the main steps of a CFD investigation. All these are applied to the mass, momentum, species, energy and potential conservation equations, equations that govern the operation of Proton Exchange Membrane (PEM) fuel cells. The importance of spatial discretization and of interpolation schemes used in CFD investigations is point out by analysing few parameters with impact on the fuel cell operation. Two cases have been considered. First case based on a fuel cell with a simplified configuration, namely a single serpentine channel, revealed the influence of spatial discretization on the accuracy of the simulation results with regards to current density, pressure and temperature. The second case based on a lab-scale fuel cell with two configurations for channels (7 serpentine and 7 parallel) have been used to analyse the effect of three interpolation schemes (first order, second order, QUICK) on the PEM fuel cell operation; therefore, pressure, hydrogen and water mass fraction profiles were considered for comparison. It was found out that besides the differences in the results accuracy due to spatial discretization and interpolation schemes, the design/geometry used in the CFD investigation may or may not emphasize these differences. If for the 7-serpentine channels fuel cell the interpolation scheme did not show much changes in the accuracy of the results not the same conclusion was drawn for the 7-parallel channels fuel cell where the accuracy of the results improved with increasing the order of the interpolation scheme. A mesh-independent solution on a well-posed problem will provide valuable and accurate results only if the numerical methods are appropriate and the interpolation schemes are of high order. The modeling of fuel cells using CFD techniques, as of any other device, can be an important alternative to the experiment, providing information that is critical to design, operation

E. Carcadea (✉) · M. Varlam
National Research and Development Institute for Cryogenics and Isotopic Technologies, ICSI, Rm. Valcea, Romania
e-mail: elena.carcadea@icsi.ro

M. Varlam
e-mail: mihai.varlam@icsi.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_13

and optimization, the requirement being to use appropriate model, assumptions and boundary conditions and, of course, an adequate numerical method.

Keywords Numerical methods · Spatial discretization · Interpolation · Fuel cell · Optimization

Glossary

c	Concentration, mol/m^3
D	Diffusivity, m^2/s
F	Faraday's constant, C/mol
i	Current, A
j	Current, density, A/m^2
j_0^{ref}	Reference current density, A/m^2
k	Thermal conductivity, W/(m K)
K	Absolute permeability, m^2
M	Molecular weight, g/mol
R	Universal gas constant, J/(mol K)
S	Source term
p	Pressure, Pa
T	Temperature, K
Y	Mass fraction
v	Velocity, m/s
V_{OC}	Open circuit voltage, V

Greek Symbols

α	Charge transfer coefficient
ε	Porosity
ζ	Specific active surface area, 1/m
η	Overpotential, V
μ	Dynamic viscosity, Pa s
φ	Potential, V
ρ	Density, kg/m^3
$\sigma_{sol/mem}$	Electric/membrane conductivity, S/m

Subscripts and Superscripts

a	Anode
c	Cathode
sol	Solid
mem	Membrane

1 Introduction

Electrochemical devices, such as: fuel cells and electrolysers, are complex physical systems capable of transforming the chemical energy of the continuously supplied reactants into electrical energy or of generating chemical reactions due to electrical energy provided. Gas kinetics phenomena, diffusion phenomena between the porous layers, mass transport phenomena, along with complex geometries and various material properties make numerical investigations an effective tool in the optimization process and a viable alternative to experiments [1–3]. The large number of parameters that influence the fluid flow and the complex processes that are taking place inside electrochemical devices are described by non-linear partial differential equations (PDE) that do not allow analytical solutions except for some simple or simplified models [4–6]. Obtaining solutions in the other cases requires calculations based on Computational Fluid Dynamics (CFD) [7, 8], that is replacing the PDE systems by a set of algebraic equations which can be solved using numerical methods [9–11].

One well-known discretization method used for numerical simulation of different conservation equations is the Finite Volume Method (FVM) [12–14], a method capable of dealing with all kinds of complex phenomena and systems, being applied to address some of the challenges that emerge in the operation of the electrochemical devices in order to optimize their performance, increase durability and reduce costs [15–17].

The FVM is used in many Computational Fluid Dynamic (CFD) software tools due to its advantages, such as: the conservation of fluxes through a particular control volume, the ability to use structured and unstructured grid for the computational domain, the implementation of different boundary conditions, robustness and proper convergence. The chapter gives a general overview of the FVM, discusses all the steps involved in CFD simulations and presents a numerical algorithm developed for solving a generic diffusion-convection transport equation. This algorithm is applied to the mass, momentum, species, energy and potential conservation equations that are taking place inside electrochemical devices and the results are presented in detail in the chapter in case studies. The aim is to analyse the phenomena of interests (fluid flow, heat and mass transfer, chemical reactions) and to see the influence of CFD steps on the accuracy of the results. The final goal in fuel cell investigations is to obtain optimized geometries, adequate operating parameters and materials, all for an improved performance; therefore, CFD investigations can be an effective tool in

reaching these purposes. A versatile code based on FVM, ANSYS Multiphysics, have been used to show how the spatial discretization and interpolation schemes can affect the results of a numerical investigation, the results being discussed and displayed as profiles taking into account relevant parameters for fuel cell operation (pressure, temperature, hydrogen and water mass fractions).

2 Basics of Finite Volume Method

The FVM is based on the integral form of governing equations that are discretized directly in the computational domain. This computational domain is based on small volumes surrounding each node called "finite volumes". The volume integrals are replaced by surface integrals according to Gauss divergence theorem and as a result a system of algebraic equations for each control volume of the domain is obtained and solved numerically. This system can be solved by using direct (Gauss elimination, LU Decomposition) and iterative methods (Jacobi Method, Gauss–Seidel Method), few of them being briefly presented hereafter.

The chapter introduces the finite volume method by using a general scalar transport, Eq. 1, followed by the description of the main steps required for a reliable numerical investigation, namely: defining the modeling goals, pre-processing with geometry development and grid generation, a mathematical and numerical model with boundary and initial conditions, an iterative or simultaneous solver for the equations, and a post-processor for analysing the results.

$$\frac{\partial(\rho\varphi)}{\partial t} + \nabla \cdot \left(\rho\varphi \vec{u}\right) = \nabla \cdot (\Gamma \nabla\varphi) + S_\varphi \tag{1}$$

where:

ρ density
$\vec{u}$ velocity vector ($=u_x \vec{i} + u_y \vec{j} + u_z \vec{k}$)
Γ diffusion coefficient for the scalar φ
$\nabla\varphi$ gradient of scalar φ, $\nabla\varphi = \frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}$
S_φ source of φ per unit volume.

By integrating the general scalar equation over an arbitrary volume, V_i, and by applying the Gauss divergence theorem, the integral form is obtained:

$$\int_{V_i} \frac{\partial(\rho\varphi)}{\partial t} dV + \oint \rho\varphi \vec{u} \cdot d\vec{A} = \oint \Gamma \nabla\varphi \cdot d\vec{A} + \int_{V_i} S_\varphi dV \tag{2}$$

where: $\vec{A}$ is the surface area vector.

The Eq. 2 is applied to each control volume or cell, in 3D or 2D modeling, respectively. An example of a domain used to solve a set of scalar transport equations

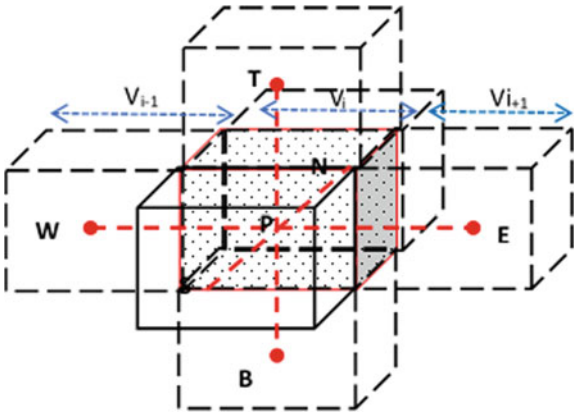


Fig. 1 Control volumes used for discretization of a scalar transport equation

is presented in Fig. 1. By discretizing the Eq. 2 on a control volume it results:

$$\frac{\partial(\rho\varphi)}{\partial t}V + \sum_{f}^{N_f} \rho_f \overrightarrow{v}_f \cdot \overrightarrow{A}_f = \sum_{f}^{N_f} \Gamma_\varphi \nabla\varphi_f \cdot \overrightarrow{A}_f + S_\varphi V \tag{3}$$

where

N_f	number of faces enclosing cell.
φ_f	value of φ convected through face f.
$\rho_f \overrightarrow{v}_f \cdot \overrightarrow{A}_f$	mass flux through the face f.
V	cell volume.

The discretized scalar transport Eq. 3 for the variable φ takes into account the surrounding neighbour cells, as presented in Fig. 1. A linearized form of the Eq. 3 is given by the following expression:

$$a_P\varphi = \sum_{nb} a_{nb}\varphi_{nb} + b \tag{4}$$

where the subscript nb refers to the neighbour cells (W, E, N, S, T, B), their number depending on the mesh developed and being equal to the number of faces surrounding the cell. The coefficients a_P and a_{nb} are the linearized coefficients for scalar variables φ and φ_{nb}. For each cell in the mesh similar equations can be written resulting in a set of algebraic equations that can be solved using several numerical methods.

3 Numerical Methods for Solving Linear Systems

The direct methods for solving the linear systems, such as: $Ax = b$, are based on a finite number of arithmetic operations to find the exact solution of the system, considering also a round off error. A unique solution of the system is determined only if the inverse matrix for A exists, which is equivalent to fulfilling the condition: $\det(A) \neq 0$, the vector $x = A^{-1}b$ being the solution of the system.

The Gauss elimination method, named about the famous German mathematician Carl Friedrich Gauss who wrote about this method, but did not invent it [18], is one method used to find the solutions of linear systems and is based on variables elimination from the equations until one equation having one variable is obtained and solved. The idea of the algorithm is to eliminate the matrix element under the diagonal, obtaining zeros, transforming the matrix A in an upper triangular matrix U and to solve the upper triangular system of equations by a back substitution process. In this way, the $x_n = \frac{b_n}{U_{nn}}$ is found and proceeding in an upward manner the x_{n-1}, x_{n-2}, …, x_1 are determined and the solution of linear equations system is obtained. The general form of the solution x_i can be expressed by the following formula:

$$x_i = \frac{b_i}{a_{ii}} - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} x_j \tag{5}$$

In the case of large number of equations, the Gauss elimination method can be expensive as regarding the computational effort since it gives the solution of one set of linear equations at a time. It is good as any other numerical method, being considered the fundamental solution algorithm.

The LU Decomposition method of the polish mathematician Tadeusz Banachiewicz, introduced in 1938 [19], is based on the factorization of the matrix A in a lower triangular matrix having unit diagonals (L) and an upper triangular matrix (U). It is basically a modified form of Gauss elimination method. Following, the $Ax = b$ system become $LUx = b$ and the solution can be easily obtained by solving two linear systems. By noting $y = Ux$ the system solved first is the lower triangular system $Ly = b$ and the solution y is determined, followed by solving the upper triangular system $Ux = y$ to obtain the solution x. The general form of the solutions y_i and x_i can be expressed by the following formulas:

$$y_1 = \frac{b_1}{l_{11}} \tag{6}$$

$$y_i = \frac{b_i}{l_{ii}} - \sum_{j=1}^{i-1} l_{ij} y_j, \text{ for } i = 2, \ldots, n \tag{7}$$

$$x_n = \frac{y_n}{u_{nn}} \tag{8}$$

$$x_i = \frac{y_i}{u_{ii}} - \sum_{j=i+1}^{n} u_{ij} x_j, \text{ for } i = n-1, \ldots 1 \tag{9}$$

The LU decomposition method is efficient in the case of repeatedly solving of one set of linear equations but it becomes less effective and more complex than the Gauss elimination method.

The Gauss-Seidel iterative method [21], named after the German mathematicians Carl Friedrich Gauss and Philipp Ludwig von Seidel, is used for solving a square system with n linear equations and is based on an initial approximation of the solution vector, namely $x^{(1)}$, to start calculation for the next estimates $x^{(k+1)}$, for $k = 1, 2, \ldots$, resulting in:

$$x_i^{(k+1)} = \frac{b_i}{a_{ii}} - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(k+1)} - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k)} \tag{10}$$

In Gauss-Seidel iterative method each new calculated solution (x_i) is replacing the previous calculated value. From Eq. 10 it can be noticed that for the computation of $x_i^{(k+1)}$ there are used the previously $(i-1)$ elements computed at $(k+1)$ step and only the elements from $(i+1)$ to n computed in the k iteration. The iteration continues until the absolute relative approximate error is less than a tolerance specified for all unknowns. This method can be advantageous for very large problems.

4 Computational Fluid Dynamics

Computational Fluid Dynamics (CFD) is an interdisciplinary science at the interface between numerical analysis, fluid mechanics and computer science providing a numerical approximation of the equations that govern the fluid dynamics. Fluid flow, heat and mass transfer, chemical reactions, and related phenomena can be predicted by solving the conservation equations for mass, momentum, species, energy and potential. There are many books [20–22] and papers [23–25] written on CFD investigations since is considered an indispensable tool for many fields of research and industries with applications spanning from environmental aspects, weather simulations and up to engineering processes where a multitude of phenomena from aerospace, automotive and up to energy and power generation can be analysed in all stages of an engineering process from: design, product development, optimization and, followed by re-design if is required. The CFD is considered a reliable tool in identifying, understanding, predicting, controlling and optimizing various transport processes and phenomena that occur in such applications and an alternative for costly experimental tests.

The steps required for analysing a fluid flow problem are presented in the algorithm shown in Fig. 2.

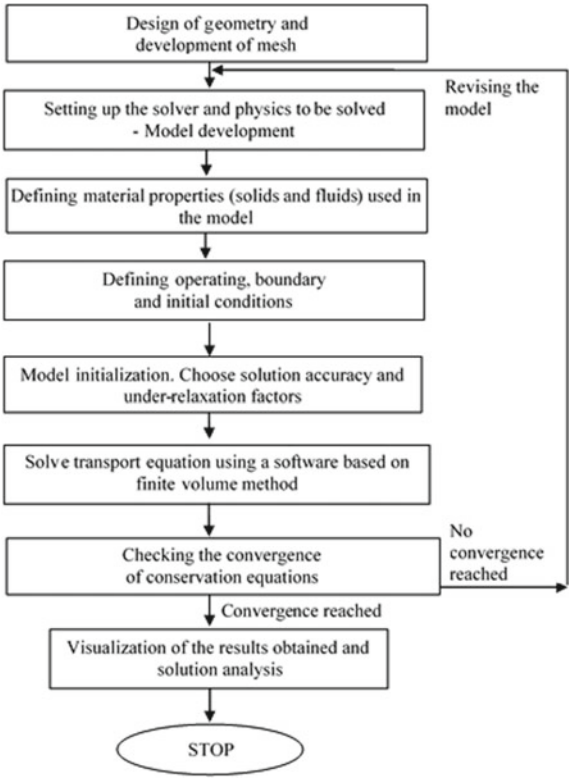


Fig. 2 CFD analysis steps

A short description of the steps involved in a CFD investigation will be presented below and an example on applying this procedure will be discussed in the next section where an investigation for electrochemical devices will be presented.

The design of a geometry and mesh generation is often the most complex process in solving a problem. Finding a numerical solution accurate and grid independent is a process that depends on the number of discretization volumes. These volumes can have any shape: hexahedrons, tetrahedrons, prisms, pyramids and so on, but the requirement is that the faces of the volumes to be planar and each element to be convex. If a small number of volumes or discretization nodes are used, then there will be a significant difference between the discrete solution and the exact solution. A good numerical scheme will lead to the exact solution as long as the number of volumes in the mesh is large enough.

For the spatial discretization, several interpolation schemes are available, including central differencing, first order and second order upwind, QUICK—Quadratic Upstream Interpolation for Convective Kinetics, Power-law, the accuracy and stability of the solution being influenced by their use. Among them, the first order and second order upwind schemes are the most used in CFD simulations being very

simple and extremely stable. Although higher-order schemes offer improved accuracy in the computations, they can lead to numerical instabilities due to unphysical oscillations in regions of steep gradients [26].

As mentioned in Sect. 2, the strategy of numerical modeling consists in replacing the problem defined on a continuous domain with a problem of the same type that has a discrete domain. Obtaining an approximate numerical solution is based on the use of a discretization method that replaces the system of differential equations by a system of algebraic equations, typically solved by one of the methods discussed in Sect. 3, such as Gauss–Seidel or Gauss elimination.

Mathematical model formulation involves several steps related to setting up the solver and physics to be solve, defining operating and boundary conditions, defining material properties, followed by the model initialization, accuracy and under-relaxation factors establishing in order to solve the transport equations. When defining the modeling goals it is important to determine what types of results are expected, what are the assumptions that can be taken into account and do not affect the phenomena investigated, what physical models must be included or what accuracy is required. The conservation of mass, momentum, energy and species will be customized according to the problem needed to be solved (laminar or turbulent, 2D or 3D investigation) and the boundary conditions will specify the fluid behaviour and properties at the domain boundaries.

It must be mentioned that any attempt to produce a general method of solving the conservation equations, applicable to all types of fluid flow, is impracticable, perhaps even impossible. Two numerical methods, a pressure-based solver and a density-based solver, are used in CFD codes in order to solve sequentially and iteratively the transport equations. The default solver in most of the CFD solvers is the pressure-based solver. Two methods are available for this solver, namely: segregated and coupled method that solves the conservation equations one after another for the first method and simultaneously for the second. In both methods, the velocity field is obtained from the momentum conservation equation. After updating the mass flux, pressure and velocity the additional scalar equations for energy, species, potential are solved repeatedly until the solution converges, with given boundary and initial conditions. The most common boundary conditions used are Dirichlet (value for the scalar φ is specified) and Neumann (flux for the scalar is specified, $\partial\varphi/\partial n$). Also, for reaching the convergence in solving the system of linear algebraic equations it is important to monitor the residuals,R^{φ}, taking into account the under-relaxation of variables and a precise convergence criteria. Using the Eq. (4), the residual R^{φ} can be calculated based on the following expression [26]:

$$R^{\varphi} = \sum_{cells\ P} \left| \sum_{nb} a_{nb}\varphi_{nb} + b - a_P\varphi_P \right| \tag{11}$$

where the subscript *nb* refers to the cells, the coefficients a_P and a_{nb} are the linearized coefficients for scalar variables φ and φ_{nb}.

The iterative process for reaching a convergent solution will stop when the total residual error, as sum of the absolute residuals ($\sum_{cells} |R^{\varphi}|$) or as root-mean-square error $\left(\sqrt{\frac{1}{N}\sum_{cells}(R^{\varphi})^2}\right)$, will go to zero or to a small value (i.e. 10^{-4}) chosen as a convergence criterion.

As mentioned before, under-relaxation of variables φ can be used to reduce the changes produced during each iteration, so that the new value of the variable φ within a cell is calculated taking into account the previous value, φ^{prev}, the computed change in the variable φ,$\Delta\varphi$, and the under-relaxation factor, α, according to the following expression:

$$\varphi = \varphi^{prev} + \alpha\Delta\varphi \tag{12}$$

Most of the CFD investigations requires under relaxed values since helps in increasing the convergence rate of the solution and can minimize the number of iterations required, while ensuring the solution stability. The under-relaxation factor α has the value between zero and one. In any CFD software is recommended to use the default under-relaxation factors, chosen to be suitable for many problems, and to decrease these values only if the residuals continue to increase after a few iterations.

After obtaining a converged solution, the post-processing step it follows in order to analyse the results and visualize the important parameters and phenomena investigated (fluid flow, heat and mass transfer).

5 Numerical Investigation of an Electrochemical Device—A Study Case for PEM Fuel Cells

Electrochemical devices such as PEM fuel cell have proven to be viable energy conversion sources for various applications in the past decades due to high efficiency, low emissions, adaptability, high reliability, low maintenance costs, excellent performance, modularity and the ability to use a variety of fuels. A fuel cell is a complex system that is continuously fed with reactants and under the action of a potential difference it transforms the chemical energy of reactant into electrical energy, having as by products water and heat. Gas kinetics phenomena, diffusion phenomena between the porous layers, heat and mass transport phenomena that are taking place inside complex geometries are just few of the processes that occur inside the fuel cells and can be investigated by numerical modeling for performance optimization and costs reduction.

The fuel cell is an electrochemical device that uses hydrogen and oxygen to produce electricity, heat and water. It consists of an anode, a cathode and a polymeric membrane. Both the anode and the cathode have the following components: a current collector, a bipolar plate with gas channels, a gas diffusion layer and a catalyst layer.

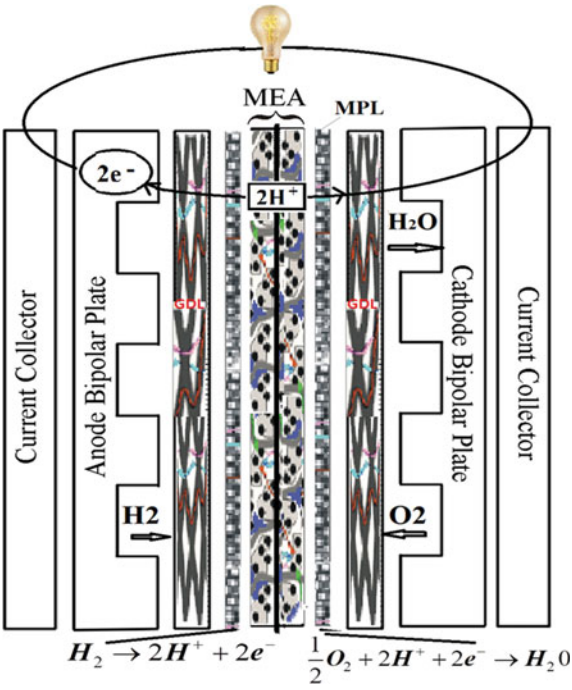


Fig. 3 PEM fuel cell components and operating principle

The gas flow channels have a specific geometry so that the contact surface between the gas and the membrane is as large as possible. The hydrogen from the anode side is split into protons and electrons. The protons are passing through the membrane into the cathode side, while the electrons are directed through an outer circuit, due to the potential difference between the anode and the cathode. In this way, the electrons reach the cathode side where combines with the hydrogen protons and the oxygen (from the air), in the presence of catalysts to form water, thus completing the circuit [27], as presented in Fig. 3.

The overall reaction taking place inside PEM fuel cell is given by:

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \tag{13}$$

The mathematical models and the numerical simulations for fuel cells investigations have undergone an intense development lately, being used to understand the internal phenomena, to optimize the geometry of the fluid flow channels or to improve the performance and durability. A detailed model should include a complete set of equations to describe fluid flow in porous media, complex processes of diffusion of chemical species, phase changes and multiphase flows, chemical and electrochemical reactions occurring in the catalyst layer, proton transport through membrane, electrical potential and current density, mass and heat transfer.

The mathematical model developed for investigating the physical and electrochemical phenomena that take place in a fuel cell is based on the conservation of mass, moment, species, energy and electrical charge. Using the general scalar transport Eq. 1, the mathematical model for a PEM fuel cell is developed for each component, the finite volume and numerical methods described before are applied in order to have a comprehensive CFD investigation.

5.1 *Finite Volume Method Applied for PEM Fuel Cells Investigation*

The mathematical model developed hereafter for all components of a fuel cell is based on the general scalar transport Eq. 1, taking into account: gas channels that provide reactants and remove the by-products; the bipolar plates that conduct the electrons; the gas diffusion layers that helps in gasses diffusion; catalyst layers where the electrochemical reactions occur and a polymeric membrane that separates the two sides of a fuel cell, the anode and the cathode, and allows the transport of water and protons. Both the anode and the cathode of a PEM fuel cell have the same components, therefore the equations of the model included in the Tables 1, 2, 3, 4 and 5 are valid for both sides, with the mention that where the value of a variable has a different expression, it is specified.

The variables are explained in the nomenclature section and the closure relations are given below:

Table 1 Equations solved in PEM fuel cell channels

Equation solved	φ	ε	Γ	S_φ
Mass conservation	ρ	1	0	0
Momentum conservation on x-axis	u	1	μ	0
Momentum conservation on y-axis	v	1	μ	0
Momentum conservation on z-axis	w	1	μ	0
Species conservation—H_2	Y_{H_2}	1	D_{H_2}	0
Species conservation—O_2	Y_{O_2}	1	D_{O_2}	0
Species conservation—H_2O	Y_{H_2O}	1	D_{H_2O}	0
Energy conservation	T	1	k	0

Table 2 Equations solved in PEM fuel cell bipolar plates

Equation solved	φ	ε	Γ	S_φ
Energy conservation	T	1	k	i_{sol}^2 / σ_{sol}
Solid potential conservation	φ_{sol}	1	σ_{sol}	j

Table 3 Equations solved in PEM fuel cell gas diffusion layers

Equation solved	φ	ε	Γ	S_φ
Mass conservation	ρ	1	0	0
Momentum conservation on x-axis	u	ε	0	$-\mu\varepsilon u/K$
Momentum conservation on y-axis	v	ε	0	$-\mu\varepsilon v/K$
Momentum conservation on z-axis	w	ε	0	$-\mu\varepsilon w/K$
Species conservation—H_2	Y_{H_2}	ε	D_{H_2}	0
Species conservation—O_2	Y_{O_2}	ε	D_{O_2}	0
Species conservation—H_2O	Y_{H_2O}	ε	D_{H_2O}	0
Energy conservation	T	ε	k	i_{sol}^2/σ_{sol}
Solid potential conservation	φ_{sol}	ε	σ_{sol}	0

Table 4 Equations solved in PEM fuel cell catalyst layers

Equation solved	φ	ε	Γ	S_φ
Mass conservation	ρ	1	0	$\begin{cases} -(M_{H_2}/2F)j - M_{H_2O}\cdot\alpha \quad in \quad anode \\ -(M_{O_2}/4F)j + M_{H_2O}\cdot\alpha + (M_{H_2O}/2F)j \; in \; cathode \end{cases}$
Momentum conservation on x-axis	u	ε	0	$-\mu\varepsilon u/K$
Momentum conservation on y-axis	v	ε	0	$-\mu\varepsilon v/K$
Momentum conservation on z-axis	w	ε	0	$-\mu\varepsilon w/K$
Species conservation—H_2	Y_{H_2}	ε	D_{H_2}	$-(M_{H_2}/2F)j$
Species conservation—O_2	Y_{O_2}	ε	D_{O_2}	$-(M_{O_2}/4F)j$
Species conservation—H_2O	Y_{H_2O}	ε	D_{H_2O}	$\begin{cases} -M_{H_2O}\cdot\alpha \quad in \quad anode \\ M_{H_2O}\cdot\alpha + (M_{H_2O}/2F)j \quad in \quad cathode \end{cases}$
Energy conservation	T	ε	k	$\begin{cases} j(\eta - T\Delta S/2F) + i_{sol}^2/\sigma_{sol} + i_{mem}^2/\sigma_{mem} \quad in \; anode \\ j(-\eta - T\Delta S/2F) + i_{sol}^2/\sigma_{sol} + i_{mem}^2/\sigma_{mem} \; in \; cathode \end{cases}$
Solid potential conservation	φ_{sol}	ε	σ_{sol}	j
Protonic potential conservation	φ_{mem}	ε	σ_{mem}	$-j$

Table 5 Equations solved in PEM fuel cell membrane

Equation solved	φ	ε	Γ	S_φ
Energy conservation	T	ε	k	i_{mem}^2 / σ_{mem}
Species conservation	Y_{H^+}	ε	D_{H^+}	
Protonic potential conservation	φ_{mem}	ε	σ_{mem}	$-j$

$$j_a = \left(\zeta_a \cdot j_{0,a}^{ref}\right)\left(c_{H_2} \Big/ c_{H_2}^{ref}\right)^{\gamma_a}\left[\exp\left(\alpha_a F \eta_a / RT\right) - \exp\left(-\alpha_c F \eta_c / RT\right)\right] \tag{14}$$

$$j_c = \left(\zeta_c \cdot j_{0,c}^{ref}\right)\left(c_{O_2} \Big/ c_{O_2}^{ref}\right)^{\gamma_c}\left[-\exp\left(\alpha_a F \eta_c / RT\right) + \exp\left(-\alpha_c F \eta_c / RT\right)\right] \tag{15}$$

$$\eta_a = \varphi_{sol} - \varphi_{mem} \tag{16}$$

$$\eta_c = \varphi_{sol} - \varphi_{mem} - V_{oc} \tag{17}$$

More details about the previously PEM fuel cell model developed and used to investigate the influence on performance of various channels configurations, various properties of porous media or catalyst microstructure can be found elsewhere [24, 25] and are not repeated here.

5.2 Numerical Modeling—CFD Implementation

CFD analysis provides predictions of fluid flows, giving insights on information that are difficult, expensive or impossible to be obtained by experiments. The steps involved in such investigations, and presented in Fig. 2, must be implemented in a software package that can apply numerical methods in order to find an accurate approximation of the solution of the system of partial differential equations which cannot be solved analytically. The reliability and accuracy of the CFD investigation depends on all steps involved: from domain discretization and up to the solution methods and controls chosen in simulation.

There are many CFD software tools that can be used to investigate the complex processes taking place inside PEM fuel cells, from open source [28, 29] to commercial software [30, 31], having similar capabilities such as multidimensional effects of phenomena or specific approaches such as the catalyst microstructure investigation. These software packages have the conservation equations and many numerical techniques and methods built in, being ready to analyse fluid dynamics, solid mechanics and thermodynamics problems. A variety of applications involving all kinds of complex physics and real-world conditions can be analyse using ANSYS Fluent, one of the most preferred commercial software available since 1983 [30]. Up to 50 User Defined Scalar (UDS) transport equations, as Eq. 1, can be solved in the

same model using ANSYS Fluent. Boundary conditions, properties and parameters can be used to customize the model to particular phenomena of interest.

In this chapter, the ANSYS Multiphysics with Fluent software have been used for investigating the electrochemical, thermodynamics and fluid dynamics inside the PEM fuel cell. A special tool developed by ANSYS, the Fuel Cell Module, incorporates all the equation presented in Tables 1, 2, 3, 4 and 5 and can be used for such investigations as default application or can be enhanced by developing UDFs to include specific characteristics that are not available in the default options.

A CFD analysis includes the following steps: problem definition, mathematical model formulation, geometry (domain) designing and mesh generation, solving, and post-processing. The solving step involves several actions that must be repeated several times for reaching the convergence and obtaining reliable results.

5.2.1 Problem Definition and PEM Fuel Cell Model Development

In the last 20 years, many models have been developed for PEM fuel cells due to complexity of the phenomena involved, such as: electrochemical reactions coupled with mass and heat transfer, multiphase and multi-species flows. Most models focus only on one aspect [32–34] or a certain region of the fuel cell [35, 36], but to obtain a detailed model this is not enough. At present, there is no general consensus for describing the phenomena that occur in the PEM fuel cell layers. Many mathematical models or experimental data are obtained for particular cases, which are often idealized and unrealistic. Most models require uniform conditions for fuel supply [37] and an isothermal temperature [38], for example, but both simplifications are describing particular aspects of the fuel cell behaviour and are independent of the other effects that may occur, leading to several shortcomings. Therefore, a detailed model incorporating the three-dimensional study of the phenomena involved and their dynamic effects is necessary.

The first models developed were one-dimensional, isothermal and analysed only the gas diffusion layer [25], the catalytic layer [24] or the membrane [39]. At present, there are numerous papers discussing aspects regarding the material properties [40], the fluid flow [41], heat transfer [42], and chemical and electrochemical reactions [43] that occur in fuel cells using advanced numerical techniques for fluid dynamics analysis.

Before developing a model for a fuel cell it is very important to establish the main characteristics, such as: model approach (theoretical, empirical, and semi-empirical), system state (steady state or transient), system boundaries (single fuel cell, stack, and system), size of domain (from zero to three dimensional), complexity and degree of detail (electrochemical, thermodynamic or fluid dynamics), accuracy, flexibility, validation.

The PEM fuel cells models have become increasingly complex, so most of the models are solved numerically using specialized software such as ANSYS Fluent [30] or COMSOL Multiphysics [31]. Finding a suitable numerical technique is not the only problem in PEM fuel cell modeling, a good understanding of the physical and

chemical processes that occurs, the need to link the macroscopic with the microscopic processes are few of the challenges that must be overcome and are discussed in several articles.

5.2.2 Geometry Design and Spatial Discretization

The complex geometries used in electrochemical devices and different scales used from nano- to meter- scale make the discretization process a challenge. In solving a problem using numerical algorithms the basic idea is to generate the mesh by splitting the domain into small controls volumes. The accuracy of the numerical solution is dependent on the number of volumes; therefore, the approximation will tend to the exact solution as long as the number of nodes in 2D and volumes in 3D of the mesh is large enough.

Several tools can be used in designing and meshing the geometries of the PEM fuel cells, such as: Gambit, SpaceClaim, DesignModeler, ICEM-CFD, SolidWorks or Catia. The pre-processor Gambit 2.4.6 has been used in this study to develop the geometries and the meshes, often considered the most complex process because the accuracy of the results is affected by the discrete domain generated.

The mesh used to find the approximate solution of the PEMFC model should be dense enough to ensure its grid independency, but also coarse enough to keep to a minimum the computational requirements. Therefore, a mesh refinement study must be done when designing a geometry and performing a numerical investigation, and best practice recommendations need to be considered in order to take into account some specific characteristics of the model. Figure 4 presents two flow field patterns

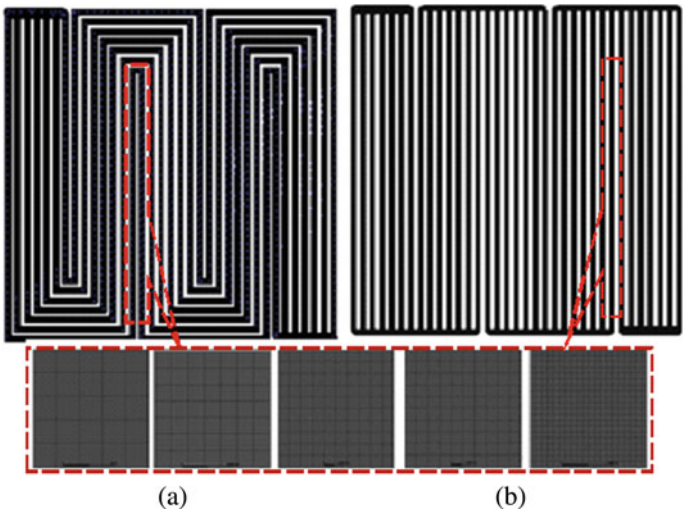


Fig. 4 Mesh developed for a PEM fuel cell with **a** serpentine channels and **b** parallel channels for different number of cells in the channel cross-section

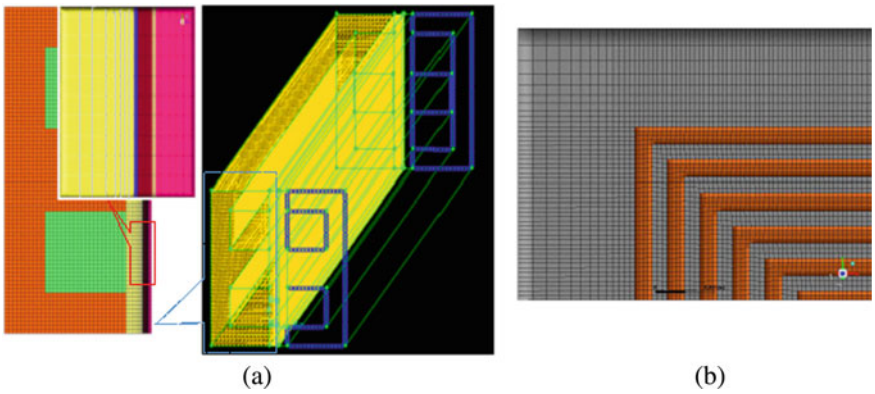


Fig. 5 **a** Fuel cell with a zoom for mesh density in the thickness direction, **b** mesh density in U-bends zone

with 5 cross sections for the channels with 5, 7, 10, 15, 20 cells in the width direction. Since the thickness direction of the fuel cell is much smaller than the length and width directions, it is very important that the mesh density to be different, much higher in the thickness direction compared with the other two directions, as can be seen in Fig. 5a. Also, there are some cases, i.e. serpentine channels, where a flow field channel has several U-patterns that needs a mesh density higher compared to the mid-section of each channel, as presented in Fig. 5b, configuration that comes with a recommendation to coarsen the mesh density along the channel length because it does not affect the accuracy of the solution [30].

A simple geometry consisting in a single serpentine channel, i.e. the red highlighted channel in Fig. 4, is taken into account in the grid independency analysis. 5 meshes have been developed using for the channel depth/width a grid with 5, 7, 10, 15 and 20 cells and the same number of cells in the length direction, resulting in a mesh for the PEM fuel cell geometry with a total number of 110,800, 140,400, 248,000, 475,800 and 736,000 cells, respectively. These geometries will be used further in determining the mesh-independent solution.

5.2.3 Model Solving

The mathematical model used to investigate the fluid dynamic inside PEM fuel cells is three-dimensional, non-isothermal and is dealing with a multi-physics and multiphase flow. For solving the set of equations describing the conservation of mass, momentum, species, energy and charge, equations presented in Sect. 5.2.1, ANSYS Fluent solver with PEM Fuel Cell Module have been used and some assumptions and boundary conditions have been taken into account. It must be mentioned that Gauss–Seidel method was chosen for solving the system of equations that govern the PEM fuel cell operation.

Regarding the assumptions used to solve the equations, the following are considered, being also the most used in numerical investigations of PEM fuel cells: steady-state and non-isothermal operation, ideal gas mixtures, laminar flow, isotropic and homogenous components, membrane is impermeable to the gas phase, only protons are passing by.

Since the fuel cell is developed as a single domain, no boundary conditions are required between the layers, only boundary conditions for the external surfaces are needed. Dirichlet boundary conditions are prescribed for gas flow inlets: mass flow rate, species mass fractions and temperature. Also, on the external walls of the bipolar plates, a potential difference is set up between the anode and the cathode, zero volts on the anode wall and a constant voltage between zero and open circuit value on the cathode terminal wall. This is called potentiostatic boundary condition and is the most realistic boundary condition that leads to a robust solution, result that can be compared with experimental data.

After setting relevant values for operating pressure and material properties, as in Table 6, the flow is initialized and the simulation starts. The best practice recommendation is to start with a high voltage, close to the open circuit voltage, and then reduce the voltage by 0.05–0.1 V between each converged solution. At each step, an average current density will be calculated at both the anode and the cathode catalyst layers. If the difference between these values is less than 10^{-4} and the residuals for the other scalar variables taken into consideration in the model will go to zero or to a small value (i.e. 10^{-6}) then the iterations will stop and the solution is considered convergent. If there are some convergence problems, then it is recommended to reduce the under-relaxation factors for transport equations to values that do not affect the results accuracy.

Table 6 PEM fuel cell layers properties

Parameter	Value	Unit
Porosity of GDL/MPL/CL (ε)	0.7/0.6/0.5	–
Permeability of GDL/MPL/CL (K)	$3{\cdot}10^{-12}/1{\cdot}10^{-12}/2{\cdot}10^{-13}$	m^2
Reference exchange current density at anode ($j_{0,a}^{ref}$)	3000	A/m^2
Reference exchange current density at cathode ($j_{0,c}^{ref}$)	0.3	A/m^2
H2 molar concentration ($c_{H_2}^{ref}\, c_{H_2}^{ref}\, c_{H_2}^{ref}$)	$54.6{\cdot}10^{-3}$	$kmol/m^3$
O_2 molar concentration ($c_{O_2}^{ref}\, c_{O_2}^{ref}\, c_{O_2}^{ref}$)	$3.39{\cdot}10^{-3}$	$kmol/m^3$
Anodic/Cathodic transfer coefficient (α_a/α_c)	1/0.8	–
Contact angle GDL/MPL/CL (θ)	110/130/95	°
Diffusivity of hydrogen (D_{H_2})	$1.1028{\cdot}10^{-4}$	m^2/s
Diffusivity of oxygen (D_{O_2})	$3.234{\cdot}10^{-5}$	m^2/s
Diffusivity of water (D_{H_2O})	$1.1028{\cdot}10^{-5}$	m^2/s
Open circuit voltage (V_{OC})	0.938	V

5.2.4 Results: Post-processing and Interpretation

The last step in a CFD analysis is the post-processing of the simulation results in order to visualize the information needed to characterize the phenomena investigated, interpret the results and validate the CFD model by comparing the results with the experimental data, if are available. This section will present results from PEM fuel cell simulations, taking into account a simplified geometry and a real configuration, emphasizing the influence on performance of the following: spatial discretization and interpolation schemes used in meshing the domain.

For understanding the importance of the discretization for the accuracy of the numerical solution a simplified PEM fuel cell with a single serpentine channel was considered. The 5 meshes developed using a grid with 5, 7, 10, 15 and 20 cells for the depth/width of the channel and the same number of cells in the length direction leaded to a mesh for the PEM fuel cell geometry with a total number of 110,800 (case 1), 140,400 (case 2), 248,000 (case 3), 475,800 (case 4) and 736,000 (case 5) cells, respectively. A numerical simulation using PEM Fuel Cell Module has been run for each case and an average current density was obtained. A plot with the current density obtained at 0.4 V potential difference is presented in Fig. 6. Although the difference in current density between the cases with coarse (case 1) and fine (case 5) meshes is only 1.4%, it can be noticed that starting from the mesh with 475,800 no significant change takes place, therefore the accuracy and grid independence of the solution is ensured. To keep the computational time to a minimum the mesh with 475,800 cells is recommended to be used.

ANSYS Fluent, as any other CFD software, offers many options to post-process various variables of the PEM fuel cell model and allows the visualization of the results as contours, plots and vectors. Species mass fractions, temperature, pressure and velocities are mainly post-processed in a PEM fuel cell investigation. Figure 7

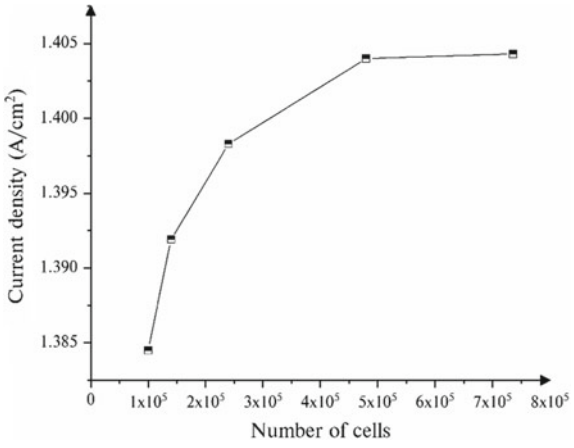


Fig. 6 Mesh size influence on average current density variation

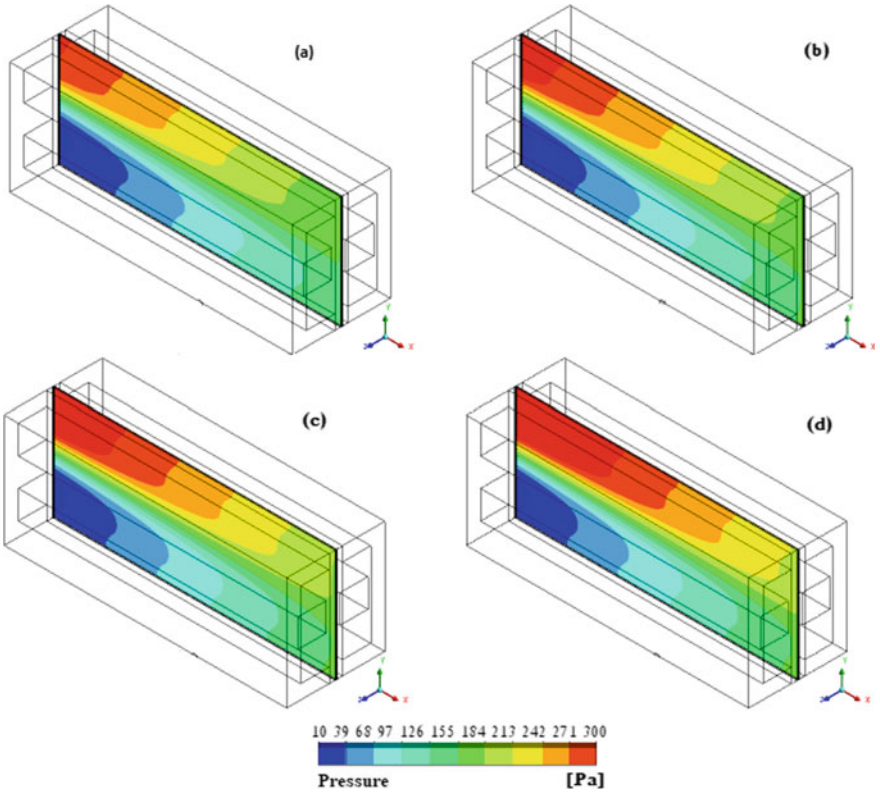


Fig. 7 Pressure distribution in a PEM fuel cell for: **a** case 1, **b** case 2, **c** case 3 and **d** case 4

is presenting the pressure distribution in the anode catalyst layer for cases 1–4. It can be seen that as the mesh become finer the accuracy of the results is improved. If the simulations will underestimate the pressure field or any other parameter due to a solution that is not mesh-independent, then an incorrect or misinterpretation of the results will be obtained. In general, in PEM fuel cell simulations, a higher pressure drop can effectively eliminate the water produced by to electrochemical reactions. In the cases 1–3 an underestimation of the pressure field is presented due to non-mesh independent solution obtained, so it is possible that incorrect results or interpretation to be given.

The study regarding the mesh independency revealed two important results related to the PEM fuel cells operation: (i) for the variables that are expected to increase due to functioning, such as: the water mass fraction, the current density and temperature, the coarse grid is leading to an overestimation; (ii) for the variables that are expected to decrease due to functioning, such as: the hydrogen and oxygen mass fraction, pressure, the coarse grid is giving an underestimation, as presented in the Figs. 7 and 8.

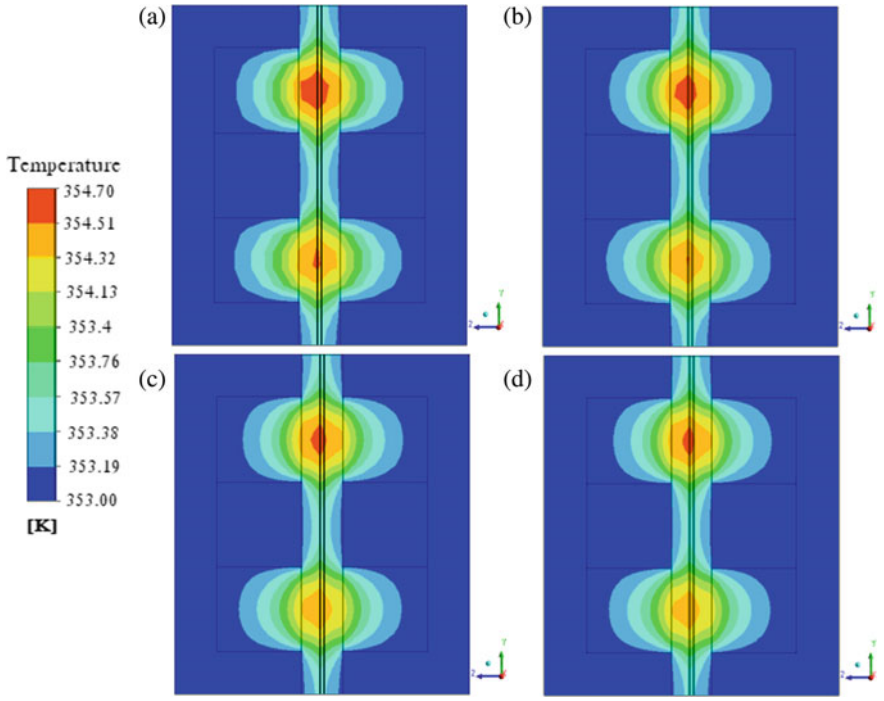


Fig. 8 Temperature profile for a PEM fuel cell in: **a** case 1, **b** case 2, **c** case 3 and **d** case 4

To sustain the first statement and taking into account that the refinement of the mesh of PEM fuel cell layers was on the y–z plane, the temperature was plotted in such plane in Fig. 8. It can be noticed a slight over-estimation for the cases with coarse mesh. Kazemi Esfeh et al. [44] in their study concluded that the concentration of reactants and products is more sensitive to the grid as compared to other parameters such as temperature. They did not notice much change because in their simulations the maximum variation of temperature was only 0.3 K. In the present study, the variation of temperature is 1.7 K and it can be noticed an overestimation for the coarse mesh. Temperature is an important parameter in PEM fuel cells investigation since can influence the performance and durability, that is why it is important to have accurate results. In general, the increase of the temperature is taking place in the active layers of a PEM fuel cell, in the region beneath the flow channels, where the electrochemical reactions occur and this behaviour can be noticed in simulation results presented in Fig. 7. A good estimation of temperature variation can help in ensuring a good durability of the PEM fuel cell since can prevent the membrane degradation, therefore finding a grid independent solution is a requirement.

Because some changes in the variables and parameters involved in PEM fuel cell operation are difficult to be visualize in simplified geometries, this study was extended to the investigation of a 7 serpentine and a 7-parallel channels fuel cells, as presented

in the Fig. 4. The 92 cm^2 active area PEM fuel cells developed previously [41] have been used further in this chapter to analyse phenomena of interests, in conjunction with interpolation schemes used in the spatial discretization. Many studies analysed simplified geometries, i.e. single serpentine channel, but the results obtained in such investigations cannot be extended to real PEM fuel cells with complex geometries and the validation with experimental data cannot be done.

Figure 9 presents the hydrogen mass fraction profile for the two configurations investigated in the anode side: channel, gas diffusion layer, catalyst layer. The general trend for reactants in PEM fuel cells is to decrease from the inlet to the outlet since they are consumed by the electrochemical reactions. Using the same model, material properties and boundary conditions, the influence of the flow channel configuration have been done and the results are displayed in Fig. 9. It can be seen a small difference, namely the 7-parallel channels do not provide a uniform profile that may lead to a decrease of performance or to degradation of components due to limited mass transport of reactants and products. Poor water management (drying or flooding), hot spots due to an increase of temperature or a decrease in the local availability of the reactants are just few issues that can appear due to an improper distribution of reactants, and therefore need to be investigated. Hence, accurate CFD analyses become essential for any product development due to the fact that are supplementing experiments and testing activities.

As mentioned, another part of the present study dealt with investigating the influence of interpolation schemes used in spatial discretization. In ANSYS Fluent [30] there are implemented the following schemes: first order upwind, second order upwind, power law and QUICK. It is known that higher-order interpolation schemes, such as the second-order upwind and QUICK, are offering improved accuracy for the numerical simulations. The basic difference between the first and second order schemes is the number of points used in the numerical simulations, in the first order

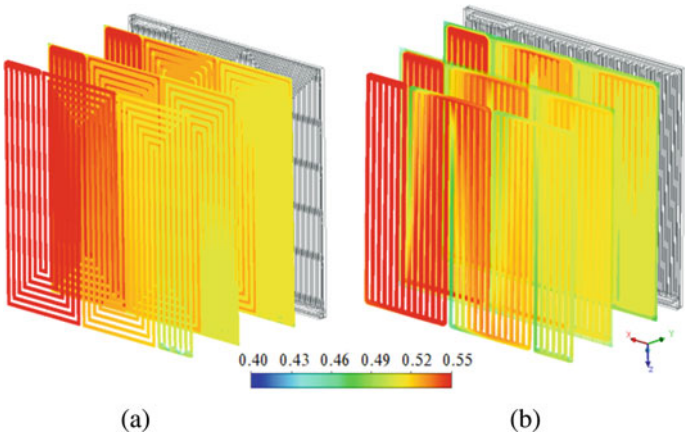


Fig. 9 Hydrogen mass fraction in anode side (channel, GDL. CL) of a PEM fuel cell with **a** serpentine channels and **b** parallel channels

being used one upstream point and in the second order two points. As regarding the QUICK scheme, a combination between the second-order upwind and the central differencing methods is used. The second order scheme is recommended to be used in CFD simulations; however, the investigation can be started with first order scheme and after obtaining a convergent solution the method can be switched to second order for better accuracy and more reliable results.

The PEM fuel cells with 7-serpentine and 7-parallel channels have been used to run simulations with 3 interpolation schemes: first order, second order and QUICK. The results plotted as pressure distribution and water mass fraction profiles are presented in Figs. 10 and 11. It can be seen that there are some differences in the results, not only the interpolation scheme improves accuracy of the results with increasing the order but also the investigated geometry can accentuate the influence of the interpolation scheme on the results, as can be seen for the 7-parallel PEM fuel cell results. Higher order schemes show improved accuracy of the computational results but suffer from the boundedness problem, meaning that the solution may display unphysical oscillations in regions of steep gradients such as the U turns from the serpentine and parallel configurations. In addition, the QUICK scheme minimizes the false diffusion problems in comparison with other schemes.

It is known that a converged and mesh-independent solution on a well-posed problem will provide valuable results only if the numerical methods are appropriate and the interpolation schemes are of high order.

The CFD model applied to these configurations has been validated previously [41] and was not considered in the present study. Briefly, a simplified single-phase, steady state, isothermal, 3D mathematical model has been used to investigate the physical and electrochemical phenomena that take place inside a fuel cell with two types of gas flow fields: 7—serpentine and 7—parallel channels. The numerical investigation revealed that both configurations can be used in experimental development since no significant difference was observed in the I-V curve, the serpentine configuration having a slightly better performance with 3.4% more than the parallel configuration in terms of current density. Since the serpentine configuration provided a uniform distribution of the reactants and of other parameters of interest, which in time can lead to an increased reliability and durability of the fuel cell, then the serpentine configuration has been manufactured and used in experiments. Moreover, our study revealed that at high current densities (above 0.8 A/cm^2) the numerical results over predict the real performance of the fuel cell and this is due to some assumptions made in the numerical model (water in vapour form, isothermal conditions). Therefore, it is important how the real operation of a device is transposed into a mathematical model, what boundary conditions and assumptions are taking into account, what numerical method is chosen, because all these are influencing the accuracy and reliability of a solution.

In general, for CFD modeling of PEM fuel cells, the I-V curve is used for model validation by comparing the results of numerical simulations with experimental data. However, this comparison is not always a good indicator of agreement between CFD and experiments since under-prediction of one parameter can lead to an over-prediction of another parameter and the I-V curve can remain unchanged. Due to

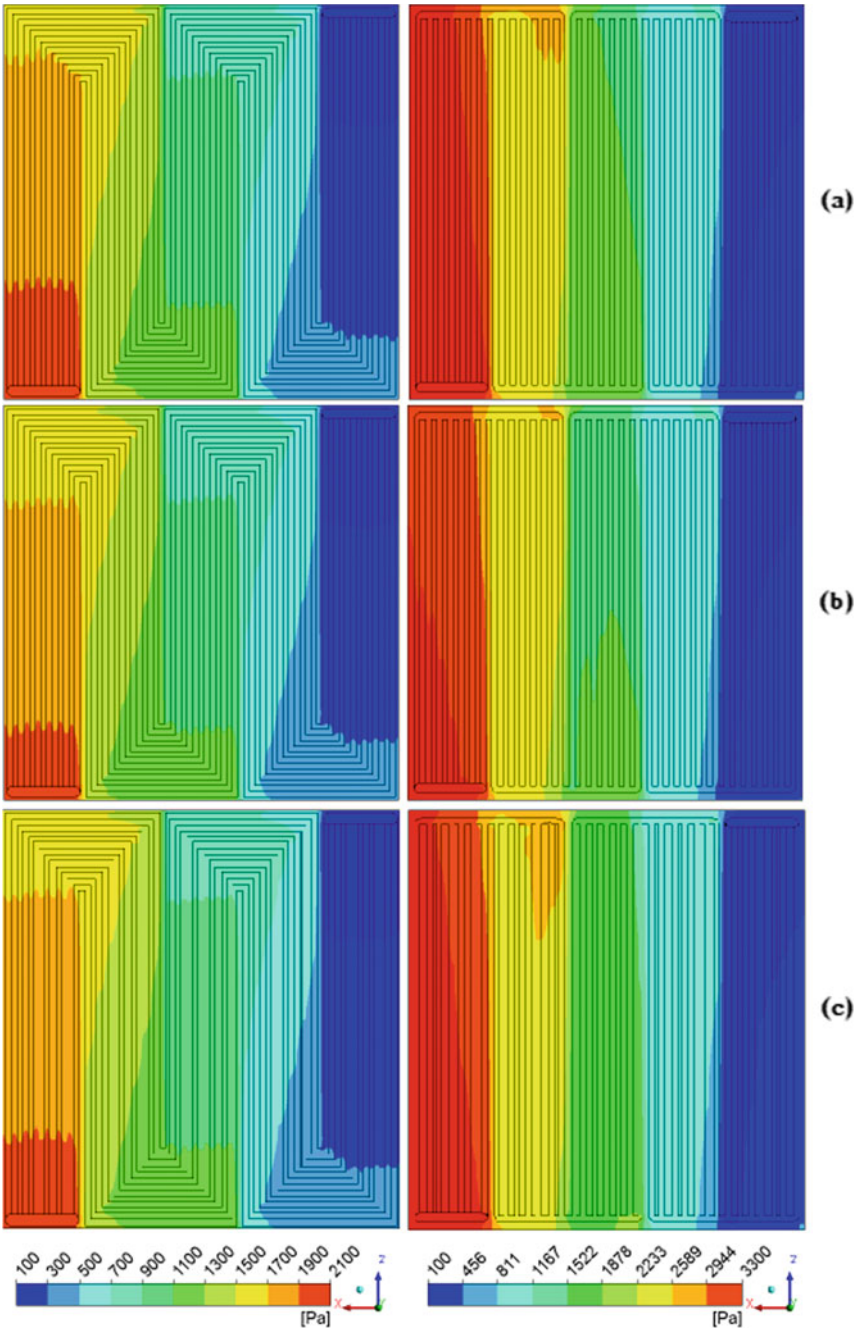


Fig. 10 Pressure distribution in a PEM fuel cell with 7-serpentine (left) and 7 parallel (right) channels for: **a** first order upwind; **b** second order upwind; **c** QUICK interpolation schemes

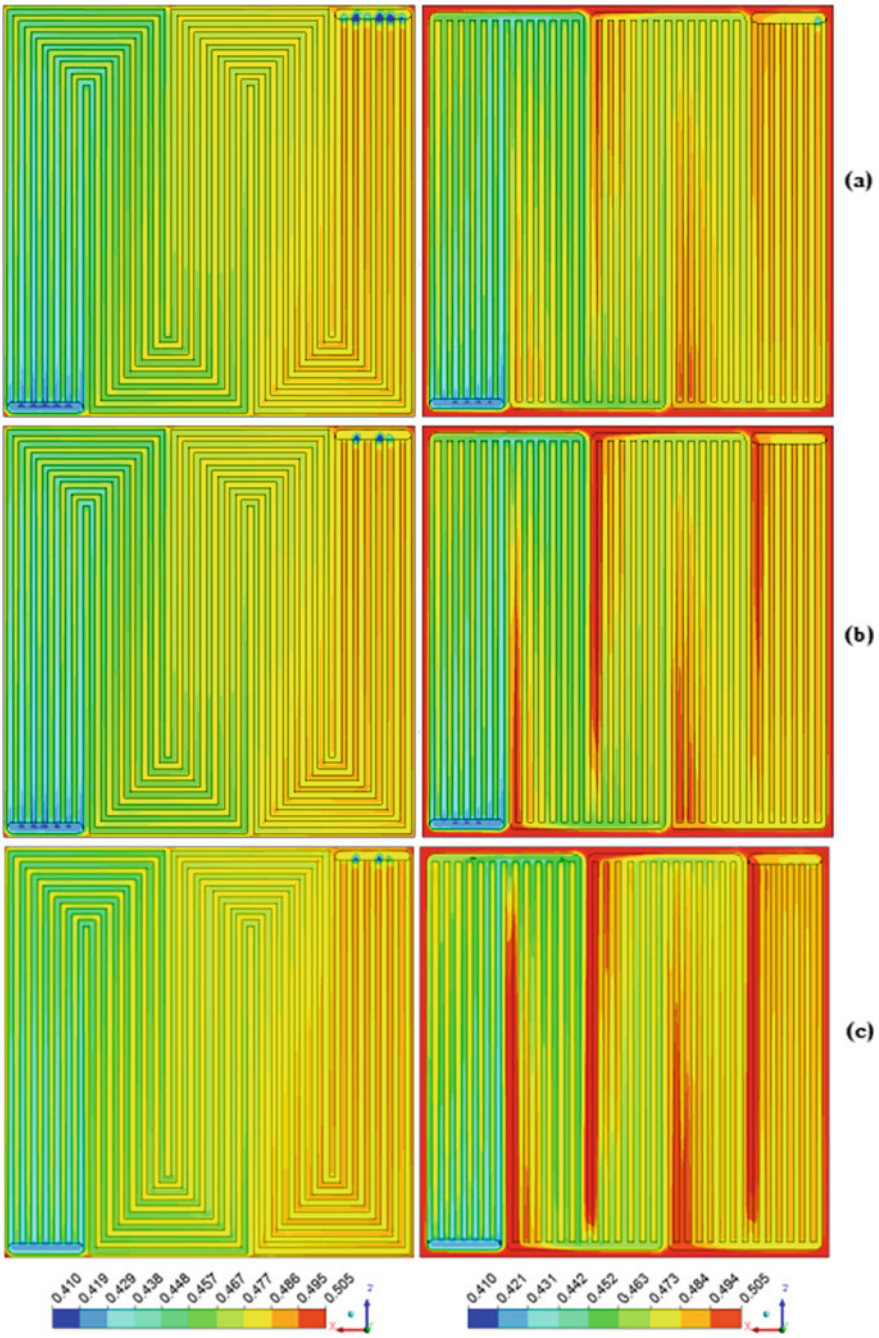


Fig. 11 Water mass fraction profile in the anode catalyst layer of a PEM fuel cell with 7-serpentine channels (left) and 7 parallel channels (right) for various interpolations schemes: **a** first order upwind; **b** second order upwind; **c** QUICK

the fact that our goal was to investigate the influence of spatial discretization and interpolation schemes on the changes of variables of interests for PEM fuel cell operation and that the differences would have been difficult to visualize by a simple I-V curve, the results of the numerical simulations have been presented in this chapter mainly as contours and profiles.

6 Conclusions

The study dealt with the FVM and its application to explore the effect and the importance of two aspects relevant for numerical investigations: the spatial discretization and the interpolation schemes, with application on analysing the operation and performance of a PEM fuel cell. The following are the main findings of the study:

(i) A general overview on the finite volume method and on the numerical methods used for solving linear system is given.
(ii) A short introduction in CFD and a description of the steps involved in such investigation is presented, the CFD analysis being afterwards applied to examine the operation of a PEM fuel cell in two case studies: one with a simplified configuration and one for a real configuration.
(iii) A mathematical model is developed for all components of the fuel cell using a general scalar transport equation, followed by the numerical investigation. Problem definition, geometry design with spatial discretization, solving and post-processing steps are discussed in the chapter underlying the influence of spatial discretization and of three interpolation schemes (first order upwind, second order upwind, and QUICK) on the accuracy of the results from a PEM fuel cell numerical investigation.
(iv) The spatial discretization study was based on a simplified geometry for a PEM fuel cell (a configuration with a single serpentine channel). 5 meshes have been developed and the study showed that it is a requirement for the accuracy of the results to ensure the grid independence of the solution. Contours and profiles for pressure and temperature have been plotted to visualize the influence of the mesh developed on the PEM fuel cell performance and durability and the results are discussed taking into account the impact of the under-estimation or over-estimation of these parameters for the overall operation process.
(v) The influence of interpolation schemes used in numerical investigations revealed that an improved accuracy of the results is obtained if higher order schemes are used. Also, it was found out that the geometry used in the investigation is another factor that may influence the results. If for the 7-serpentine configuration the interpolation schemes did not show much differences in the results plotted not the same finding was noticed for the 7-parallel configuration where the accuracy of the results improved with increasing the order of the interpolation scheme.

(vi) The chapter revealed the importance of the Computational Fluid Dynamics (CFD) for many investigations, being a reliable tool for several applications and an alternative for experimental tests since are cheaper and faster.

Acknowledgements This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI - UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0194/25 PCCDI within PNCDI III and contract 117/2016, RESTORE project.

References

1. Barbir F (2012) PEM fuel cell: theory and practice, 2nd edn. Elsevier Academic Press, Amsterdam, Boston. ISBN: 9780123877109. https://doi.org/10.1016/B978-0-12-078142-3.X5000-9
2. Berning T, Djilali N (2003) A 3D, multiphase, multicomponent model of the cathode and anode of a PEM fuel cell. J Electrochem Soc 150:A1589–A1598
3. Springer TE, Zawodzinski TA, Gottesfeld S (1991) Polymer electrolyte fuel cell model. J Electrochem Soc 138:2334–2342
4. Gurau V, Barbir F, Liu H (2000) An analytical solution of a half-cell model for PEM fuel cells. J Electrochem Soc 147(7):2468–2477. https://doi.org/10.1149/1.1393555
5. Chevalier S, Josset C, Auvity B (2018) Analytical solutions and dimensional analysis of pseudo 2D current density distribution model in PEM fuel cells. Renew Energy 125:738–746. https://doi.org/10.1016/j.renene.2018.02.120
6. Kulikovsky AA (2017) Approximate analytical solution to MHM equations for PEM fuel cell cathode performance. Electrochem Comm 77:36–39. https://doi.org/10.1016/j.elecom.2017.09.018
7. Um S, Wang CY, Chen KS (2000) Computational fluid dynamics modeling of proton exchange membrane fuel cells. J Electrochem Soc 147(12):4485–4493
8. Blazek J (2015) Computational fluid dynamics: principles and applications, 3rd edn. Butterworth-Heinemann Imprint, pp 73–120. ISBN 978–0–08–099995–1. https://doi.org/10.1016/C2013-0-19038-1
9. Ma L, Ingham DB, Pourkashanian MC (2005) Application of fluid flows through porous media in fuel cells. In: Transport phenomena in porous media III, 1st edn. Elsevier Science, pp 418–440. https://doi.org/10.1016/B978-0-08-044490-1.X5003-0
10. Heister T, Rebholz LG, Xue F (2019) Numerical analysis—an introduction. De Gruyter Textbook, pp 121–140. ISBN 978-3-11-057330-5
11. Walter E (2014) Numerical methods and optimization. Springer International Publishing, pp 17–57. ISBN 978-3-319-07670-6. https://doi.org/10.1007/978-3-319-07671-3
12. Moukalled F, Mangani L, Darwish M (2016) The finite volume method. In: Computational fluid dynamics, an advanced introduction with OpenFOAM® and Matlab®. Fluid mechanics and its applications, vol 113. Springer International Publishing, pp 103–135. ISSN 0926-5112. https://doi.org/10.1007/978-3-319-16874-6
13. Versteeg HK, Malalasekera W (1995) Computational fluid dynamics (CFD) is the application of algorithm and numerical techniques to solve fluid flow problems. Wiley, pp 85–204 ISBN 0-582-21884-5
14. Barth T, Ohlberger M (2004) Finite volume methods: foundation and analysis. In: Stein E, de Borst R, Hughes TJR (eds) Encyclopedia of computational mechanics. Part 1 fluids. Wiley. https://doi.org/10.1002/9781119176817.ecm2010
15. Marinoiu A, Cobzaru C, Carcadea E et al (2015) An experimental approach for finding low cost alternative support material in PEM fuel cells. Rev Roum Chim 61:433–440

16. Bizon N (2019) Sensitivity analysis of the fuel economy strategy based on load-following control of the fuel cell hybrid power system. Energy Convers Manag 199:111946. https://doi.org/10.1016/j.enconman.2019.111946
17. Karimi G, Baschuk JJ, Li X (2005) Performance analysis and optimization of PEM fuel cell stacks using flow network approach. J Power Sources 147(1–2):162–177. https://doi.org/10.1016/j.jpowsour.2005.01.023
18. Russel J, Cohn R (2012) Gaussian elimination. ISBN 9-785510890709
19. Mittal RC, Al-Kurdi A (2002) LU-decomposition and numerical structure for solving large sparse nonsymmetric linear systems. Comput Math Appl 43(1–2):131–155. https://doi.org/10.1016/S0898-1221(01)00279-6
20. Mazumder S (2016) Numerical methods for partial differential equations. In: Finite difference and finite volume methods, pp 1–49. ISBN 978-0-12-849894-1
21. Ferziger JH, Peric M (1997) Computational Methods for Fluid Dynamics, Springer-Verlag Berlin Heidelberg, pp 21–37, 71–89, doi: https://doi.org/10.1007/978-3-642-56026-2
22. Anderson J (1995) Computational fluid dynamics—the basics with applications. McGraw-Hill, pp 216–278. ISBN 0-07–001685-2
23. Zawawi MH, Saleha A, Salwa A et al (2018) A review: fundamentals of computational fluid dynamics (CFD). AIP Conf Proc 2030:020252. https://doi.org/10.1063/1.5066893
24. Carcadea E, Varlam M, Marinoiu A et al (2019) Influence of catalyst structure on PEM fuel cell performance—a numerical investigation. Int J Hydrogen Energy 44:12829–12841. https://doi.org/10.1016/j.ijhydene.2018.12.155
25. Carcadea E, Varlam M, Ismail M et al (2019) PEM fuel cell performance improvement through numerical optimization of the parameters of the porous layers. Int J Hydrogen Energy 45(14):7968–7980
26. Tu J, Yeoh GH, Liu C (2019) Computational fluid dynamics—a practical approach. Elsevier Ltd. ISBN 978-0-08-101127-0. https://doi.org/10.1016/C2015-0-06135-4
27. Barbir F. PEM fuel cells. In: Sammes N (ed) Fuel cell technology: reaching towards commercialization. Springer International Publishing, pp 27–51. ISBN: 978-1-84628–207-2. https://doi.org/10.1007/1-84628-207-1
28. OpenFoam Software. https://openfoam.org/
29. HELYX Sotware. https://engys.com/products/helyx
30. ANSYS. Multiphysics help. www.ansys.com
31. COMSOL Multiphysics. https://www.comsol.com/cfd-module
32. Inoue G, Matsukuma Y, Minemoto M (2006) Effect of gas channel depth on current density distribution of polymer electrolyte fuel cell by numerical analysis including gas flow through gas diffusion layer. J Power Sources 157:36–152
33. Carcadea E, Ingham DB, Stefanescu I et al (2011) The influence of permeability changes for a 7-serpentine channel PEM fuel cell performance. Int J Hydrogen Energy 36:10376–10383
34. Ismail MS, Hughes KJ, Ingham DB et al (2013) Effect of PTFE loading of gas diffusion layers on the performance of proton exchange membrane fuel cells running at high-efficiency operating conditions. Int J Energy Res 37:1592–1599
35. Jang WK, Choi J, Seo YH et al (2015) Effect of cathode flow field configuration on air-breathing proton exchange membrane fuel cell. Int J Precis Eng Man 16:1129–1134
36. Khakaz-Baboli M, Harvey DA, Pharoah JG (2013) Investigating the performance of catalyst layer micro-structure with different platinum loadings. ECS Trans 50(2):765–772
37. Raceanu M, Marinoiu A, Culcer M et al (2014) Preventing reactant starvation of a 5 kW PEM fuel cell stack during sudden load change. In: Proceedings of 6th international conference on electronics, computers and artificial intelligence, pp 55–60. https://doi.org/10.1109/ECAI.2014.7090147
38. Sanchez DG, Ruiu T, Friedrich KA et al (2016) Analysis of the influence of temperature and gas humidity on the performance stability of polymer electrolyte membrane fuel cells. J Electrochem Soc 163:F150–F159
39. Obut S, Alper E (2011) Numerical assessment of dependence of polymer electrolyte membrane fuel cell performance on cathode catalyst layer parameters. J Power Sources 196:1920–1931

40. Antunes RA, de Oliveira MCL, Ett G et al (2011) Carbon materials in composite bipolar plates for polymer electrolyte membrane fuel cells: a review of the main challenges to improve electrical performance. J Power Sources 196(6):2945–2961
41. Carcadea E, Varlam M, Stefanescu I et al (2014) (2014) Effects of flow fields on PEM fuel cell performance. Prog Cryog Isot Sep 17:81–88
42. Guilin HU, Xu Y, Zhang Z (2014) Numerical simulation of heat/mass transfer in a single proton exchange membrane fuel cell with serpentine fluid channels. Int J Electrochem Sci 9:1902–1910
43. Yuan XZR, Song C, Wang H et al (2010) PEM fuel cells and their related electrochemical fundamentals. In: Electrochemical impedance spectroscopy in PEM fuel cells. Springer, London, pp 1–37. https://doi.org/10.1007/978-1-84882-846-9_1
44. Kazemi Esfeh H, Azarafza A, Hamis MKA (2017) On the computational fluid dynamics of PEM fuel cells (PEMFCs): an investigation on mesh independence analysis. RSC Adv 7:32893–32902

Night Operation of a Solar Chimney Integrated with Spiral Heat Exchanger

Amel Dhahri, Ahmed Omri, and Jamel Orfi

Abstract Solar power is one of the most raising and encouraging renewable source of energy generation. Solar plants are playing an important role in power supplies worldwide. Nowadays, the electrical energy demand is increasing rapidly due to fast-growing daily requirements. In the last few decades, scientific researchers have focused on a novel technology called the solar chimney power plant, sometimes recognized as solar updraft tower. The solar chimney is principally composed of three main constituents, namely, a solar collector, a chimney and a wind turbine. This promising technology addresses a very challenging idea of generating electricity from free solar energy. It is categorized as a viable resource of clean energy for many non developed countries. The world's first solar chimney prototype was designed and constructed at Manzanares in Spain, as a result of a joint project between the Schlaich Bergermann partner and the Spanish government. The plant is *characterized* by a tower high 195 m with a radius of 5 m. The radius and height of the collector encircling the tower are respectively 120 m and 1.85 m. The spanish prototype built by the engineer Jorg Schlaich of Schlaich Bergermann operated without significant problems for seven years. *Several research projects* have been conducted all *over the world* to design and introduce different solar towers based on experience gained from operating the 50 kW Spanish prototype. To ensure that the energy conversion is maintained at satisfactory levels to guarantee considerable power generation, an unprecedentedly high tower and an immense collector area are needed. This plant is then based on a thermal updraft movement of hot air resulting from *natural convection*. In this chapter, the simulations were conducted using cylindrical coordinate

A. Dhahri (✉) · A. Omri
Research Unit: Materials, Energy and Renewable Energies, College of Sciences, University of Gafsa, Gafsa, Tunisia
e-mail: dhahriamel001@yahoo.fr

A. Omri
e-mail: ahom206@yahoo.fr

J. Orfi
Department of Mechanical Engineering, King Saud University, Riadh, Saudi Arabia
e-mail: orfi@ksu.edu.sa

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_14

system. The inner fluid flow is considered turbulent and simulated with the k-ε turbulent model, by means of the CFD commercial software ANSYS Fluent. Numerical data were validated by comparing them with those from experiments. The agreement between *simulation results* and the *measurements* taken from the *experimental* plant in *Manzanares* is fairly *good.* A set of mathematical models of the solar updraft power plant have been developed where a model considering the kinetic energy difference within the solar collector was proposed. The operation of such a plant is strongly dependent on the amount of solar radiation. The main disadvantage of this system is the inability to operate constantly at night. A geothermal heating device is suggested to guarantee a continuous operation during night hours. In this chapter, an auxiliary heating system constructed of in-plane spiral coil tubes is proposed to be placed above the ground under the collector. Thus, the computational model is afterward combined with a mathematical model for a geothermal heat exchanger to evaluate the effect of combining both solar and geothermal energy on the plant performance. A parametric *study of the hybrid plant* is carried out. The study *focus essentially* on the impact of the collector size, the meteorological conditions as well as the effectiveness of the heat exchanger on the air flow rate, the temperature increase within the collector and the global performance of the solar-geothermal hybrid system.

Keywords Solar energy · Solar chimney · Collector · CFD · Electric power · Geothermal energy

Nomenclatures

List of symbols

A	Area (m^2)
A_{coll}	Solar collector area (m^2)
A_0	Heat transfer area per unit tube length
D_{tube}	Tube diameter (m)
C_p	Specific heat capacity (J/kg K)
G	Solar radiation (W/m^2)
h	Heat transfer coefficient ($Wm^{-2}K^{-1}$)
$\dot{m}$	Mass flow rate (kg s^{-1})
n_{tube}	Tube number
R	Collector radius (m)
T	Temperature (K)
U_i	Total heat transfer coefficient (W/m^2 K)
V	Air flow velocity (ms^{-1})

Greek symbols

λ	Thermal conductivity (W m^{-1} K^{-1})
μ	Dynamic viscosity (kg (s m)$^{-1}$)
ρ	Density (kg m^{-3})
$\tau\ \alpha$	Transmittance-absorbtance product
α	Rate of change of the radius

Subscript

c	Collector canopy
e	Environment or exterior
f	Fluid
i	Inner
m	Average
o	External
n	Turn number
r	Storage reservoir
s	Soil or Solar
w	Geothermal water
w, in	Heat exchanger inlet
w, out	Heat exchanger exit
x	Distance along tube (m)
1	Solar collector inlet
2	Solar collector exit
ΔT	Temperature rise (K)

1 Introduction

The ever present concern of humanity and the great challenge facing the world today is to achieve sufficient supplies covering its future demand for clean energy. Renewable energy provides a growing and vital contribution to developing countries. The proper implementation of renewable energy sources in developing countries could reduce their dependence on natural gas and oil, and investments in renewable energy technologies would be more cost-effective than fossil fuels. Nowadays, geothermal and solar energies are considered very attractive because of their world's huge resources and their important *electrical production capacity*. Geothermal power may be combined with solar energy to harness the advantages of both technologies. There are several methods of hybridizing solar and geothermal technologies, and the viability of certain methods depends on factors such as location, relative geothermal

and solar resource quality. There is a wide variety of solar energy systems. Over the years, solar chimneys attracted notable attention. This system which is more appropriate to arid and semi-arid areas is an advanced way to generate electricity from solar radiation. According to the literature, several prototypes with different scales have been constructed. The first prototype was actually placed and checked in Manzanares (Spanish town 150 km from Madrid) under the direction of German engineer Jörg Schlaich, and with funding from the German Ministry research and technology [1, 2].

The operation of such a plant is strongly dependent on the amount of solar radiation. Among all solar energy technologies, Solar chimneys are attracting special interest since they are able to operate during the day and night time [3]. The main disadvantage of this plant is that it is unable to work continuously at night. Numerous studies suggested different solutions to guarantee a continuous operation during night hours. The release of thermal energy stored during the day in the ground and/or in water tanks can overcome the disturbance of power production. This guarantees that the solar chimney is operational even at night [4]. For this reason, the Engineer Jörg Schlaich proposed additional heat storage mediums to improve the storage capacity and enable the functioning of the plant during the night. The proposed solution is to lay taught water-filled pipes underneath the collector cover. The water is heated during the day and releases this heat during nighttime [1]. In order to improve the heat storage capacity, Kreetz introduced a proposition to place water-filled tubes/tanks under the collector cover [5]. Pretorius suggested the use of a secondary collector roof to enhance the performance of the plant at night [6]. The combination of a *salt-gradient solar pond with the solar collector* was investigated by Zhou et al. The goal of this alternative is a continuous night operation for solar chimney power system [7]. To enhance the plant and ensure uninterruptible power generation, Al-Kayiem et al. proposed a new approach based on coupling the updraft tower with an exhaust heat source. The aim of this technique was to use wasted flue-gas to supply air underneath the collector [8]. Some researchers focused on enhancing the ground under the collector and employing it as storage layer [9–13]. Others suggested and investigated the usage of closed water-filled tanks [14–16]. Another alternative is proposed by Zubir et al. They suggested a *waste heat energy recovery system* based on passing the flue gases through conduits inside the collector [17]. The integration of a solar chimney power plant with another source of heat has been proposed to build a day and night functional integrated hybrid system. This proposal aims to improve the plant performance and overcome the absence of solar energy during night and cloudy days [18].

Various alternatives were proposed and investigated to enhance the performance of the solar chimney but still insufficient to allow a continuous operation of the system. Recently, some researchers introduced a novel approach to improve the solar chimney performance by hybridizing the solar plant and geothermal energy system. Reference [18] proposed a hybrid solar chimney to improve power production during the day, cloudy days, or during nighttime. The main target of this novel design is to exploit underground pipes supplied by low-temperature geothermal water to heat the

air underneath the collector and at the same time employing the PV thermal panels to produce electric energy by sunlight besides solar chimney power generation.

Other few studies suggested the integration of geothermal water tubes as heating equipment. Cao et al. studied the performance of the geothermal-solar chimney having the same Manzanares prototype dimensions. The authors compared three operational modes: the solar mode, the geothermal mode and the hybrid geothermal–solar mode [19]. The authors suggested the use of water tubes as heating device. They indicated that introducing geothermal water makes greater contribution and can achieve a 24 h-operation. To attain uninterrupted power production and enhance the energy generation of the solar chimney, Dhahri et al. considered a novel heating technique to supplement the energy input by the use of geothermal energy. A radial tubes configuration is selected for the heat exchanger. A study is conducted to examine the performance of the new hybrid power generation system for different operation modes. They concluded that by using geothermal energy, the plant would be able to deliver power during night [20].

A limited number of studies have investigated the performance of solar-geothermal chimney during nighttime. This chapter is a part of a project investigating the feasibility of coupling solar technology with geothermal energy storage to provide financially viable, reliable, and resilient energy. Existing literature has focused mainly on the daytime operation mode. The Main purpose of this work is to undertake a detailed numerical study of a hybrid solar-geothermal chimney during night based on mathematical models. The present work proposed to place an auxiliary heating system constructed of in-plane spiral coil tubes above the ground. In this chapter, a 3D numerical study was performed based on the Manzanares (Spain) prototype as a reference. Firstly, a detailed thermal model of geothermal coiled pipes is built. To evaluate the performance of the solar chimney, the mathematical models of the conventional and hybrid plants are established considering the variation of kinetic energy inside the collector. Then, the model data are validated by comparing them to Manzanares experimental measurements. To accomplish the goals of the current work, the impact of the main control parameters such as the geometric parameters and operational parameters on the novel plant performance were systematically investigated.

2 System Description and Basic Principle

2.1 Conventional Solar Chimney

A conventional solar chimney (Fig. 1) converts solar energy into electricity using three basic components: A circular collector suspended at certain height from the ground, a central chimney and a wind turbine. The Air is heated by solar radiation under the transparent collector open at the periphery. The warm air moves toward and upward into the chimney resulting from air buoyancy and difference in air pressure

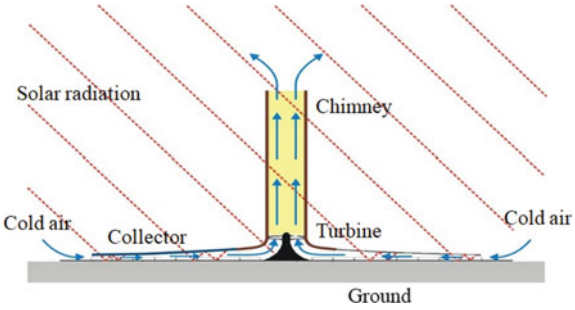


Fig. 1 Conventional solar chimney power plant scheme

between the inside and the outside of the plant [3]. The huge pressure difference between the plant and the ambient air (chimney effect), will run the wind turbine set up at the bottom of the chimney [21]. The energy contained in the updraft air is transformed into mechanical energy by the turbine and then the generator converts it into electric energy.

The solar chimney is characterized by a 195 high chimney with a constant diameter of 10 m. the solar collector area extended to a radius of 120 m from the chimney with the glazing being 1.85 m above the ground. The turbine is housed at the bottom of the chimney [5]. The concept has been verified by constructing and successfully running a prototype in Spain for more than 6 years [22].

2.2 Hybrid Power Plant

Heating air underneath the collector is needed during the night. Therefore, it is necessary to place auxiliary heating equipment. The heat supply by geothermal energy can overcome the deficits of sunshine in the absence of solar radiation (during night hours). Arid areas of Tunisia have important resources of geothermal water. The range of water temperature in southern Tunisia is between 65 and 80 °C [23]. Different techniques can be used for air heating. One of the most appropriate and advantageous heating system is the air-soil heating system with heat pipes placed on the soil. The most used and known is heating systems with pipes above the ground. In this study, the heating is ensured via spiral pipes lying on the ground. The winding of the tube leads to obtain the body of the exchanger. The material chosen for its construction is PVC. The heating system is composed of poly-circular tube arranged on the collector floor allows by exploiting the heat from the geothermal water to heat the air entering from the periphery of the collector. This system is connected to a main supply reservoir placed near the location of the geothermal well. The exploited well is situated in the center of the plant.

The operation of the hybrid power plant is simple. The air flows from the periphery (outermost turn) to the center of the collector. Water enters the inner turn located in the center of the collector with a hot temperature and exit at the end of the spiral tube

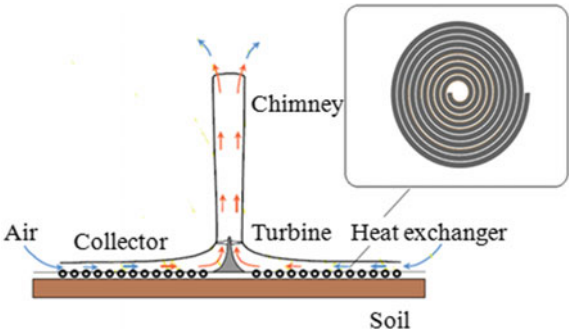


Fig. 2 Schematic diagram of a solar chimney coupled with a spiral exchanger

with a cooler temperature. The hot air passes from the solar collector to chimney base and moves up to the top of the chimney. It is important to keep in mind that the outlet temperature of the geothermal water should be as low as possible. Thus, the mode of circulation of the two fluids (air and water) is cross flow (Fig. 2).

The radius of the reservoir is equal to 15 m. The space between turns was fixed at 0.2 m. The distribution of heat is considered uniform, leading to neglect local heating heterogeneities along the radius of collector.

The heat exchanger design with its main geometric parameters is presented in Fig. 3. Moreover, its dimensions are illustrated in Table 1. This exchanger can be characterized by the diameter of its tube (d), the evolution of its radius of curvature (Ri) and the pitch between consecutive turns (b). These indications are sufficient to deduce the total length of the wound tube. A constant step between two consecutive turns is guaranteed. But, the curvature radius is inconstant. Other indications may be useful to complete the geometrical characterization: the radius (internal and/or external) of the tube and the tube thickness. There could be several turns, but it is limited by the flow developed in the tube.

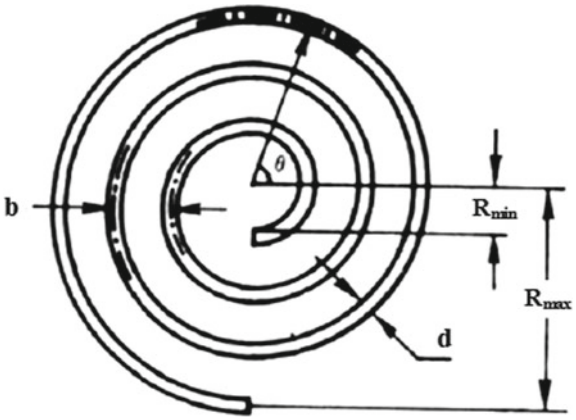


Fig. 3 Schematic diagram of a spiral exchanger

Table 1 Main dimensions

Description	Value	Unit
Tube diameter	0.2	m
Water flowrate	85.6	Kg/s
Pitch between tow turns	0.2	m
Interior radius	5	m
External radius	120	m
Turn number	92	
Inlet water temperature	70	°C

3 Thermal Modeling of the Heat Exchanger

Coiled tubes are used *commonly* in heat exchanger equipments thanks to the improved heat transfer rate resulting from secondary flows [24]. On closer inspection, higher temperature gradient is detected close to the external wall of the coiled tubes, indicating a higher transfer rate.

This is mostly contributed by the secondary flow which is induced by centrifugal forces due to the curvature of the coil [25]. As such, coiled tubing has attracted particular attention from researchers around the world.

In the case of spiral tubes, the problem is complex because the outlet temperatures vary depending on the position of the fluid in the exchanger. Theoretical energy equations are expressed for a turn of the coil. In the theoretical model, each turn of the heat exchanger is supposed to be a linear spiral. R_{i} and R_{o} are the inlet and outlet coil radius, respectively. α represents the rate of change of the radius [24, 26].

In addition, each full rotation (2π) is approximately considered circular. The mean radius of i-th turn, R_{mi}, is defined as:

$$R_{mi} = \frac{1}{2\pi} \int_0^{2\pi} R d\theta \tag{1}$$

$$R_{mi} = R_o + (2i - 1)\alpha\pi \tag{2}$$

The outlet heat transfer area (A_{i}) of i-th turn is expressed as follow:

$$A_i = 2\pi R_{mi} A_0 \tag{3}$$

The energy balance equation obtained for a small control volume in i-th turn is given by:

$$\dot{m}_w C_{p,w} dT_w = U_i A_0 [T_{\mathrm{air}} - T_w(x)] dx \tag{4}$$

The integration of (4) leads to the expression:

$$\left[T_{\text{air},i} - T_{w,i+1}\right] / \left[T_{\text{air},i} - T_{w,i}\right] = Exp(\lambda_i) \tag{5}$$

where

$$\lambda_i = (A_i U_i) / \left(\dot{m}_w C_{p,w}\right) \tag{6}$$

The equation of global energy balance developed for the i-th turn leads to:

$$\dot{m}_w C_{p,w}\left[T_{w,i} - T_{w,i+1}\right] = \dot{m}_{air} C_{p,\text{air}}\left[T_{\text{air},i} - T_{\text{air},i+1}\right] \tag{7}$$

where

$T_{air,i}$ air temperature across the i-th turn, °C
$T_{air,i+}1$ air temperature across the i + 1-th turn, °C

$$\dot{m}_{\text{air}} C_{p,\text{air}}\left[T_{\text{air},i} - T_{\text{air},i+1}\right] = \dot{m}_w C_{p,w}\left[T_{\text{air},i} - T_{w,i}\right][Exp(\lambda_i) - 1] \tag{8}$$

The approach considered in this section is to separate the geothermal heat exchanger in a sufficient number of portions. The water temperature along a plane spiral coil tube ($T_{\text{w, i+1}}$) is calculated iteratively from the air temperature in the vicinity of the coil, the temperature ($T_{\text{w, i}}$) and the heat flow rate in the preceding turn (Fig. 4). In order to determine the heat transported by geothermal water to the indoor air, the water outlet temperature is required. For this reason, an iterative calculation based on (7) and (8) must be conducted to calculate this temperature at the exit of each turn for a given inlet temperature.

The amount of heat transported by the hot water along the pipe is evaluated by:

$$\dot{Q}_{\text{geoth}} = \dot{m}_w C_{p,w}\left(T_{w,e} - T_{w,s}\right) \tag{9}$$

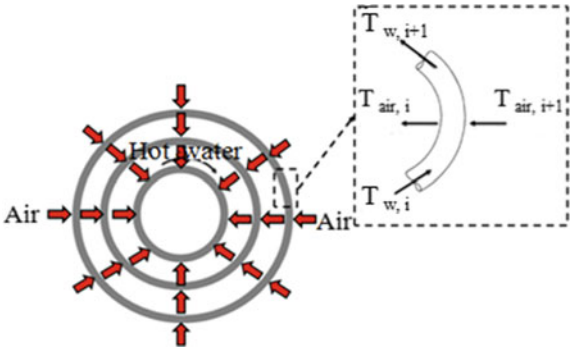


Fig. 4 Descriptive diagram of a spiral exchanger

4 Mathematical Model

The most important element of a solar chimney system is the collector. In this section, a complete theoretical model is *developed* to investigate the thermal collector performance. Both conventional and hybrid solar chimney power plant are considered in this section.

4.1 Conventional Solar Chimney

We propose to take into account the kinetic energy variation in the collector energy equation. A new term is introduced in the global energy balance of the collector. When relevant literature in solar chimney systems is reviewed, it is seen that the researchers had not considered kinetic energy term in their mathematical models. The energy balance in this case may be written as follows [20, 27]:

$$0 = \dot{Q}_s - \dot{Q}_p + \rho_1 A_1 V_1 \left(C_{p,a} T_1 + \frac{V_1^2}{2} \right) - \rho_2 A_2 V_2 \left(C_{p,a} T_2 + \frac{V_2^2}{2} \right) \tag{10}$$

The *Convective heat loss* from the collector to the environment is expressed as:

$$\dot{Q}_p = U_p A_{\text{coll}} (T_{\text{air},i} - T_{\text{amb}}) \tag{11}$$

Where A_1, A_2 et A_{coll} are expressed respectively by: $A_1 = 2\pi R h_{\text{coll}}$, $A_2 = 2\pi r h_{\text{coll}}$ and $A_{\text{coll}} = \pi R^2$.

The air outlet velocity of the solar collector can be calculated using the following relation:

$$V_2 = \sqrt[3]{-\frac{q}{2} - \frac{1}{2}\sqrt{\frac{27q^2 + 4p^3}{27}}} + \sqrt[3]{-\frac{q}{2} + \frac{1}{2}\sqrt{\frac{27q^2 + 4p^3}{27}}} \tag{12}$$

Where

$$p = 2C_p T_2 \tag{13}$$

$$q = -\frac{2G}{\rho_{\text{coll}} A_2} + \frac{2h_{\text{conv}} A_{\text{coll}} (T_c - T_{\text{air}})}{\rho_{\text{coll}} A_2} - 2V_1 \frac{A_1}{A_2} \left(C_p T_1 + \frac{V_1^2}{2} \right). \tag{14}$$

4.2 Hybrid Geothermal-Solar Chimney

The main restriction in renewable energy plants lies in the interrupted nature of the energy source and the requirement of a site with specific nature. But this restriction can frequently be surmounted with a combination of technologies. Since the sun is not available during nighttime and overcast days, we proposed to combine the solar chimney with an auxiliary heating system to guarantee a 24 h continuous operation. The current *research study focus is* on modeling the night mode of a geothermal-solar chimney with assumption of absence of solar radiation on the plant.

4.2.1 Thermal Heat Losses at the Collector

To carry out an energy equation on a solar collector, thermal losses associated with the collector should be evaluated for a given heat energy flow. The cover acts as an intermediary *between the collector zone* and the outdoor *environment*. The major collector heat losses are from the front roof. The heat transfer by natural convection between the interior air and the internal wall of the canopy is given by following equation [20].

$$\dot{Q}_{p,\mathrm{i}} = h_{\mathrm{i}} A_{\mathrm{coll}} (T_{\mathrm{air,i}} - T_c) \tag{15}$$

h_{i} is the convective heat transfer coefficient between the inside air and the inner face of the collector. h_{i} can be calculated by [20]:

$$h_{\mathrm{i}} = \left[0.2106 + 0.0026 v \left(\frac{\rho T_m}{\mu g \Delta T} \right)^{1/3} \right] \Big/ \left[\frac{\mu T_m}{g \Delta T c_p \lambda^2 \rho^2} \right]^{1/3} \tag{16}$$

Air properties including:density, thermal conductivity, dynamic viscosity and heat capacity are estimated at the average temperature. The air properties can be obtained based on (17–20) at temperatures varying between 300 and 350 K [20].

$$\rho = 1.1614 - 0.00353(T - 300) \tag{17}$$

$$\lambda = 0.0263 + 0.000074(T - 300) \tag{18}$$

$$\mu = [1.846 + 0.00472(T - 300)] \times 10^{-5} \tag{19}$$

$$C_p = [1.007 + 0.00004(T - 300)] \times 10^3 \tag{20}$$

The convective heat loss through the external face of the collector roof can be expressed as follows:

$$\dot{Q}_{p,e} = h_e A_{\text{coll}}(T_c - T_{\text{air},e}) \tag{21}$$

The correlation of McAdams is employed to evaluate the external convective heat transfer coefficient [20]:

$$h_e = 5.67 + 3.86 v_{\text{amb}} \tag{22}$$

The main role of the heat exchanger is to supply energy to *equalize* the heat losses taking place during periods when the indoor air temperature overpasses the outdoor temperature.

Analytical models generally include some simplifying presumptions that make the problem manageable. The soil beneath the collector cover acts as a storage medium. To facilitate numerical calculations, the energy stored up in the soil during daytime periods and released during nighttime will be neglected in this study. Comparing the heating technique (water tubes) with the ground, showed that the heat exchange between water pipes and water is much greater than that of soil surface and layers located beneath since the water heat capacity is almost five times superior than that of ground [28, 29]. Energy losses at the level of the collector cover will be neglected in our calculations and are considered equal to zero. The collector energy equation is represented by the expression (23).

$$\frac{\dot{m}_w C_{p,w}}{\rho_{\text{coll}} A_2}\left(T_{w,e} - T_{w,s}\right) + \frac{A_1}{A_2}\left(C_{p,a} V_1 T_1 + \frac{V_1^3}{2}\right) = C_{p,a} V_2 T_2 + \frac{V_2^3}{2} \tag{23}$$

4.2.2 Outlet Air Velocity (not Including Thermal Losses)

Rearranging (23) led to the equation below:

$$V_2^3 + 2C_{p,a} V_2 T_2 - \frac{2\dot{m}_w C_{p,w}}{\rho_{\text{coll}} A_2}\left(T_{w,e} - T_{w,s}\right) - 2\frac{A_1}{A_2}\left(C_{p,a} V_1 T_1 + \frac{V_1^3}{2}\right) = 0 \tag{24}$$

By solving (24), we find the formula of the outlet air velocity. This solution has the form:

$$V_2 = \sqrt[3]{-\frac{q}{2} - \frac{1}{2}\sqrt{\frac{27q^2 + 4p^3}{27}}} + \sqrt[3]{-\frac{q}{2} + \frac{1}{2}\sqrt{\frac{27q^2 + 4p^3}{27}}} \tag{25}$$

where

$$p = 2C_{p,a} T_2 \tag{26}$$

$$q = -\frac{2\dot{m}_w C_{p,w}}{\rho_{\text{coll}} A_2}\left(T_{w,e} - T_{w,s}\right) - 2\frac{A_1}{A_2}\left(C_{p,a} V_1 T_1 + \frac{V_1^3}{2}\right) \tag{27}$$

4.2.3 Outlet Air Velocity (with Thermal Losses)

In this part, the thermal losses of the collector are considered. The collector heat balance with inclusion of thermal heat losses is expressed by:

$$\frac{\dot{m}_w C_{p,w}}{\rho_{\text{coll}} A_2}\left(T_{w,e} - T_{w,s}\right) + \frac{A_1}{A_2}\left(C_{p,a} V_1 T_1 + \frac{V_1^3}{2}\right) - \frac{U_p A_{\text{coll}}}{\rho_{\text{coll}} A_2}\left(T_{\text{air},i} - T_{\text{amb}}\right) = C_{p,a} V_2 T_2 + \frac{V_2^3}{2} \tag{28}$$

$$V_2^3 + 2C_{p,a} V_2 T_2 - \frac{2\dot{m}_w C_{p,w}}{\rho_{\text{coll}} A_2}\left(T_{w,e} - T_{w,s}\right) - 2\frac{A_1}{A_2}\left(C_{p,a} V_1 T_1 + \frac{V_1^3}{2}\right) + \frac{2U_p A_{\text{coll}}}{\rho_{\text{coll}} A_2}\left(T_{\text{air},i} - T_{\text{amb}}\right) = 0 \tag{29}$$

The resolution of (29) conducts to the find out of the outlet air velocity.

5 Numerical Approach and Boundary Conditions

In the presented study, a finite volume method was applied to resolve the governing equations using the CFD software FLUENT. The pre-processor software GAMBIT was employed for creating mesh, assigning boundary conditions and determining the computational domain. The computational domain is illustrated in Fig. 5. The mesh quality can be *better controlled* with geometric splitting process. In this study, the whole computational domain is partitioned into *multiple* sub-domains (collector, transition part, and chimney), each of which is meshed independently. The part with extremely elevated pressure gradients necessitates a *fine grid* resolution. Therefore, the tetrahedral meshes were used in the transition section. The two other sub-regions are meshed with hexahedral elements [27].

After grid independence study, it was noticed that a mesh with 1,228,452 cells is adequate for numerical analysis (Fig. 5). The numerical simulation of turbulent flow is performed for 3D case of a solar chimney plant. During simulation, the turbulence model k-ε standard was used.

The solver pressure-based is specified in ANSYS Fluent. The SIMPLE scheme is used for velocity-pressure coupling [27]. The relaxation factors were as follows: pressure = 0.3, velocity = 0.7, energy = 1 and radiation = 1. A second order upwind scheme was applied for *discretization* of the velocity and temperature fields, while the

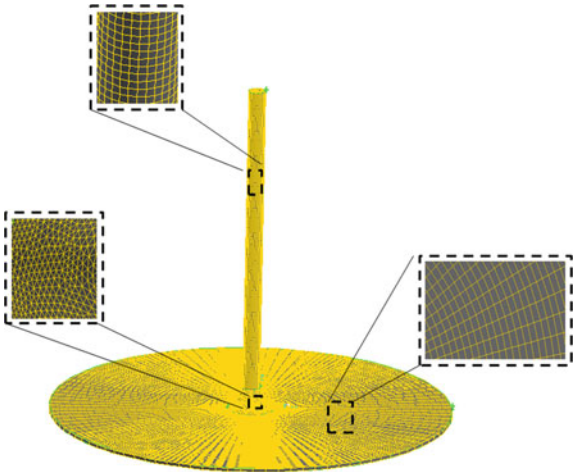


Fig. 5 Numerical grid

PRESTO *discretization* scheme was used for pressure field. The Discrete Ordinate (*DO*) *model* is used for *radiation.* In this study, the Discrete Ordinate (*DO*) *model* is used to compute *radiation.* This model is appropriate to solve the radiation through a semi-transparent medium.

For residuals of continuity and momentum equations, a convergence criterion of 10^{-3} was required while 10^{-6} as energy and radiation equations residuals.

6 Results and Discussions

This section presents the results of the numerical simulation under different geometrical and operational conditions. Several key parameters affecting flow behavior and heat transfer performance have been evaluated and examined. Four main parameters used in the parametric study, namely, the effect of collector radius, coil pitch, tube diameter and water inlet temperature are investigated in details.

6.1 Validation

The validation is performed by comparing the CFD results with experimental data from the Spanish prototype. The previous *measurements are taken* on September 1982. The calculation results and the experimental data are illustrated in Table 2. We conclude that our calculated values have shown close agreement with the experimental data from the Spanish plant.

The air exit velocity were calculated and presented in Fig. 6 for several solar irradiation values varying between 100 and 1000 W/m^2. The examination of Fig. 6

Table 2 Comparison between numerical and experimental results

Solar radiation (W/m^2)		Numerical data	Experimental data
800	ΔT (°C)	17.783	17
1000	ΔT (°C)	19.277	20

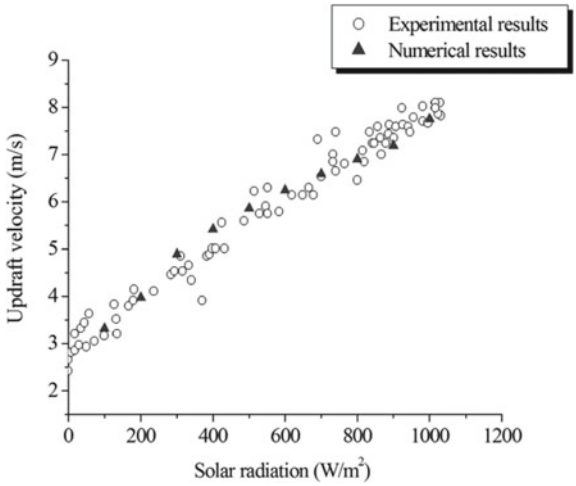


Fig. 6 Effect of solar radiation on outlet air velocity

shows that the numerical data are in a satisfactory agreement with the experimental measurements [27]. The small difference between the calculated and experimentally obtained data at the outlet of the collector corroborates the validity of the results.

6.2 The Impact of Collector Radius on the Plant Performance

The collector radius is an *important* factor that greatly influences the solar plant performance. To investigate the impact of the collector radius on the hybrid plant performance, different values of radius were considered (40, 60, 80, 100, 120 and 140 m) and main dimensions of the prototype are kept constant apart from the radius of the collector. Figure 7 displays the distribution of outlet air velocity against the collector radius. A constant water temperature equal to 70 °C was considered. In this section, the study is carried out for a given total water flow rate (85.6 kg/s).

These results highlight a marked influence of the collector radius. After various numerical tests, it was found that *during* the *night*, the collector exit air velocity drop significantly as the radius increases.

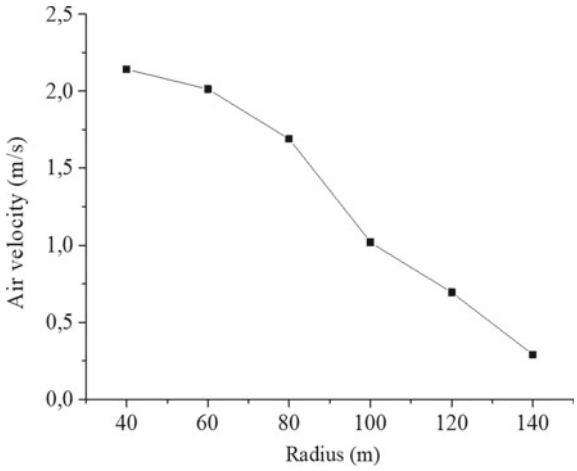


Fig. 7 Air velocity variation against collector radius

6.3 Investigation of Coil Pitch

Spacing plays an essential role in the design of spiral heat exchangers. The analyses of the heat flow rate for different coil pitch are important. Tests have been made to study the effect of the variation of the spacing between consecutive turns on the amount of heat exchanged between the geothermal water and the air circulating underneath the collector. Figure 8 shows the change in heat flow rate for different spacing between the spiral turns.

We chose to vary the pitch between 0.2 and 1 m. Note that increasing the pitch leads to a reduction of the number of turns which varies from 92 to 19. For a pitch ranging from 0.2 to 0.4 m, it can be seen that rising the pitch coil induces the augmentation of heat transfer rate. Apparently, the in-plane spiral heat exchanger has a smaller mean

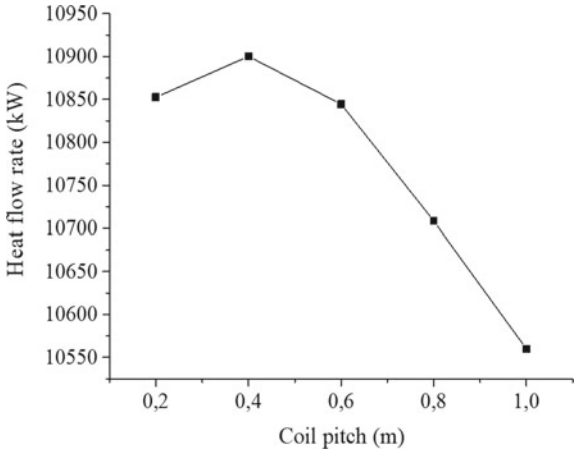


Fig. 8 Variation of heat flow rate against the coil pitch

bend radius and a shorter tube length for a bigger coil pitch. Small curve radius is advantageous to the heat exchange process. Therefore, the heat flux rises with the increment of the coil pitch.

Beyond these values, a dramatic change in the heat flow rate is observed. The increase in the coil pitch will result in a lower amount of heat. The calculated data show clearly the influence of the pitch variation on the performance of the spiral heat exchanger. So maintaining a constant pitch is one of the key parameters that guarantee a good performance. In addition, it is necessary to maintain a sufficient spacing so that air can circulate easily between the tubes.

6.4 Investigation of Tube Diameter

The Fig. 9 shows the variation of the geothermal heat flux against the tube diameter. Nine cases were considered to evaluate the impact of the tube diameter on the heat flux by varying the diameter from 0.2 to 1 m. In addition, the water flow rate, the coil pitch and the collector size are assumed to be constant.

It can be seen that the increase in tube diameter can significantly affect the rate of heat transferred to the internal air. For a diameter value less than 0.4 m, the heating capacity increases. It should be noted that beyond this value, the heat flow recovered from the hot water decreases. This may be due to the reduction of the effect of centrifugal force (as a result of the pipe curve). This leads ultimately to the reduction of the secondary flow (Dean Effect) for largest coil diameter. The intensity of the secondary flow created in the pipe is dependent on coil geometry and pipe diameter [30–33].

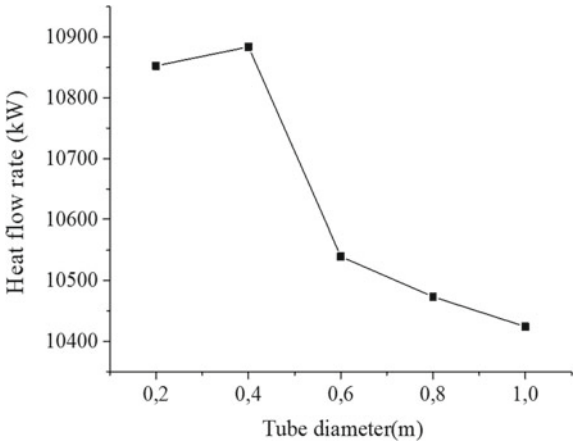


Fig. 9 Heat flow rate variation versus tube diameter

6.5 Geothermal Water Flow Investigation

With the aim of examining the impact of the indoor water flow rate on the thermal performance of the coiled tube heat exchanger, a study is performed assuming a constant geothermal water temperature (70 °C) and for a constant collector radius (120 m).

The constructional parameters of the heat exchanger are held fixed values (R_i = 5 m, R_0 = 120 m, b = 0.2 m, d = 0.2 m). Figure 10 displays the variation of the amount of heat in terms of the water flow along the spiral coil pipes. The flow rates of water considered are in the 85.6- 214 kg/s range.

As expected, we can notice in Fig. 10 that it is advantageous to increase the total flow of geothermal water. All calculations were made with the same geometrical parameters of Manzanares prototype. A total of 7 cases were considered in this study by varying the flow rate with an increment of 21.4 kg/s. The bigger the water flow rate, the greater the amount of heat recovered by the air passing through the collector. The Fig. 10 also confirms that it is preferable to increase the water inlet flow rate to obtain higher air velocities.

6.6 Effect of Water Inlet Temperature

To evaluate the impact of varying inlet water temperature on heat flow rate, the numerical tests are performed for a water temperature in the range of 55–75 °C.

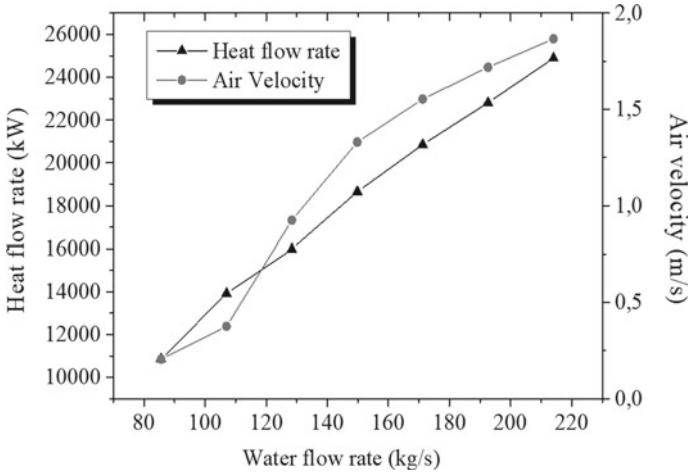


Fig. 10 Effect of water flow rate

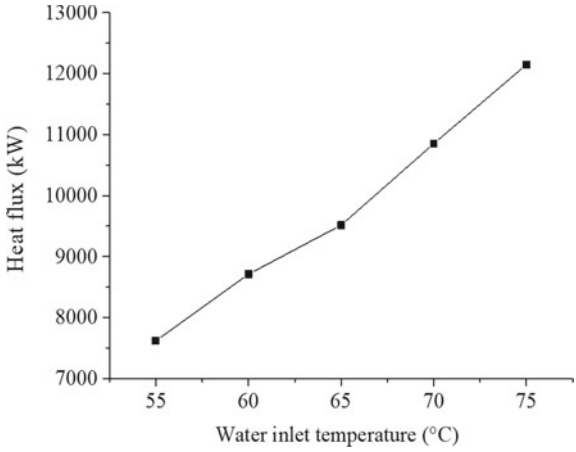


Fig. 11 Heat Flow rate as a function of water inlet temperature

The value of geothermal water flow is kept to be constant (85.6 kg/s). Except the water inlet temperature, all parameters were fixed as constant values. Computational simulations were conducted for the same geometric model of the Spanish prototype. Figure 11 shows a net increase of the geothermal power and a significant effect of inlet water temperature. In addition, this figure indicates that the difference between the quantities of heat supplied by the heat exchanger can reach 12.38% on average. It can be concluded that the more the operating temperatures and input flow rate are well chosen, the more efficiently the heat exchanger will operate. The Fig. 11 shows that for a stated water mass flow rate, the overall rate of heat transfer increases with the increase of the inlet water temperature.

6.7 Variation of Internal Air and Geothermal Water Temperatures

It seems important to investigate the impact of the heater size on the temperatures of the two working fluids. To investigate the effect of collector size, we have varied the radius from 40 to 140 m. Figure 12 shows the variation of the water and air temperatures at the exit of each turn for a constant geothermal water flow. The number of turns corresponding to this case is 92. The water enters at the inlet of the tube located in the center of the collector with a hot temperature (70 °C) and comes out at the end of the spiral with a colder temperature (39 °C). According to Fig. 12, the water outlet temperature drops quickly initially, but then slows down. Thus, the air temperature near the tube is cooler on the outside and warmer inwards (radially), as it approaches the water exit temperature. This temperature increases less rapidly. As it can be seen from Fig. 12, for a number of turns equal to 40, the temperatures of the working fluids are almost stable.

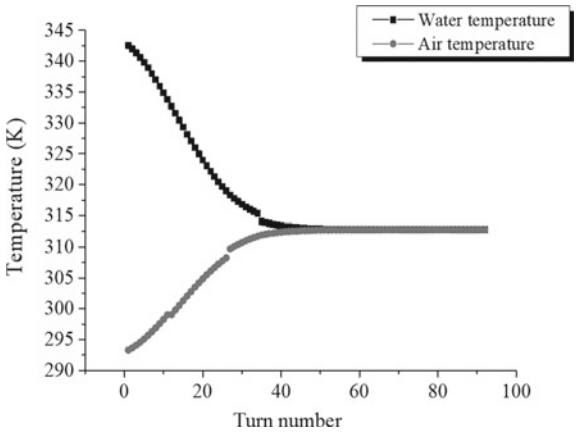


Fig. 12 Variation of air and geothermal water temperatures against turn number

The obtained results can be related to the development of secondary flow as a result of centrifugal force. This force has long been identified as the most important feature of coiled tubes and it is created due to the pipe curvature effect [34]. Heat Transfer associated with spiral ducts is more difficult than that occurring in straight pipes. This difference depends on the occurrence of secondary flows caused by centrifugal forces that considerably influence the heat and mass transfer. It is also interesting to note that stronger secondary flow intensity results in greater rates of heat transfer. But the increase in heat transfer is more marked in the laminar regime. Since the regime flow considered in this study is turbulent, the increase in heat transfer is less important than in the case of laminar flow [35]. Turbulence effects dominate over secondary high Reynolds flow, making the heat transfer coefficient equivalent to straight tubes [36]. The secondary flow is found to have a considerable influence on the evolution of temperature profiles and flow rate in the spiral duct [37].

6.8 Collector Heat Losses

To examine the impact of radius variation on heat dissipation rate in the collector, the radius was varying from 40 to 140 m. It should be noted that for a radius less than 80 m, any augmentation of the radius induces the increase of the collector energy losses (Fig. 13). In the case of a radius equal to 100 m, the lower energy losses are mainly due to the decrease of convective losses in the outer surface of the collector cover. Exceeding 120 m, the temperature of indoor air is lower than the outside air. A temperature inversion appears and convective exchange from the cover to the outside air in the opposite direction is observed. This phenomenon occurs when thermal losses through the cover are very high and the cover temperature may drop several degrees below the outside air temperature.

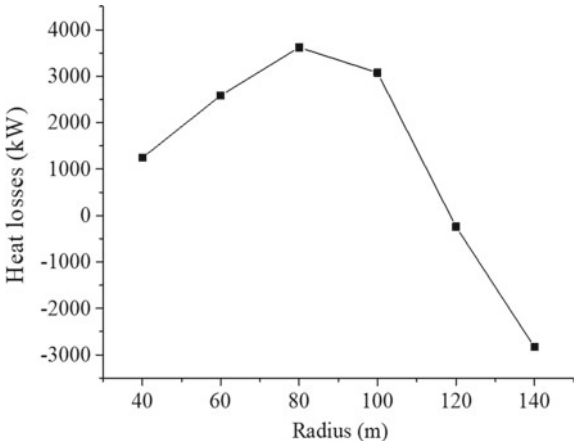


Fig. 13 Collector heat losses as a function of radius

7 Conclusion

A computational study was conducted to examine the night operation of a solar chimney combined to a spiral heat exchanger. The use of geothermal water is an advantageous solution due to its ability to ensure a continuous operation of the plant by fulfilling the nocturnal power demand. This work has identified the parameters affecting the system performance overnight. The Analysis of the combined system behavior leads to the following findings:

1. The air exit velocity at night depends strongly on the size of the collector.
2. The use of geothermal energy may improve the performance of the solar chimney overnight.
3. The geometrical dimensions of the heating device (coil pitch and tube diameter) have a prominent impact on the heat transfer process.
4. The rate of heat transferred *to the interior air* is greatly determined by the exchanger operational parameters (inner temperature and water flow rate). The higher these parameters, the greater the collector outlet air velocity.
5. *Canopy thermal losses* increase notably once the collector radius is beyond 80 m. for a radius exceeding 120 m, a temperature inversion phenomenon appears and convective exchange from the cover to the outside air in the opposite direction is observed.

References

1. Schlaich J (1995) The solar chimney: electricity from the, sun edn. Axel Menges, Felbach
2. Goswami DY, Zhao Y (2009) Proceedings of ISES world congress 2007, vol 1–vol 5. Solar energy and human settlement. Springer Science & Business Media
3. Chikere AO, Al-Kayiem HH, Karim ZAA (2011) Review on the enhancement techniques and introduction of an alternate enhancement technique of solar chimney power plant. J Appl Sci 11(11):1877–1884
4. Fasel HF, Meng F, Shams E, Gross A (2013) CFD analysis for solar chimney power plants. Solar Energy Part A 98:12–22
5. Kreetz H (1997) Theoretische Untersuchungen und Auslegung eines temporärenWasserspeichers für das Aufwindkraftwerk, dissertation, Technical University Berlin
6. Fathi N, McDaniel P, Aleyasin SS, Robinson M, Vorobieff P, Rodriguez S, de Oliveira C (2018) Efficiency enhancement of solar chimney power plant by use of waste heat from nuclear power plant. J Clean Prod 180:407–416
7. Kasaeian AB, Molana S, Rahmani K et al (2017) A review on solar chimney systems. Renew Sustain Energy Rev 67:954–987
8. Al-Kayiem HH et al (2019) Performance evaluation of hybrid solar chimney for uninterrupted power generation. Energy 166:490–505
9. Hurtado FJ, Kaiser AS, Zamora B (2012) Evaluation of the influence of soil thermal inertia on the performance of a solar chimney power plant. Energy 47(1):213–224
10. Zheng Y, Ming TZ, Zhou Z, Yu XF, Wang HY, Pan Y, Liu W (2010) Unsteady numerical simulation of solar chimney power plant system with energy storage layer. J Energy Inst 83(2):86–92
11. Ming TZ, Zheng Y, Liu W, Huang XM (2009) Unsteady numerical conjugate simulation of the solar chimney power generation systems. J Eng Thermophys 30:4
12. Fanlong M, Tingzhen M, Yuan P (2011) A method of decreasing power output fluctuation of solar chimney power generating systems. In: 2011 3rd international conference on measuring technology and mechatronics automation (ICMTMA), vol 1, pp 114–118
13. Yan Zhou XHL (2011) Unsteady conjugate numerical simulation of the solar chimney power plant system with vertical heat collector. Mater Sci Forum 704–705:535–540
14. Robert AL, Craig RB, Eckhard AG et al (2012) Alternative heat rejection methods for power plants. Appl Energy 92:17–25
15. Barbier E (2002) Geothermal energy technology and current status: an overview. Renew Sustain Energy Rev 6(1–2):3–65
16. Al-Kayiem HH, Brebbia CA, Zubir SS (2014) Energy and sustainability V. WIT Press, UK
17. Zubir SS, Brebbia A (2013) The sustainable city VIII (2 volume set): urban regeneration and sustainability. WIT Press, UK
18. Aurybi MA, Al-Kayiem HH, Gilani SIU, Ismaeel AA (2017) Numerical analysis of solar updraft power plant integrated with external heat source. In: UTP-UMP symposium on energy systems 2017 (SES 2017), vol 131, (MATEC Web Conf.), pp 7
19. Cao F, Li H, Ma Q, Zhao L (2014) Design and simulation of a geothermal–solar combined chimney power plant. Energy Convers Manag 84:186–195
20. Dhahri A, Omri A, Orfi J (2018) Theoretical analysis of the performance of a solar chimney coupled with a geothermal heat exchanger. In: Driss Z, Necib B, Zhang HC (eds) CFD techniques and energy applications. Springer, Cham
21. Elmagid WMA, Kepple I (2017) Axial flow turbine for solar chimney. Hung Agric Eng J 32:29–37
22. Weinrebe G, Bergermann R, Schlaich J (2012) solar updraft towers. In: Meyers RA (ed) Encyclopedia of sustainability science and technology. Springer, New York, NY
23. Verlodt H (1983) Amélioration du bilan thermique sous abri-serre. Tropicultura 1(2):59–69
24. Ho JC, Wijeysundera NE (1996) Study of a compact spiral-coil cooling and dehu midifying heat exchanger. Appl Therm Eng 16(10):777–790

25. Nguyen DK, San JY (2016) Heat transfer and exergy analysis of a spiral heat exchanger. Heat Trans Eng 37(12):1013–1026
26. Wijeysundera NE, Ho JC, Rajasekar S (1996) The effectiveness of a spiral coil heat exchanger. Int Commun Heat Mass Transfer 23(5):623–631
27. Dhahri A, Omri A, Orfi J (2014) Numerical study of a solar chimney power plant. Res J Appl Sci Eng Technol 8:1953–1965
28. Cuce E, Cuce PM (2019) Performance assessment of solar chimneys: part i–impact of chimney height on power output. Energy Res J 10:11–19
29. Bernardes MAdS (2004) Technische, ökonomische und ökologische Analyse von Aufwind-kraftwerken, Ph.D. Thesis. Universität Stuttgart
30. Deshpande M et al (2012) Study of hydrodynamics of horizontal spiral coil tube. Int J Adv Eng Res Stud 1(3):112–114
31. Purandare PS, Lel MM, Gupta R (2012) Parametric analysis of helical coil heat exchanger. Int J of Eng Res Technol 1(8)
32. Mukesh Kumar PC, Chandrasekar M (2019) CFD analysis on heat and flow characteristics of double helically coiled tube heat exchanger handling MWCNT/water nanofluids. Heliyon 5(7):e02030
33. Chaware P, Sewatkar CM (2018) Effects of tangential and radial velocity on fluid flow and heat transfer for flow through a pipe with twisted tape insert—laminar flow. Sādhanā 43(9):1
34. Yoo G, Choi H, Dong W (2012) Fluid flow and heat transfer characteristics of spiral coiled tube: effects of reynolds number and curvature ratio. J Cent South Univ Technol 19:471–476
35. Bergman TL, Incropera FP (2011) Introduction to heat transfer. John Wiley & Sons
36. Pablo Coronel KPS (2008) Heat transfer coefficient in helical heat exchangers under turbulent flow conditions. Int Jof Food Eng 4(1)
37. Naphon P, Suwagrai J (2007) Effect of curvature ratios on the heat transfer and flow developments in the horizontal spirally coiled tubes. Int J Heat Mass Transfer 50:444–451

Incorporating of IPFC in Multi-machine Power System Phillips-Heffron Model

Nemat Talebi and Abas Moghadasi

Abstract The development of power networks has led to low-frequency oscillations in the power system. The relatively small and sudden disturbances in the network cause such oscillations in the system. In the normal case, the oscillations damp rapidly and the amplitude of oscillations does not exceed a certain value. But depending on the operating point conditions and the values of the system parameters, these oscillations may continue for a long time and at worst increase their amplitude. These oscillations can affect the power transmission capability and stability of the power system. PSS (Power System Stabilizer) and FACTS (Flexible AC Transmission Systems) devices are commonly used for damping low-frequency oscillations. The Interline Power Flow Controller (IPFC) among the FACTS device is suitable for the power flow control and stability of power systems. This chapter establishes an approach to derive the dynamical model of a multi-machine power system with embedded IPFC devices. Derivations about the stability analysis and the incorporating of IPFC to enhance the damping of low-frequency oscillations in a multi-machine power system based on Modified Phillips-Heffron modeling has been done. To demonstrate the application and efficiency of the developed models, a case study on a three-machine test power system with adding IPFC has been presented. Numerical results with MATLAB for dynamical simulations show the significant effects of IPFC on damping the low-frequency oscillations and especially validates the modeling procedure.

Keywords IPFC · Phillips-Heffron model · FACTS Devices · Stability of Multi-machine power systems

N. Talebi (✉) · A. Moghadasi
Department of Electrical Engineering, Tehran South Branch, Islamic Azad University, Tehran, Iran
e-mail: nemat-talebi@azad.ac.ir

A. Moghadasi
e-mail: abasmoghadasi@gmail.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_15

Nomenclature

δ	Is the rotor angle in electric radians,
ω	Is generator rotor angle velocity (rad/s),
V_{Bt}	Is the terminal voltage of IPFC (p.u),
E'_q	Is internal quadrature-axis Voltage generator (p.u),
E_{fd}	Is the field voltage (p.u),
V_{DC}	Is the DC voltage (p.u),
$M = 2H$,	H Is the inertia constant (s),
P_m	Is the mechanical input power (p.u),
P_e	Is the electrical output power (p.u),
D	Is the damping torque coefficient,
x_d, x'_d	Are the d-axis reactance and transient reactance (p.u),
T'_{do}	Is the open circuit d axis time constant (s),
I_d, I_q	Are the direct and quadrature axis components of stator current in p.u,
V_{td}, V_{tq},	Are the d and q axis components of terminal voltage in p.u.,
T_a	Are time constant and gain of AVR,
m_{B1} and m_{B2}	Are The amplitude of the modulation ratios of VSC 1 and 2,
$\delta_{B1}\delta_{B2}$	Are the phase angles of VSC 1 and 2.

1 Introduction

The electrical power system is an interconnected large scale system with non-linearity in nature. Usually, the damping of electromechanical oscillations is weak on the large scale power systems. Due to insufficient damping torque of the generators, the low-frequency Electro-mechanical oscillations happen and they have two type modes: local mode (1–2 Hz) and the inter-area mode (more than 0.1 Hz) [1, 2]. If the damping of these modes becomes too small, then it can impose severe restrictions on the operation of power systems. Thus, identifying the nature of these modes to determine the stability boundaries and in many cases using additional control signals to prevent instability is necessary [2, 3].

For years Power System Stabilizers (PSS) have been used to enhance the dynamical stability of the power systems. So, PSS by adding an additional signal to the generator excitation control loop improves power or speed oscillation damping [4, 5]. The PSS usually has been designed for damping increment of the local modes and in large power systems, but especially in heavy load conditions with long transmission lines cannot maintain enough damping for inter-area modes [2, 6]. In recent years, The Flexible AC Transmission System (FACTS) devices have been introduced to increase control options in power systems and enhance the capability and flexibility of power transmission systems [7, 8]

In literature dedicated to a power system with embedded FACTS devices, the main attention reveals to control techniques. Most of the efforts have been done on the linear model and small-signal stability using Phillips- Heffron model, especially with a single machine connected to the infinite bus [9, 10]. Thus Phillips-Heffron model with implementing various FACTS devices (TCSC, SSSC, and UPFC) in the power system as a single machine connected to the infinite bus and a variety of classical and fuzzy controllers have been proposed to improve stability and increasing damping [11, 12].

Furthermore, coordination between PSS and various FACTS devices controls has been addressed [13, 14]. Linear multi-machine Philips-Heffron model with the application of SSSC and UPFC through appropriate controllers have been implemented to enhance stability [15].

The Interline Power Flow Controller (IPFC) is a multi-terminal FACTS device which aimed to control the power transmission over lines [16]. IPFC has been made from a few numbers of Voltage Source Converters (VSC) which, are installed in series with transmission lines [17, 18]. Three types of static modeling of the IPFC are; Voltage source model [19], Current source model [20] and Power injection model [21, 22]. Like other FACTS devices, IPFC can also be used as a dynamical controller to increase system stability and oscillations damping [23, 24].

This chapter is organized as follows: In Sect. 2 nonlinear modeling of multi-machine with installed IPFC modeling has been carried out. Section 3 to combine the modeling of the IPFC and generators and its effects on both local and inter-area oscillations, the stability of the power system within the linear multi-machine Phillips- Heffron model have been investigated. In Sect. 4 the controllers for IPFC are designed separately by using the generator outputs. In Sect. 5 evaluation of the simulation results indicates efficient modeling and proper controller assignment which improves stability and damps oscillations perfectly. Finally, conclusion and advantages of the proposed model are remarked in Sect. 6.

2 Nonlinear Model of Multi-machine Power System with Embedded IPFC

A typical installation of IPFC in a multi-machine system is shown in Fig. 1 . It is assumed that an IPFC is installed on branches 1, 2 and 1, 3.

The IPFC consists of two boosting transformers (BT), two three-phase GTO based voltage source converters (VSCs) and a DC link capacitor. In Fig. 1, mB1, mB2, δB1, and δB2 are the amplitude of the modulation ratios and phase angles of the control signal of each VSC respectively, which are the input control signals to the IPFC. If the resistance and transients of transformers of the IPFC are not considered here, the dynamic model of the IPFC is given as:

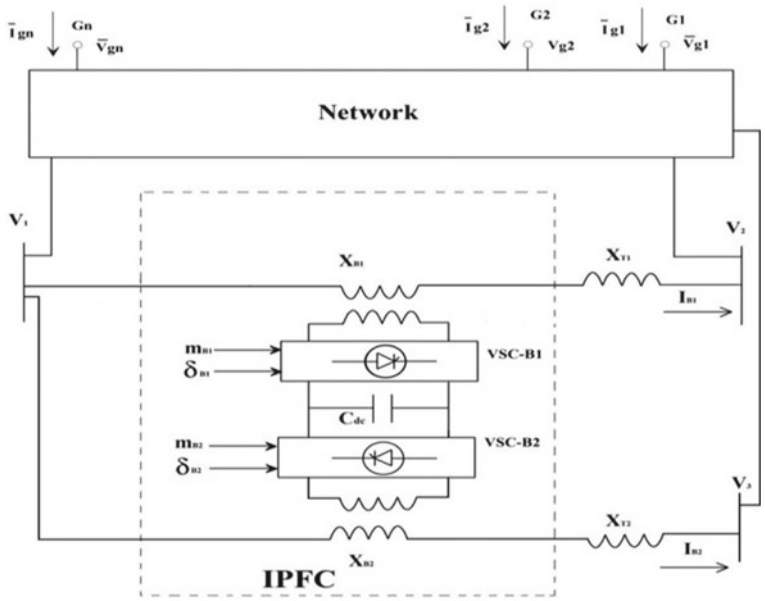


Fig. 1. n-machine power system within IPFC installation

$$\frac{dv_{DC}}{dt} = \frac{3m_{B1}}{4C_{DC}}[\cos\delta_{B1}\ \sin\delta_{B1}]\begin{bmatrix} i_{Bd1} \\ i_{Bq1} \end{bmatrix} + \frac{3m_{B2}}{4C_{DC}}[\cos\delta_{B2}\ \sin\delta_{B2}]\begin{bmatrix} i_{Bd2} \\ i_{Bq2} \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} v_{Btd1} \\ v_{Btq1} \end{bmatrix} = \begin{bmatrix} 0 & -x_{B1} \\ x_{B1} & 0 \end{bmatrix}\begin{bmatrix} i_{Bd1} \\ i_{Bq1} \end{bmatrix} + \begin{bmatrix} \frac{m_{B1}\cos\delta_{B1}v_{DC}}{2} \\ \frac{m_{B2}\sin\delta_{B2}v_{DC}}{2} \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} v_{Btd2} \\ v_{Btq2} \end{bmatrix} = \begin{bmatrix} 0 & -x_{B2} \\ x_{B2} & 0 \end{bmatrix}\begin{bmatrix} i_{Bd2} \\ i_{Bq2} \end{bmatrix} + \begin{bmatrix} \frac{m_{B2}\cos\delta_{B2}v_{DC}}{2} \\ \frac{m_{B2}\sin\delta_{B2}v_{DC}}{2} \end{bmatrix} \tag{3}$$

Equations (2, 3) in the compact form can be written as

$$\overline{V}Bt1 = jx\overline{I}B1 + \overline{V}B1 \tag{4}$$

$$\overline{V}_{Bt2} = jx_{B2}\overline{I}_{B2} + \overline{V}_{B2} \tag{5}$$

where

$$\overline{I}_{B1} = i_{Bd1} + ji_{Bq1} \quad \overline{I}_{B2} = i_{Bd2} + ji_{Bq2} \tag{6}$$

$$\begin{aligned}\overline{V}_{B1} &= \tfrac{m_{B1}v_{DC}}{2}(\cos\delta_{B1} + j\sin\delta_{B1}) = \tfrac{m_{B1}v_{DC}}{2}e^{j\delta_{B1}} \\ \overline{V}_{B2} &= \tfrac{m_{B2}v_{DC}}{2}(\cos\delta_{B2} + j\sin\delta_{B2}) = \tfrac{m_{B2}v_{DC}}{2}e^{j\delta_{B2}}\end{aligned} \tag{7}$$

$$\overline{I_{B1}} = \frac{1}{j(X_{T1} + X_{B1})}(V_1 - V_2 - V_{B1}) \tag{8}$$

$$\overline{I_{B2}} = \frac{1}{j(X_{T2} + X_{B2})}(V_1 - V_3 - V_{B2}) \tag{9}$$

The network admittance Yt is formed before the IPFC has been installed, where only n generator nodes plus 3 IPFC terminal nodes (Assigned node 1, 2 and 3 for IPFC) are kept and the other nodes are reduced. The circuit KVL and KCL equation are:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \overline{I_g} \end{bmatrix} = \begin{bmatrix} \overline{Y_{11}} & \overline{Y_{12}} & \overline{Y_{13}} & \overline{Y_{14}} \\ \overline{Y_{21}} & \overline{Y_{22}} & \overline{Y_{23}} & \overline{Y_{24}} \\ \overline{Y_{31}} & \overline{Y_{32}} & \overline{Y_{33}} & \overline{Y_{34}} \\ \overline{Y_{41}} & \overline{Y_{42}} & \overline{Y_{43}} & \overline{Y_{44}} \end{bmatrix} \begin{bmatrix} \overline{V_1} \\ \overline{V_2} \\ \overline{V_3} \\ \overline{V_g} \end{bmatrix} = \overline{Y_t} \begin{bmatrix} \overline{V_1} \\ \overline{V_2} \\ \overline{V_3} \\ \overline{V_g} \end{bmatrix} \tag{10}$$

where

$$\overline{I}_g = \begin{bmatrix} \overline{I}_{g1} & \overline{I}_{g2} & \dots & \overline{I}_{gn} \end{bmatrix}^T \quad \overline{V}_g = \begin{bmatrix} \overline{V}_{g1} & \overline{V}_{g2} & \dots & V_{gn} \end{bmatrix}^T \tag{11}$$

With installing of the IPFC on node 1, 2 and 3 the equations of the network and IPFC are combined together:

$$\begin{bmatrix} -(\overline{I_{B1IPFC}} + \overline{I_{B2IPFC}}) \\ \overline{I_{B1IPFC}} \\ \overline{I_{B2IPFC}} \\ \overline{I_g} \end{bmatrix} = \begin{bmatrix} \overline{Y_{11}} & \overline{Y_{12}} & \overline{Y_{13}} & \overline{Y_{14}} \\ \overline{Y_{21}} & \overline{Y_{22}} & \overline{Y_{23}} & \overline{Y_{24}} \\ \overline{Y_{31}} & \overline{Y_{32}} & \overline{Y_{33}} & \overline{Y_{34}} \\ \overline{Y_{41}} & \overline{Y_{42}} & \overline{Y_{43}} & \overline{Y_{44}} \end{bmatrix} \begin{bmatrix} \overline{V_1} \\ \overline{V_2} \\ \overline{V_3} \\ \overline{V_g} \end{bmatrix} = \overline{Y_t} \begin{bmatrix} \overline{V_1} \\ \overline{V_2} \\ \overline{V_3} \\ \overline{V_g} \end{bmatrix} \tag{12}$$

$$\overline{Y_{11}} = \overline{Y_{11}} + \frac{1}{j(X_{T1} + X_{B1}) + j(X_{T2} + X_{B2})} \tag{13}$$

$$\overline{Y_{12}} = \frac{-1}{j(X_{T1} + X_{B1})} \tag{14}$$

$$\overline{Y_{21}} = \frac{-1}{j(X_{T1} + X_{B1})} \tag{15}$$

$$\overline{Y_{22}} = \overline{Y_{22}} + \frac{1}{j(X_{T1} + X_{B1})} \tag{16}$$

$$\overline{Y_{23}} = \overline{Y_{23}} \text{ and } \overline{Y_{32}} = \overline{Y_{32}} \tag{17}$$

$$\overline{Y_{31}} = \frac{-1}{j(X_{T2} + X_{B2})} \tag{18}$$

$$\overline{Y_{33}} = \overline{Y_{33}} + \frac{1}{j(X_{T2} + X_{B2})} \tag{19}$$

$$\overline{I_{B1IPFC}} + \overline{I_{B2IPFC}} = \frac{1}{j(X_{T1} + X_{B1})} V_{B1} + \frac{1}{j(X_{T2} + X_{B2})} V_{B2} \tag{20}$$

$$\overline{I_{B1IPFC}} = \frac{1}{j(X_{T1} + X_{B1})} V_{B1} \tag{21}$$

$$\overline{I_{B2IPFC}} = \frac{1}{j(X_{T2} + X_{B2})} V_{B2} \tag{22}$$

Then by substituting (20)–(22) and reduction of the node 1, 2 and 3 the result in the compact form will be as:

$$\overline{I}_g = \overline{CV_g} + \overline{F_{B1}V_{B1}} + \overline{F_{B2}V_{B2}} \tag{23}$$

where

$$\begin{aligned} \overline{I_g} = &\left[\overline{Y_{44}} - \begin{bmatrix} \overline{Y_{41}} & \overline{Y_{42}} & \overline{Y_{43}} \end{bmatrix} \begin{bmatrix} \overline{Y_{11}} & \overline{Y_{12}} & \overline{Y_{13}} \\ \overline{Y_{21}} & \overline{Y_{22}} & \overline{Y_{23}} \\ \overline{Y_{31}} & \overline{Y_{32}} & \overline{Y_{33}} \end{bmatrix}^{-1} \begin{bmatrix} \overline{Y_{14}} \\ \overline{Y_{24}} \\ \overline{Y_{34}} \end{bmatrix} \right] \overline{V_g} \\ &+ \begin{bmatrix} \overline{Y_{41}} & \overline{Y_{42}} & \overline{Y_{43}} \end{bmatrix} \begin{bmatrix} \overline{Y_{11}} & \overline{Y_{12}} & \overline{Y_{13}} \\ \overline{Y_{21}} & \overline{Y_{22}} & \overline{Y_{23}} \\ \overline{Y_{31}} & \overline{Y_{32}} & \overline{Y_{33}} \end{bmatrix}^{-1} \begin{bmatrix} \overline{I_{B1IPFC}} + \overline{I_{B2IPFC}} \\ \overline{I_{B1IPFC}} \\ \overline{I_{B2IPFC}} \end{bmatrix} \end{aligned} \tag{24}$$

In (11), $\overline{I_g}$ and $\overline{V_g}$ are expressed on x–y co-ordinate. That is,

$$\begin{aligned} &\overline{I}_{gi} = I_{gxi} + jI_{gyi} \quad \overline{V}_{gi} = V_{gxi} + jV_{gyi} \quad i = 1, 2, \ldots, n \\ &\overline{V}_E = V_{Ex} + jV_{Ey} \quad \overline{V}_B = V_{Bx} + jV_{By} \end{aligned} \tag{25}$$

So with the arrangement of (10), it becomes

$$I_G = Y_G V_G + Y_{B1} \begin{bmatrix} V_{B1x} \\ V_{B1y} \end{bmatrix} + Y_{B2} \begin{bmatrix} V_{B2x} \\ V_{B2y} \end{bmatrix} \tag{26}$$

where

$$\begin{aligned} I_G &= \begin{bmatrix} I_{gx1} & I_{gy1} & I_{gx2} & I_{gy2} & \ldots & I_{gxn} & I_{gyn} \end{bmatrix}^T \\ V_G &= \begin{bmatrix} V_{gx1} & V_{gy1} & V_{gx2} & V_{gy2} & \ldots & V_{gxn} & V_{gyn} \end{bmatrix}^T \end{aligned} \tag{27}$$

Transforming Eq. (27) into the *d-q* co-ordinates yields,

$$I_{Gdq} = T I_G \ , \ V_{Gdq} = T V_G \tag{28}$$

where

$$\begin{aligned} I_{Gdq} &= \left[I_{d1} \ I_{q1} \ I_{d2} \ I_{q2} \ \ldots \ I_{dn} \ I_{qn} \right]^T \\ V_G &= \left[V_{d1} \ V_{q1} \ V_{d2} \ V_{q2} \ \ldots \ V_{dn} \ V_{qn} \right]^T \\ T &= diag(T_i) \quad T_i = \begin{bmatrix} \cos\delta_i & \sin\delta_i \\ \sin\delta_i & -\cos\delta_i \end{bmatrix} \end{aligned} \tag{29}$$

With substituting Eq. (29) into (26) The relationship of the voltages and the currents of generator nodes in d-q co-ordinates are obtained as:

$$I_{Gdq} = M_G V_{Gdq} + M_E \begin{bmatrix} V_{Ex} \\ V_{Ey} \end{bmatrix} + M_B \begin{bmatrix} V_{Bx} \\ V_{By} \end{bmatrix} \tag{30}$$

where

$$M_G = T Y_G T^{-1}, M_E = T Y_E, M_B = T Y_B \tag{31}$$

In an n-machine power system, the full dynamics of a synchronous generator is described by the following equations in d-q co-ordinate [25].

$$\begin{aligned} \begin{bmatrix} \dot{E}'_{qi} \\ \dot{E}''_{qi} \\ \dot{E}''_{di} \end{bmatrix} &= \begin{bmatrix} -\frac{1}{T'_{doi}} & 0 & 0 \\ \frac{T'_{doi}-T''_{doi}}{T'_{doi}T''_{doi}} & -\frac{1}{T''_{doi}} & 0 \\ 0 & 0 & -\frac{1}{T''_{qoi}} \end{bmatrix} \begin{bmatrix} E'_{qi} \\ E''_{qi} \\ E''_{di} \end{bmatrix} + \\ &\begin{bmatrix} -\frac{1}{T'_{doi}} & 0 & -\frac{x_{di}-x'_{di}}{T'_{doj}} \\ -\frac{1}{T'_{doi}} & 0 & -\frac{T''_{doi}(x_{di}-x'_{di})-T'_{doi}(x_{qi}-x''_{qi})}{T'_{doi}T''_{doi}} \\ 0 & \frac{x_{qi}-x''_{qi}}{T''_{qoi}} & 0 \end{bmatrix} \\ &\times \begin{bmatrix} E_{qei} \\ I_{qi} \\ I_{di} \end{bmatrix} \end{aligned} \tag{32}$$

$$\begin{bmatrix} \dot{\delta}_i \\ \dot{\omega}_i \end{bmatrix} = \begin{bmatrix} 0 & \omega_0 \\ 0 & -\frac{D_i}{2H_i} \end{bmatrix} \begin{bmatrix} \delta_i \\ \omega_i \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2H_i} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2H_i} \end{bmatrix} (T_{Mi} - T_{ei}) \tag{33}$$

$$Tei = \psi_{di} I_{qi} - \psi_{qi} I_{di} \tag{34}$$

The dynamic model of the above generator with adding the governor and AVR control loops is completed as follows [25]:

$$T_{mi} = G_{ovi}(s)\omega_i,\ E_{qei} = A_{vri}(s)(V_{oi} - \sqrt{V_{di}^2 - V_{qi}^2}) \tag{35}$$

where Govi(s) and Avri(s) are the transfer function of the governor and AVR control loops, respectively.

3 Linearized Model

The linear model of power systems has been successfully used for dynamic stability analysis and design of low-frequency oscillation controllers such as PSS-based stabilizers and FACTS. The linear model of the n-machine power system including IPFC is obtained by linearization (1) and (30)–(35) (nonlinear power system model) around an operating point. This will produce a general linear model of the power system [26].

$$\begin{aligned} &\dot{\delta} = \omega_0\omega \quad , \quad \dot{\omega} = M^{-1}(T_M - T_E - D_\omega), \\ &\dot{E}'_q = T_{DO}^{\prime -1}[(X_D - X'_D)I_D - E'_q + E_{QE}] \end{aligned} \tag{36}$$

where

$$T_E = I_Q V_{TQ} + I_D V_{TD} \quad , \quad V_{TD} = X_Q I_Q \ V_{TQ} = \dot{E}'_q - X'_D I_D \tag{37}$$

$$\begin{aligned} &\delta = \left[\delta_1\ \delta_2 \ldots \delta_n\right]^T \ , \ \ \omega = \left[\omega_1\ \omega_2 \ldots \omega_n\right]^T \ , \\ &E'_q = \left[E'_{q1}\ E'_{q2} \ldots E'_{qn}\right]^T \\ &I_D = \left[I_{d1}\ I_{d2} \ldots I_{dn}\right]^T \ , \ \ I_Q = \left[I_{q1}\ I_{q2} \ldots I_{qn}\right]^T \\ &V_{TD} = \left[V_{d1}\ V_{d2} \ldots V_{dn}\right]^T \ , \ \ V_{TQ} = \left[V_{q1}\ V_{q2} \ldots V_{qn}\right]^T \\ &M = diag(2H_i) \ \ D = diag(D_i) \ \ T'_{DO} = diag(T'_{do}) \\ &X_D = diag(x_{di}) \ \ X'_D = diag(x'_d) \end{aligned} \tag{38}$$

For the n-machine power system, the terminal voltage of generators can also be expressed in x–y co-ordinates as follows [26]:

$$\overline{V}_g = \overline{E}'_q - jX'_D\overline{I}_g - j(X_Q - X'_D)\overline{I}_Q \tag{39}$$

Substituting (39) into (23) results in

$$I_g = \overline{C}_d\left[\overline{E}'_q - j(X_Q - X'_D)\overline{I}_Q + \overline{C}_{B1}\overline{V}_{B1} + \overline{C}_{B2}\overline{V}_{B2}\right] \tag{40}$$

where

$$C_d = (\overline{C}^{-1} + jX'_D)^{-1} \quad \overline{C}_{B1} = \overline{C}^{-1}\overline{F}_{B1} \quad \overline{C}_{B2} = \overline{C}^{-1}\overline{F}_{B2} \tag{41}$$

So in d-q co-ordinates it will be:

$$\begin{aligned} &\overline{I}_{Gi} = \overline{I}_{gi}e^{j\delta_i} = \sum_{k=1}^{n} \overline{C}_{dik}\left[E'_{qk}e^{j(90^\circ+\delta_k-\delta_i)}\right]+ \\ &(x_{qk} - x'_{dk})e^{j(\delta_k-\delta_i)}I_{qk} + \overline{C}_{B1k}\overline{V}_E e^{j\delta_i} + \overline{C}_{B2k}\overline{V}_B e^{j\delta_i} \end{aligned} \tag{42}$$

By denoting

$$\begin{aligned} \overline{C}_{dik} &= \overline{C}_{dik}e^{j\beta dik} \\ \overline{C}_{B1k} &= \overline{C}_{B1k}e^{j\beta B1k} \\ \overline{C}_{B2k} &= \overline{C}_{B2k}e^{j\beta B2k} \end{aligned} \tag{43}$$

Equation (42) so becomes

$$\begin{aligned} I_{di} &= \sum_{k=1}^{n} C_{dik}[-E'_{qk}\sin\delta_{ikd} + (x_{qk} - x'_{dk})I_{qk}\cos\delta_{ikd} \\ &+C_{B1K}V_{B1}\cos\delta_{B1K} + C_{B2k}V_{B2}\cos\delta_{B2k}] \\ I_{qi} &= \sum_{k=1}^{n} C_{dik}[-E'_{qk}\sin\delta_{ikd} + (x_{qk} - x'_{dk})I_{qk}\sin\delta_{ikd} \\ &+C_{B1K}V_{B1}\sin\delta_{B1K} + C_{B2k}V_{B2}\sin\delta_{B2k}] \end{aligned} \tag{44}$$

where

$$\begin{aligned} \delta_{ikd} &= \delta_k - \delta_i + \beta_{dik} \\ \delta_{B1k} &= \delta_{B1} + \delta_i + \beta_{dik} + \beta_{B1i} \\ \delta_{B2k} &= \delta_{B2} + \delta_i + \beta_{dik} + \beta_{B2i} \end{aligned} \tag{45}$$

Then, with linearization of the above equations is obtained:

$$\begin{aligned} \Delta I_d &= Y_d\Delta\delta + F_d\Delta E'_q + G_d\Delta v_{DC} + H_{B1d}\Delta m_{B1} \\ &+H_{B2d}\Delta m_{B2} + R_{B1d}\Delta\delta_{B1} + R_{B2d}\Delta\delta_{B2} \\ \Delta I_q &= Y_q\Delta\delta + F_q\Delta E'_q + G_q\Delta v_{DC} + H_{B1q}\Delta m_{B1} \\ &+H_{B2q}\Delta m_{B2} + R_{B1q}\Delta\delta_{B1} + R_{B2q}\Delta\delta_{B2} \end{aligned} \tag{46}$$

By ignoring the transfer function of governors (TM=0) and assuming a simple one-order AVR $[Avri(s) = K_A/1 + sT_{Ai}]$ with linearization (36) one gets [21]:

$$\begin{aligned} \Delta\dot{\delta} &= \omega_0\Delta\omega \qquad \Delta\dot{\omega} = M^{-1}(-\Delta T_E - D\Delta\omega) \\ \Delta\dot{E}'_q &= T'^{-1}_{Do}[(X_D - X'_D)\Delta I_D - \Delta E'_q + \Delta E_{QE}] \end{aligned} \tag{47}$$

$$\begin{aligned}
&\Delta\dot{E}_{QE} = (-\Delta E_{QE} - K_A \Delta V_T) T_A^{-1} \\
&\Delta T_E = \Delta I_Q \dot{E}'_{qo} + I_{QO} \Delta \dot{E}'_q \\
&+\Delta I_Q (X_Q - X'_D) I_{DO} + I_{QO} (X_Q - X'_D) \Delta I_D \\
&\Delta V_{TD} = X_Q \Delta I_Q \quad \Delta V_{TQ} = \Delta \dot{E}'_q - X'_D \Delta I_D
\end{aligned} \tag{48}$$

By substituting (46) into (48) can obtain:

$$\begin{aligned}
&\Delta T_e = K_1 \Delta\delta + K_2 \Delta E'_q + K_{pd} \Delta v_{DC} \\
&+K_{pb1} \Delta m_{B1} + K_{pdb1} \Delta\delta_{B1} + K_{pb2} \Delta m_{B2} \\
&+K_{pdb2} \Delta\delta_{B2} \\
&\Delta E_q = K_4 \Delta\delta + K_3 \Delta E'_q + K_{qd} \Delta v_{DC} \\
&+K_{qb1} \Delta m_{B1} + K_{qdb1} \Delta\delta_{B1} + K_{qb2} \Delta m_{B2} \\
&+K_{qdb2} \Delta\delta_{B2} \\
&\Delta V_t = K_5 \Delta\delta + K_6 \Delta E'_q + K_{vd} \Delta v_{DC} \\
&+K_{vb1} \Delta m_{B1} + K_{vdb1} \Delta\delta_{B1} + K_{vb2} \Delta m_{B2} \\
&+K_{vdb2} \Delta\delta_{B2}
\end{aligned} \tag{49}$$

Substituting Eq. (46) into Eq. (44) results in:

$$\begin{aligned}
&\begin{bmatrix} \Delta\dot{\delta} \\ \Delta\dot{\omega} \\ \Delta\dot{E}'_q \\ \Delta\dot{E}_{fd} \end{bmatrix} = \begin{bmatrix} 0 & \omega_0 I & 0 & 0 \\ -M^{-1}K_1 & -M^{-1}D & -M^{-1}K_2 & 0 \\ -T_{do}^{'-1}K_4 & 0 & -T_{do}^{'-1}K_3 & T_{do}^{'-1} \\ -T_A^{-1}K_A K_5 & 0 & -T_A^{-1}K_A K_6 & -T_A^{-1} \end{bmatrix} \\
&\begin{bmatrix} \Delta\delta \\ \Delta\omega \\ \Delta E'_q \\ \Delta E_{fd} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -M^{-1}K_{pd} & -M^{-1}K_{pb1} & -M^{-1}K_{pdb1} \\ -T_{do}^{'-1}K_{qd} & -T_{do}^{'-1}K_{qb1} & -T_{do}^{'-1}K_{qdb2} \\ -T_A^{-1}K_A K_{vd} & -T_A^{-1}K_A K_{vb1} & -T_A^{-1}K_A K_{vdb1} \end{bmatrix} \\
&\begin{bmatrix} 0 & 0 \\ -M^{-1}K_{pb2} & -M^{-1}K_{pdb2} \\ -T_{do}^{'-1}K_{qb2} & -T_{do}^{'-1}K_{qdb2} \\ -T_A^{-1}K_A K_{vb2} & -T_A^{-1}K_A K_{vdb2} \end{bmatrix} \begin{bmatrix} \Delta v_{DC} \\ \Delta m_{B1} \\ \Delta\delta_{B1} \\ \Delta m_{B2} \\ \Delta\delta_{B2} \end{bmatrix}
\end{aligned} \tag{50}$$

From (8) and (9) we have:

$$\begin{aligned}
i_{B1d} &= \frac{1}{(X_{T1} + X_{B1})} \left[V_{t1q} - V_{t2q} - \frac{m_{B1} v_{DC}}{2} \sin\delta_{B1} \right] \\
i_{B1q} &= \frac{1}{(X_{T1} + X_{B1})} \left[V_{t1q} - V_{t2q} - \frac{m_{B1} v_{DC}}{2} \cos\delta_{B1} \right]
\end{aligned} \tag{51}$$

$$
\begin{aligned}
i_{B2d} &= \frac{1}{(X_{T2}+X_{B2})}\left[V_{t1q} - V_{t3q} - \frac{m_{B2}v_{DC}}{2}\sin\delta_{B2}\right] \\
i_{B2q} &= \frac{1}{(X_{T2}+X_{B2})}\left[V_{t1q} - V_{t3q} - \frac{m_{B2}v_{DC}}{2}\cos\delta_{B2}\right]
\end{aligned} \tag{52}
$$

where

$$
\begin{aligned}
V_{tjq} &= \Delta E'_{qj} - x'_{dj}\Delta I_{dj} \\
V_{tjd} &= x_{qj}\Delta I_{qj} \qquad j = 1, 2, \ldots
\end{aligned} \tag{53}
$$

Linearization (1) by using (51) and (52) gets

$$
\begin{aligned}
&\Delta v_{DC} = \tfrac{1}{K_9+s}(K_7\Delta\delta + K_8\,\Delta E'_q + K_{cb1}\Delta m_{B1} \\
&+K_{c\delta1}\Delta\delta_{B1} + K_{cb2}\Delta m_{B2} + K_{c\delta b}2\Delta\delta_{B2})
\end{aligned} \tag{54}
$$

Substituting (54) into (50) results in (55). Where Δu could be ΔmB1, ΔmB2, ΔδB1, and ΔδB2, which are the deviations of the input control signals of the IPFC. Treating with the IPFC as an external controller installed in the network and its effect is included in the network current–voltage equation can be shown by Fig. 2. In the modeling of the IPFC, mainly the DC capacitor is included by expressing line

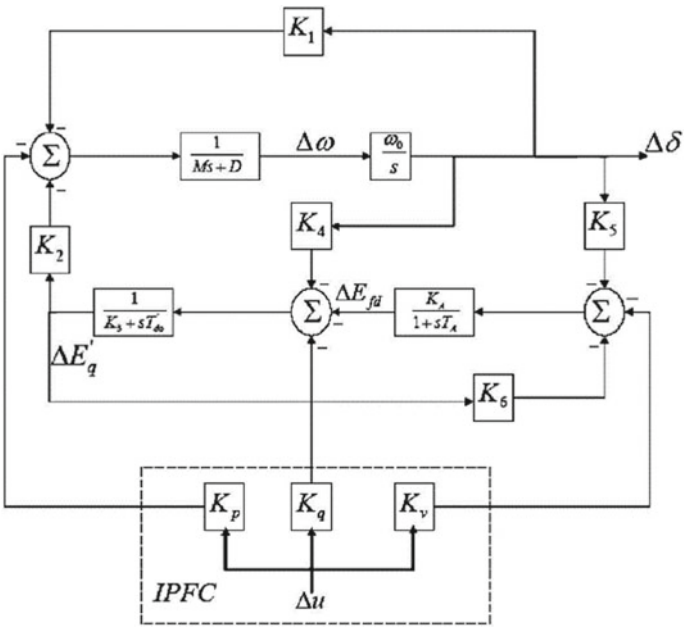


Fig. 2 The linear Phillips-Heffron model of the multi-machine power system with the installed IPFC

current explicitly. Thus the final form of the linear model of the IPFC shown in (55) is obtained. And (55) is expressed to the formal state-space form (56).

$$\begin{bmatrix} \Delta\dot{\delta} \\ \Delta\dot{\omega} \\ \Delta\dot{E}'_q \\ \Delta\dot{E}_{fd} \\ \Delta\dot{v}_{DC} \end{bmatrix} = \begin{bmatrix} 0 & \omega_0 I \\ -M^{-1}K_1 & -M^{-1}D \\ -T_{do}^{'-1}K_4 & 0 \\ -T_A^{-1}K_A K_5 & 0 \\ K_7 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 & 0 \\ -M^{-1}K_2 & 0 & -M^{-1}K_{pd} \\ -T_{do}^{'-1}K_3 & T_{do}^{'-1} & -T_{do}^{'-1}K_{qd} \\ -T_A^{-1}K_A K_6 & -T_A^{-1} & -T_A^{-1}K_A K_{vd} \\ K_8 & 0 & -K_9 \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \Delta\omega \\ \Delta E'_q \\ \Delta E_{fd} \\ \Delta v_{DC} \end{bmatrix} +$$
$$\begin{bmatrix} 0 & 0 \\ -M^{-1}K_{pb1} & -M^{-1}K_{pdb1} \\ -T_{do}^{'-1}K_{qb1} & -T_{do}^{'-1}K_{qdb1} \\ -T_A^{-1}K_A K_{vb1} & -T_A^{-1}K_A K_{vdb1} \\ K_{cb1} & K_{c\delta b1} \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 \\ -M^{-1}K_{pb2} & -M^{-1}K_{pdb2} \\ -T_{do}^{'-1}K_{qb2} & -T_{do}^{'-1}K_{qdb2} \\ -T_A^{-1}K_A K_{vb2} & -T_A^{-1}K_A K_{vdb2} \\ K_{cb2} & K_{c\delta b2} \end{bmatrix} \begin{bmatrix} \Delta m_{B1} \\ \Delta\delta_{B1} \\ \Delta m_{B2} \\ \Delta\delta_{B2} \end{bmatrix} \tag{55}$$

$$\dot{X} = AX + BU \tag{56}$$

4 Power System Oscillation Damping Controller

Each generator may have a damping controller as illustrated in Fig. 3 which is provided to improve the damping of power system oscillations. This damping controller consists of three parallel-connected blocks: Proportional, Integral and Derivative. The input signal of the controller is the angular velocity deviation of the generator and the output signal is control signals of the IPFC. To improve the damping oscillation and stability of the generators during disturbances in the power system, the controller coefficients are computed and coordinated with the gradient method. So after disturbances, the generators and power system reaches steady-state in a new operating point at the least possible time.

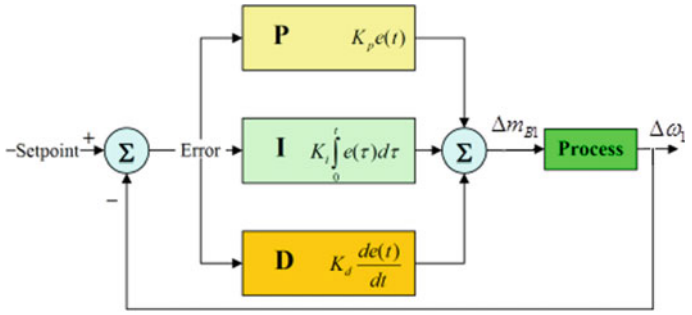


Fig. 3 structure of power system oscillation damping controller

5 Digital Simulation

In order to show the effects of the IPFC on low-frequency oscillations damping, digital simulations using MATLAB are done based on two cases; stable power system and unstable power system.

5.1 Stable Power System

The parameters of two-machines of the power system in Fig. 4 are given in the Appendix. Figure 4 shows that the IPFC is embedded in the lines 2–4 and 2–5 near

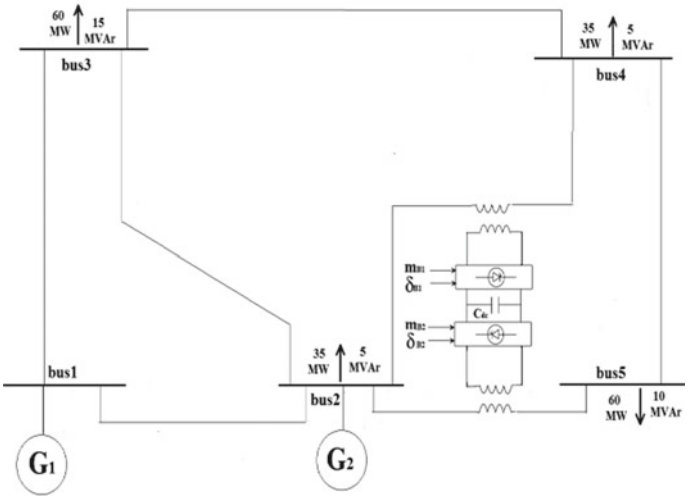


Fig. 4 Two-machine power system with the installed IPFC

the bus no. 2. In this paper, it is assumed that both active and reactive power flow of line 2–4 are controlled simultaneously. In other words, line 2–4 is called major line. The other line 2–5, which one of its active or reactive powers can only be controlled, is called the minor line. For static modeling of the IPFC in the power flow program, power injection modeling of the FACTS devices is used. In the case of a stable power system, it is planned that an IPFC regulates the power flow between bus 2 and 4–50 MW and 26 MVAr and between bus 2 and 5–70 MW.

After performing the linear Heffron-Philips power system model at the stable condition, eigenvalues of the system are presented in Table 1. From Table 1 can be concluded that the operating point is stable and all the eigenvalues of matrix A are located on the left side of the imaginary axis.

For dynamic simulation it is assumed, a three-phase fault occurred and after 0.1 sec, the fault cleared and the result of the fault, a step-change in mechanical power ($\Delta P_m = 0.1$ pu) is obtained. The Performance of the IPFC in two cases has been simulated and presented here.

First, the IPFC has no controller. In this case, the angular velocity deviations of two machines are shown in Figs. 5, 6 . It is clear from the figures that, the oscillation of the angular velocity deviation is poorly damped. Second, the simulation is performed with a step-change in mechanical input power, but the IPFC has the controller. In this condition to increase the damping of the generators separately, a PID controller for IPFC is maintained.

The input signal of the controller is the angular velocity deviation of each generator and the output signal is the control signal of IPFC (Δm_{B1} and Δm_{B2}). Also, for the fault case study in the power system, PID controller coefficients for the IPFC are computed and coordinated by the gradient method. Figures 7 and 8 show the effects of the designed controller for the IPFC on the angular velocity deviation damping of two generators when the IPFC has a tuned controller and without the controller. The results show the application and tuned control of inputs increase damping rate and it improves the damping of oscillation faster in comparison when there is no control on IPFC.

Table. 1 Eigenvalues of linear Phillips-Heffron model of test power system with the installed IPFC

35.23 + j0.5275
9.86 + j21.10
0.705 − j13.95
−4.88 − j7.0435
10.0 + j0.296
10.0 + j0.332
0.0346 − j0.023
0.027 − j0.3369
0.0233 − j0.328

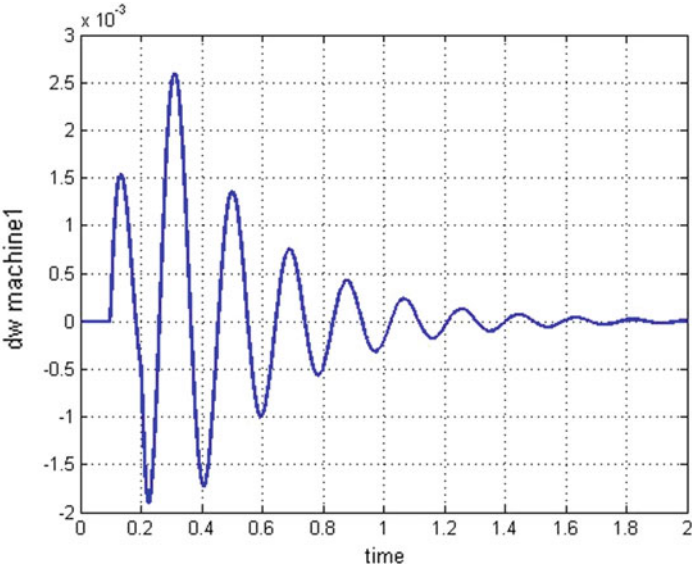


Fig. 5 Angular velocity deviation of machine1 during step change in mechanical input power without the controller

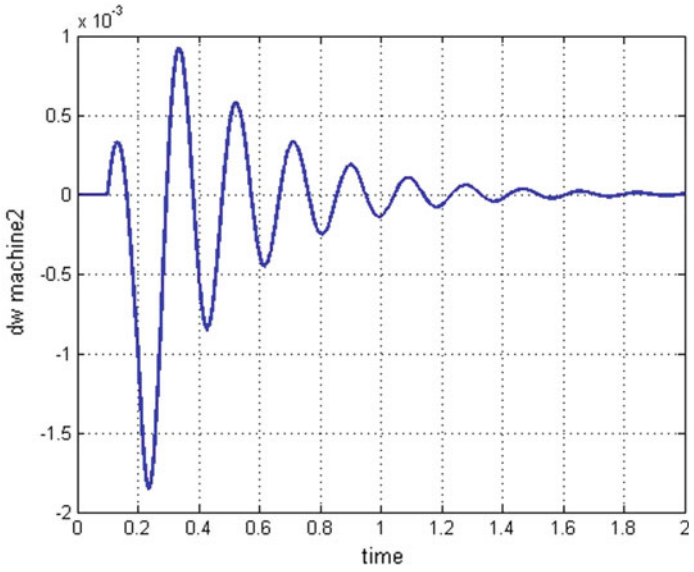


Fig. 6 Angular velocity deviation of machine 2 during step change in mechanical input power without the controller

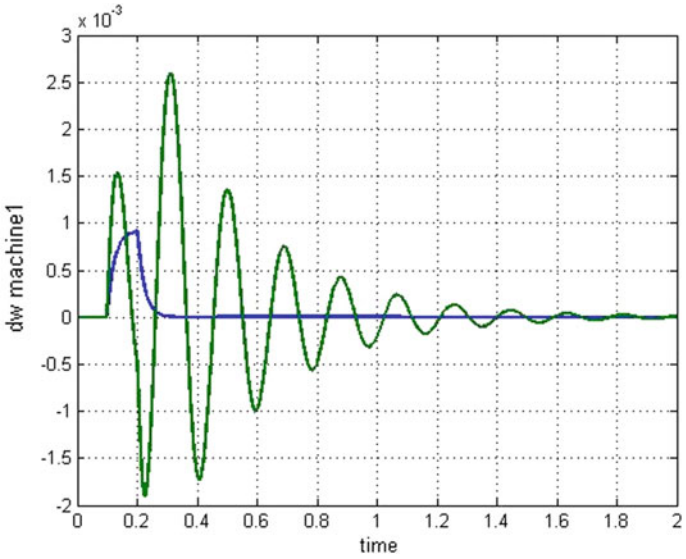


Fig. 7 Angular velocity deviation of machine1 during step change in mechanical input power with PID controller (curve with one overshoot)

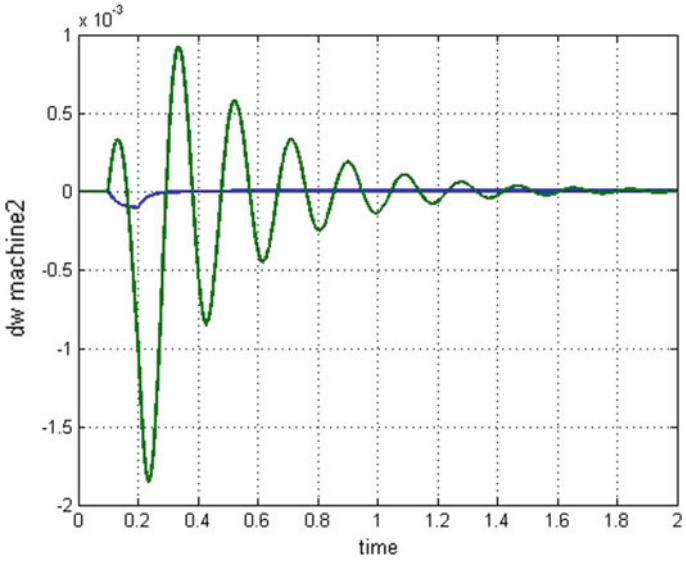


Fig. 8 Angular velocity deviation of machine2 during step change in mechanical input power with PID controller (curve with one undershoot)

Table 2 The eigenvalues of linear Phillips-Heffron model of test power system with the installed IPFC

13.5 – 24.5i
0.0741 + 0.0627i
−0.0253 − 0.334i
−0.0277 − 0.331i
−10.0 + 0.334i
−10.0 + 0.339i
−11.80 + 29.1i
−15.1 − 0.0627i
−32.30 − 3.84i

5.2 Unstable Power System

In the unstable power system case, it has been assumed that the IPFC is installed with a controller to regulate the power flow between bus 2 and 4–52 MW and 19 MVAr and between bus 2 and 5–74 MW. After performing linear Philips-Heffron power system modeling, the eigenvalues of the system is derived and presented in Table 2. It is clear from Table 2 that two eigenvalues of matrix A are located on the right side of the imaginary axis.

A three-phase fault occurred and after 0.1 s cleared and the result of the fault, a step-change in mechanical power ($\Delta P_m = 0.1$ pu) is obtained. The Performance of the IPFC in two cases has been simulated and presented. Like the previous case (A), at first, the IPFC has no controller. In this case, the angular velocity deviations of the generators are shown in Figs. 9 and 10 . As the figures show, the oscillations of the angular velocity deviations of the generators are not damped as the same as a previous case study.

Next, on the IPFC a PID controller is added which has the angular velocity deviation as input signals and the output signal is control signal of IPFC (Δm_{B1}. and Δm_{B2}). During fault occurrence in the power system, the PID controller coefficients are tuned with the gradient method. Figures 11 and 12 show the effects of the designed controller on the angular velocity deviations of generators. The results show the embedded IPFC and the tuned controller of inputs can stable the unstable power system which has eigenvalues that are located on the right side of the imaginary axes and in addition tune the PID controller coefficients increase the damping rate of power system oscillations.

6 Conclusion

In dynamical studies of the power system in the presence of FACTS device usually use a single machine infinite bus model, effects of one generators and FACTS device simulate and effects of other generators on the stability ignore. For the exact study

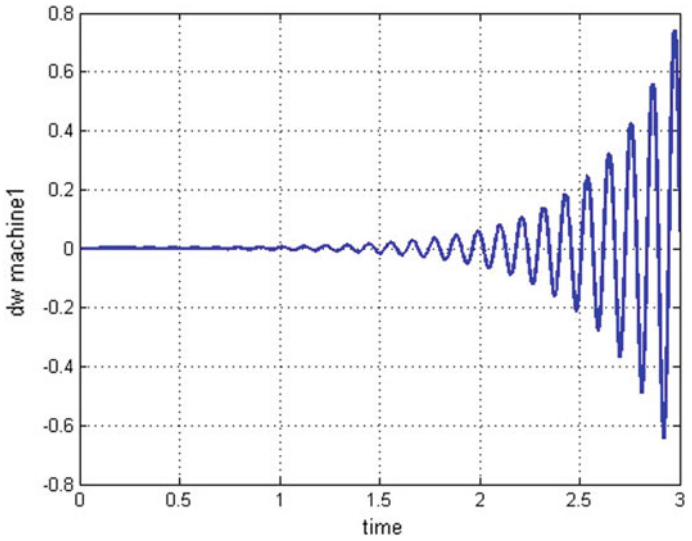


Fig. 9 Angular velocity deviation of machine1 during step change in mechanical input power without the controller

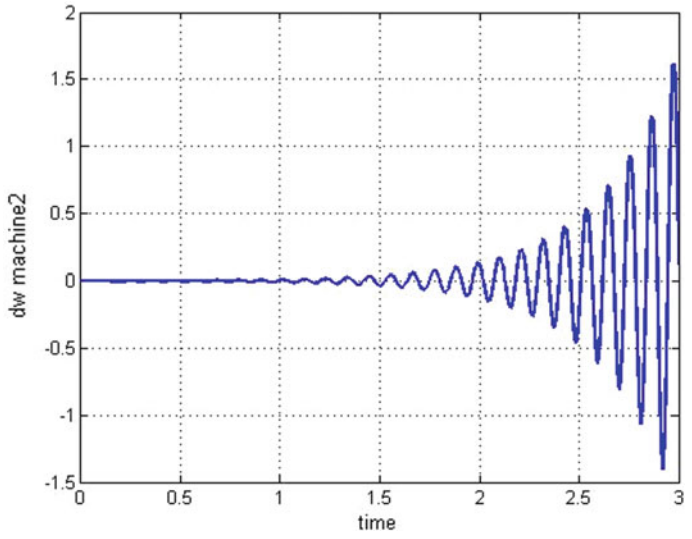


Fig. 10 Angular velocity deviation of machine 2 during step change in mechanical input power without the controller

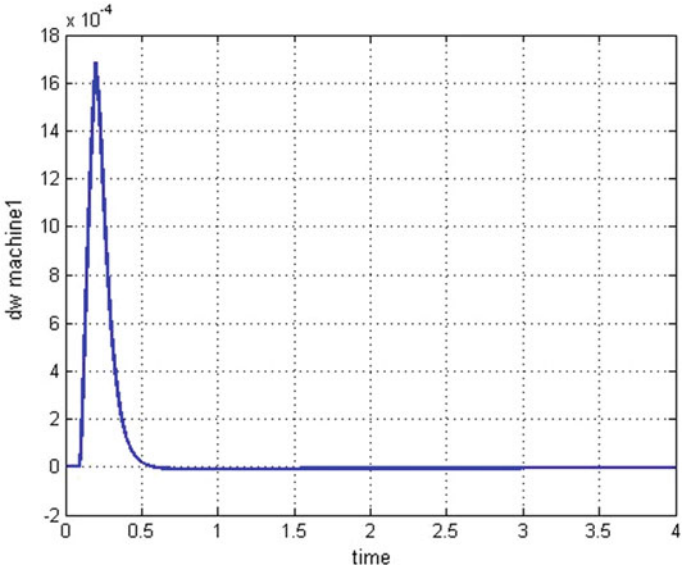


Fig. 11 Angular velocity deviation of machine1 during step change in mechanical input power with PID controller

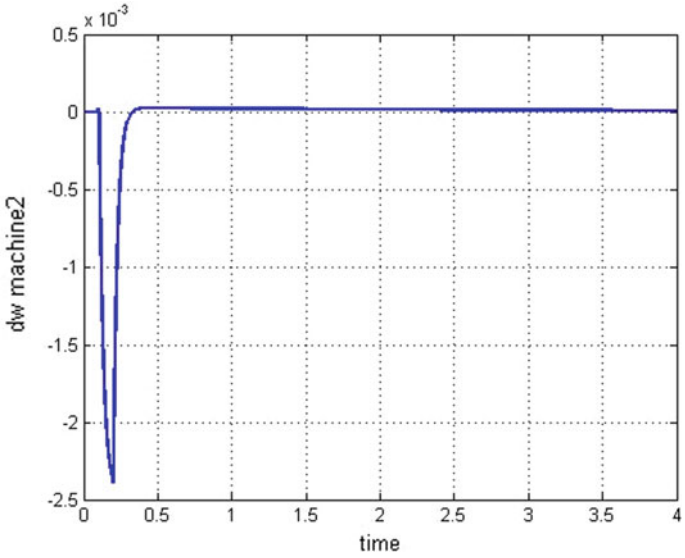


Fig. 12 Angular velocity deviation of machine2 during step change in mechanical input power with PID controller

should use the multi-machine model of the power system, the application of a nonlinear multimachine power system with FACTS device and linearized model has very complicated calculations. Although the duty of IPFC is power flow control in the power system, their potential of damping low-frequency oscillations has attracted interests.

Modeling is a three steps process that initiates with a conceptual, physical and mathematical approach. In this chapter, a linear multi-machine Phillips-Heffron mathematical model with the IPFC embedded in the power system transmission network is established. Mathematical modeling of the IPFC and incorporating it into the power system model is the basic aim of this work. Then by Using MATLAB software, the power flow of the test system in the presence of IPFC is carried out and the coefficients of the linear Phillips-Heffron model with including IPFC are derived. Based on this model, optimized PID damping controllers are suggested and are coordinated to improve the damping of the generators. In stable case angular velocity deviations of machine1 and machine 2 without controller damp in 2 s but with use of optimized controllers angular velocity deviations damp in less than 0.3 s. In addition, the amplitudes of angular velocity deviations limit to 10^{-3}, -0.5×10^{-3}. In the unstable case, angular velocity deviations don't damp and with the increase of time, the damping of machines decreases. With the use of designed optimized controllers, the stability improves and angular velocity deviations in 0.5 s reach zero and their amplitude restricted to 18×10^{-4} and -2.5×10^{-3} that shows the excellent application of IPFC.

The analysis of Eigenvalues and linear simulation on the test power system carried out in two cases (stable and unstable) shows that the modeling is efficient and also, IPFC is a powerful transmission network controller device that can properly be applied to damp low-frequency oscillations.

The advantages of this approach are:

1. The linear multi-machine Phillips-Heffron mathematical model with embedded IPFC is derived. Many studies often only consider the effect of one machine on stability (local modes), but in this process, in addition to local modes, the effects of other machines (inter-modal modes) are examined more comprehensively.
2. Application of IPFC in the multimachine power system and design tuned controllers of its inputs can increase the damping of low-frequency oscillations.
3. Lastly, IPFC can, in addition to controlling the load flow, cause the stability of unstable systems (with eigenvalues to the right side of the imaginary axis) that indicate the importance of IPFC in power systems performance.

References

1. Padiyar KR (2009) Facts controllers in power transmission and distribution. Anshan
2. Pai MA, Gupta DPS, Padiyar KR (2004) Small signal analysis of power systems. Alpha Science International

3. Quang NN et al (2019) Inter-area oscillations in the 500-kV Vietnamese power system. In: 2019 IEEE international conference on environment and electrical engineering and 2019 IEEE industrial and commercial power systems Europe (EEEIC/I&CPS Europe)
4. Kerperin A, Assenkamp A, Kreischer C (2019) PSS modification to stabilize synchronous machines during events with high rate of change of frequency. In: 2019 IEEE milan powertech
5. Renedo J, Sigrist L, Rouco L (2019) Design of power system stabilizers to damp low frequency inter-area oscillations with limited information. In: 2019 IEEE milan powertech
6. Ramirez-Gonzalez M, Malik OP (2008) Power system stabilizer design using an online adaptive neurofuzzy controller with adaptive input link weights. IEEE Trans Energy Convers 23(3):914–922
7. Wu W et al (2018) A novel damping strategy for low frequency oscillation suppression with MMC-type unified power flow controller. In: 2018 IEEE international conference on industrial technology (ICIT)
8. Nahak N, Sahoo SR, Mallick RK (2018) Design of dual optimal UPFC based PI controller to damp low frequency oscillation in power system. In: 2018 technologies for smart-city energy security and power (ICSESP)
9. Banaei MR, Kami A (2011) Interline power flow controller (IPFC) based damping recurrent neural network controllers for enhancing stability. Energy Convers Manage 52(7):2629–2636
10. Cai H et al (2013) Phillips-Heffron power system model for SmartPark and suppression of inter-area oscillations. In: 2nd IET renewable power generation conference (RPG 2013)
11. Pandey RK, Singh NK (2009) UPFC control parameter identification for effective power oscillation damping. Int J Electr Power Energy Syst 31(6):269–276
12. Hongesombut K, Kerdphol T, Weerakamaeng Y (2013) Robust interline power flow controller using phase-plane fuzzy logic control. In: 2013 10th international conference on electrical engineering/electronics, computer, telecommunications and information technology
13. Pradeep KM, Mini V (2018) Mitigation of inter area oscillations in a two area system using PSS and STATCOM. In: 2018 international conference on power, instrumentation, control and computing (PICC)
14. Naghshbandy AH, Faraji A (2019) Coordinated design of PSS and unified power flow controller using the combination of CWT and Prony methods with the help of SPEA II multi-objective optimisation algorithm. IET Gener Transm Distrib 13(21):4900–4909
15. Wang HF (1999) Design of SSSC damping controller to improve power system oscillation stability in Africon. IEEE, Cape Town
16. Hosseini SMH, Talebi N, Dehghanpour A (2011) Interline power flow controller and impact on ATC improvement. In: Othman M, Kasim RSR (eds) International conference on software technology and engineering, 3rd edn (ICSTE 2011). ASME Press, p 0
17. Xiao-Ping Zhang CR, Pal B (2008) Flexible AC transmission systems: modelling and control, Berlin-Germany.
18. Ebeed M, Kamel S, Jurado F (2016) Determination of IPFC operating constraints in power flow analysis. Int J Electr Power Energy Syst 81:299–307
19. Strzelecki R, Bojarski J, Benyesk G (2002) Probabilistic method for parallel filters power selection in interline power flow controller. In: 7th international conference 2002. Naples
20. Anton Vinkovic RM (2009) A current-based model of an IPFC for Newton-Raphson power flow electric power systems research 79(1):1247–1254
21. Yankui Zhang YZ, Chen C (2006) A novel power injection model of IPFC for power flow analysis inclusive of practical constraints. IEEE Trans Power Syst 21(4):1550–1556
22. Talebi N, Abedi A (2009) Automatic generation control using interline power flow controller. In: 2009 2nd international conference on power electronics and intelligent transportation system (PEITS)
23. Vural AM, Hamad MS (2018) Comparison of dynamic performances of IPFC, UPFC and back to back HVDC transmission on local and inter-area oscillation damping in power systems. In: 2018 5th international conference on electrical and electronic engineering (ICEEE)
24. Bahravar S et al (2018) Modeling and coordinated control of IPFC and SVC to improve the low frequency oscillation damping. In: Conference on in Electrical Engineering (ICEE), Iranian

25. Padiyar KR (2004) Power system dynamics: stability and control. Anshan
26. Wang HF (1999) Applications of modelling UPFC into multi-machine power systems Generation. Trans Distrib, IEE Proc 146(3):306–312

Techno-Economical Analysis of Energy Storage Systems in Conventional Distribution Networks

Amin Foroughi Nematollahi, Abolfazl Rahiminejad, and Behrooz Vahidi

Abstract Growing trend of integrating renewable energy resources in conventional Distribution Networks and their intermittent output increase the need for Energy Storage Systems. In this chapter, a closed form equation is proposed for optimal management of multi Energy Storage System in a conventional Distribution Network, analytically. Using the proposed approach, Energy Storage System can be embedded and utilized optimally in today's Distribution Network using the existing facilities. In order to find optimal performance of the storage system, the objective function is solved analytically and a closed form equation is achieved for storage system performance. The Energy Storage System management is performed in order to minimize total cost of daily energy loss and energy supply of the system. In addition, technical assessment of Energy Storage System is also taken into account to obtain a situation in which utilization of Energy Storage System is economical. In the Optimization problem, energy price, storage utilization duration, amount of load demand, power loss of the system, costs, limits and characteristics of storage system are integrated in the objective function. The proposed approach is applied to two test systems with different load levels. Obtained results indicate that the proposed approach can be successfully applied to practical networks and enhance efficiency of the distribution systems.

Keywords Energy storage systems (ESS) · Technical and economic model · Energy storage management · Power distribution network

A. Foroughi Nematollahi · B. Vahidi (✉)
Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran
e-mail: vahidi@aut.ac.ir

A. Foroughi Nematollahi
e-mail: amin.forooghi@aut.ac.ir

A. Rahiminejad
Department of Electrical and Computer Science, Esfarayen University of Technology, Esfarayen, North Khorasan, Iran
e-mail: abolfazl.rahiminezhad@gmail.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_16

Nomenclature

A. Acronyms

ESS	Energy storage system
DN	Distribution network
PV	Photovoltaic
WTG	Wind turbine generators
TN	Transmission networks
SLDs	Storage-like devices
FC	Fuel cell
PEVs	Plug-in electric vehicles
HESS	Hybrid energy storage system
DGs	Distributed generations
DR	Demand response
BES	Battery energy storage
ST	Smart transformer
HEMS	Home energy management system
MILP	Mixed-integer linear programing
ADN	Active distribution network
MILP	Mixed-integer linear programing
OF	Objective function
SOC	State of charge
MaxCap	Maximum capacity of ESS
MinCap,	Minimum capacity of ESS
$P_{storage}$	The charge or discharge power of ESS
AvailableCap	The ESS capacity which can be used
LF	Load factor
Ch	Charging
D.CH	Discharging
ESSE	Energy storage system efficiency

B. Symbols/parameters

Cost	Network energy cost
C and D	Denote charging and discharging period
T_C	Duration of ESS charging
T_D	Duration of ESS discharging
Pr	Power price
PL	Active power loss
T	Time duration
P^{Ch} and P^{DCh}	Charging and discharging power of ESS, respectively
$V_i \angle \delta_i$	Voltage of ith Bus,
$r_{ij} + jx_{ij} = z_{ij}$,	The element (i,j) of impedance matrix
P_i and Pj	Net injected power at ith and jth buses
Es	The energy stored in the ESS

P_{Lch}	The active power loss of the system during charge period of ESS
η_s	The discharging efficiency
P_{LD}	Active power loss of the system during discharging period
$V^i_{\min} \leq V_i \leq V^i_{\max}$	The bus voltage and indices max and min denote the upper and lower bounds. k
Limitk	The thermal limit of line k, and
Sk	The power flow through line

1 Introduction

Today, ESSs are one of the fundamental devices that ensure durability of energy supply and improve the reliability of the power system. ESSs can be utilized in different forms and sizes. The characteristics of an ESS highly depend on the form of the stored energy. Energy can be stored as potential, kinetic, chemical, electromagnetic, thermal, etc. [1–3].

1.1 Motivation

Integration of renewable energy resources in DNs has attracted attentions, significantly. However, the most challenging issue related to these kinds of resources is their variable outputs which highly depend on weather conditions. Utilization of ESS is one the most appropriate ways to deal with this issue.

ESS refers to the apparatuses which shift generation and consumption of energy over time. These apparatuses save energy (charge) when there is excess energy so that it can be consumed (discharged) when required. It should be mentioned that the charging and discharging are asynchronous and there is a transform coefficient ($\eta < 1$) for both charging and discharging of the ESS [2, 4, 5].

In recent years, utilization of storage systems has increased dramatically due to growth of electric load, environmental issues, financial problems and development of technologies [6–12]. One of the most important advantages of ESS is reducing generation investment. To meet the load demands, 20% higher than estimated load should be generated and only 55% of total capacity is used annually [13]. However, by optimal management and employment of ESS, more power plants can operate close to their maximum capacity [14].

These systems which can be installed in both DN and TN, can enhance efficiency of utilizing both renewable resources and conventional fossil fuel generators [9–12, 15]. Lots of efforts have been performed for optimal utilization of ESS using different objective functions [16–19]. The most well-known targets that can be achieved using

ESSs are load leveling, stability, reliability, power quality, frequency control, loss reduction and source intermittency reduction.

ESS can be located near renewable energy-based generators such as PV and WTG in order to improve their unpredictable behavior. In fact, ESS with controllable behavior can be employed besides renewable energy resources to tackle output power fluctuations. Lots of efforts have been performed on optimal design of hybrid systems which mainly consist of renewable energy resources (PV and WTG) and ESSs (battery/hydrogen tank) [20].

The concept of smart grids and micro-grids are introduced as the future networks in which controllability and monitoring of the power system can be enhanced through employing communication infrastructure [21]. In such systems, distributed renewable energy generators are easily embedded into network and utilized effectively. Moreover, these systems provide an appropriate platform for utilizing SLDs including energy storage systems such as batteries and devices with similar properties such as PEVs. Optimal utilization of SLDs in smart grids can be performed easily due to existence of adequate infrastructure [15, 21, 22].

In addition to smoothing out renewable energy output and optimal utilization of ESS in smart grids, ESSs can also be integrated and utilized optimally in conventional DNs. They can be charged by the power drawn from the main grid and discharged to supply DN loads so that performance of the conventional network is enhanced economically and technically. Many researches have been conducted in this context [12, 16–19]. In this chapter, a new closed form equation is proposed for optimal utilization of ESS in conventional DNs in order to enhance efficiency of the DNs using analytical methods.

1.2 Literature Survey

Nearly, all the attempts regarding ESS utilization can be considered in four different areas including hybrid systems, smart grids, HESS, and conventional networks.

1. Hybrid system- lots of efforts have been performed for optimal design of hybrid systems which mainly include renewable energy resources and ESSs. These systems which are known as micro-grids can operate in two modes including stand-alone and grid-connected modes. Micro-grids consist of a set of DGs and ESS. These systems can supply a building (residential, commercial, or office), a market, or even a village. The stand-alone hybrid systems are mainly designed to supply the load demands in remote areas. Since these areas are far from the main grid, the more economical method for supplying the load is to establish a micro-grid consisting of generators which are based on renewable energies such as solar and wind. Due to unpredictable output of PVs and WTGs, presence of ESS is almost obligatory. These systems must be designed so that the load is supplied all the time, while the system is low cost. Thus, ESS must be designed optimally so that it can store an amount of energy which is sufficient for supplying

the load demand when there is no generation. In this situation, the number or the capacity of ESS must be obtained optimally and technically. Thus, there is no optimal placement or utilization for ESS. In other words, ESS must be controlled so that excess energy of generators is stored in ESSs, then they are discharged to load whenever needed. ESS which is mainly used for these systems is battery or hydrogen tank which is used to supply a FC. Some efforts in this area can be found in [23–30].

2. Smart grids—smart environment of distribution networks, micro-grids and even homes, facilitates optimal utilization of renewable energy resources, conventional resources, and ESS. In such networks, in addition to conventional ESS, SLD such as PEVs can also be operated as ESS [31–33]. In [34], Demand Response (DR), Smart management of PEVs, and smart management of thermal energy storage are performed to reduce the cost of energy consumption for residential sector. Game theory is used for optimal control of ESS and demand-side management in a smart power system in [35]. A stochastic energy management of PVs, PEV and as ESS has been performed in a smart home in [36]. The authors have tried to charge the PEVs in low tariff electricity time and supply the load with the lowest possible cost. Authors of [37] have implemented optimal sizing of BESS in a smart grid equipped with ST to reduce the rate of ST converters. The paper showed that the BESS incorporating with ST can operate more efficiently. A review on HEMS including management of generators especially renewable ones, ESS, and SLD such PEVs can be found in [38]. Mixed-integer linear programing (MILP) has been used for sizing of a hybrid system consisting of renewable DGs and ESS for a smart home in [39]. They have used MILP for long term tech-economical HEMS. Real-time optimal management of charge/discharge of ESS has been performed in [40] for a smart grid. The authors have predicted load and scheduled an optimal ESS management, then, they have proposed an on-line algorithm for real time energy management. Establishing a balance between energy saving and a comfortable lifestyle aiming for low cost energy consumption, user convenience rates and thermal comfort level in a smart home has been done in [41]. Lots of other efforts conducted on ESS and energy management in smart grids and smart homes, which are equipped with smart metering devices and communication infrastructure, can be found in the literature. It is completely obvious that in such systems, optimal management of ESS can be performed in real-time manner easily. Thus, high efficient operation of ESS is quite expected.
3. HESS—different storage devices have different features which determine their application and location of use. However, in some cases, a hybrid energy storage system is an efficient way for storing excess energy and supplying the load in a different situation. For instance, due to high energy density of batteries and high power density of ultra-capacitors, an HESS consisting of batteries and ultra-capacitors is an adequate ESS for a PEV. However, it is very important to manage such systems to have an optimal operation of HESS. In a majority of works performed in this area, the main objective is to reduce the current stress of batteries and increase their life time. Some of the attempts done in HESS management can be found in [42–46].

4. Conventional network- the three important concerns of today's societies are energy, environmental and financial concerns [47]. Researchers and experts have tried a lot to deal with these crises especially in the field of electricity energy production and management. Loss reduction methods, distributed generation incorporation, power market, micro-grids and smart grids are some of the proposed methods to reduce the cost, fuel consumption, and the environmental threats. Some of these methods such as DG integration and loss reduction methods are now applied to the power systems and some of them are not. ESS management is a novel method for increasing efficiency of today's distribution systems. As mentioned before, ESS can be easily managed in smart grids which actually refer to the future networks. However, in conventional networks (i.e., nowadays networks), there is no direct communication between producers and consumers. Thus, ESS management in conventional networks with existing facilities is very important for enhancing efficiency of the networks and fossil fuel producers. ESS management means managing charge and discharge of these devices to move the generation and consumption over time aiming to enhance network performance. Using ESSs, large and even small power plants can operate in their exact optimal operation points; while, the load variation can be smooth out by means of ESSs.

A real-time BES management has been done in [48]. In that paper, the load is forecasted, BES scheduling is performed, and the errors are met in real-time. Shen et al. have discussed the ADN expansion in the presence of ESS [49]. The charging and discharging of ESS has been performed aiming to shift peak and enhance reliability. Optimal control of ESS in a micro-grid integrated with renewable energy resources has been discussed in [50]. In that paper, an optimal power flow by which the optimal control of ESS is obtained has been introduced. Optimal power flow considering ESS has also been discussed in [51]. They have integrated ESS in their proposed OPF by adding simple charge/discharge dynamics of ESS. BES also have proposed an AC optimal power flow for optimal placing of large ESS in a power system in the presence of wind turbine generators [52]. In that paper, changes in storage allocation in the network are studied as a function of total storage budget and transmission line-flow constraints. Nik et al. have used a mixed integer second-order cone programming to solve optimal power flow of ADN for optimal sitting and sizing of distributed ESS [53]. The problem is solved as a multi-objective optimization problem in which the network voltage deviations; line congestion; power loss, ESS cost, load supplying cost; etc., are considered as the objectives. Tang et al. have performed the optimal placement and capacity of ESS in distribution networks aiming to minimize energy loss [54]. They have simulated the distribution network as a tree and have shown that the best place for ESS is near the leaves far from the substation.

As can be seen, lots of efforts have been done in the field of optimal allocation and management of ESS in conventional networks, smart grids and stand-alone micro-grids. This subject is one of the most interesting fields of research which motivated the authors to perform the current study.

1.3 Research Gap

Role of ESSs on performance enhancement of power systems cannot be denied. They reduce power losses by leveling out the load profile, increase reliability by storing the power in energy excess time and supplying the load in time of low generation, enhance the power quality by smoothing out the load variation and provide a better platform for renewable DG penetration by tackling their intermittent output. As mentioned before, lots of efforts have been done in the field of ESS management or allocation. Numerical and evolutionary methods are used for optimal management or ESS setting; however, there is no guarantee for optimal performance of ESS using these methods [55–57]. Moreover, the simulation for finding the optimal management of ESSs must be performed once again as the load varies. Since a series of optimization problems must be solved in each iteration and according to complexity of the problem, solving these problems would be time consuming and unreasonable. Furthermore, in the previous studies on optimal management of ESS, the energy price and duration of each level are considered to be constant.

Besides all appropriate efforts done in ESS management of conventional network, and to overcome the aforementioned shortcomings, an analytical approach for ESS management is proposed in this chapter. A novel OF which is a closed form equation is proposed. Using this closed form equation for optimal management of ESS, the exact optimal operation point can be achieved easily using analytical methods. The novelties of this chapter can be addressed as follows:

1. A new formulation is introduced for operation cost of the system in the presence of ESS.
2. A closed form equation is proposed for charging and discharging ESS with the aim of cost minimization.
3. Analytical solution is employed to solve the optimization problem.
4. The global optimal operation point of the system is reached mathematically.
5. Optimal management of ESS is achieved in real-time.
6. The energy cost based on consumption time is taken into account.
7. Using the proposed approach, effects of ESS parameters such as efficiency and charging rate on optimal performance of the ESS can be investigated easily.
8. A critical efficiency is calculated for each ESS by means of the proposed closed form equation. This critical efficiency illustrates whether utilization of ESS is beneficial for the system or not.

1.4 Chapter Organization

In Sect. 2, problem formulation is illustrated and the proposed objective function which is a closed form equation for ESS management is extracted. In this section, the algorithm for optimal allocation and management of ESS is presented. In Sect. 3, the simulation and results are presented for two test cases and the optimal management of ESS is discussed. Finally, the chapter is concluded in Sect. 4.

2 Problem Formulation

In both conventional and smart distribution grids, the power price varies in different load levels and periods. In other words, in a deregulated environment, power price is determined by the market participants based on network conditions. The system operator tries to minimize the overall cost of the system. In passive DNs, the overall cost consists of power consumption cost and active power loss. However, for the active DNs, the total costs include not only the cost of power consumption and active power loss, but also the cost of energy produced by different resources. Moreover, in such a network, revenue is obtained from selling output power of resources. For a network equipped with ESS, cost of charging and revenue of discharging are also added to system cost. Thus, the total cost of a system equipped with ESS can be formulated as follows (Eq. 1):

$$Cost = T_C \times \underset{C}{\Pr} \times P_{LC} + T_D \times \underset{D}{\Pr} \times P_{LD} + T_C \times \underset{C}{\Pr} \times P^{Ch} - T_D \times \underset{D}{\Pr} \times P^{DCh} \tag{1}$$

In this equation, indices C and D denote charging and discharging period, P_r, P_L, and T are power price, active power loss and time duration, and P^{Ch} and P^{DCh} are charging and discharging power of ESS, respectively. The first and second parts of the equation are costs of active power loss of the system in both charging and discharging periods, respectively. The third part is the cost of purchasing power for charging ESS, and the fourth part is the revenue of selling power by discharging the ESS.

The active power loss can be formulated as follows (Eq. 2) [58–60]:

$$P_l = \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\alpha_{ij} \left(P_i P_j + Q_i Q_j \right) + \beta_{ij} \left(Q_i P_j - P_i Q_j \right) \right] \tag{2}$$

where

$$\alpha_{ij} = \frac{r_{ij}}{V_i V_j} \cos\left(\delta_i - \delta_j\right) \tag{3}$$

$$\beta_{ij} = \frac{r_{ij}}{V_i V_j} \sin(\delta_i - \delta_j) \tag{4}$$

And $V_i \angle \delta_i$ is the Voltage of ith Bus, $r_{ij} + jx_{ij} = z_{ij}$, is the element (i,j) of impedance matrix, and P_i and P_j are net injected power at ith and jth buses, respectively,

Now, imagine that ESS is installed at bus m. Thus, the active power loss of the system during charge period of ESS can be written as follows (Eq. 5):

$$P_{LCh} = \sum_{j=1, j\neq m}^{N} \sum_{i=1, i\neq m}^{N} \left[\alpha_{Cij} \left(P_{Ci} P_{Cj} + Q_{Ci} Q_{Cj} \right) + \beta_{Cij} \left(Q_{Ci} P_{Cj} - P_{Ci} Q_{Cj} \right) \right]$$

$$+\alpha_{Cmm}\left((P_{Cm}+P^{Ch})^2+Q_{Cm}^2\right)$$
$$+\sum_{j=1,j\neq m}^{N}\left[\alpha_{Cmj}\left((P_{Cm}+P^{Ch})P_{Cj}+Q_{Ci}Q_{Cj}\right)+\beta_{Cij}\left(Q_{Ci}P_{Cj}-(P_{Cm}+P^{Ch})Q_{Cj}\right)\right] \quad (5)$$

The load level during charging period is different from the one in discharging period. Index 1 denotes charging period. The charging period lasts for T^1 and the energy stored in the ESS in this period is $Es = T_C \times P^{Ch}$. The stored energy is injected to the network in discharging period with time duration of T_D. It should be mentioned that T_D does not have to be continuous. In other words, the charged energy can be discharged whenever required. Considering the discharging efficiency of η_s the discharged power can be formulated as follows (Eq. 6):

$$P^{DCh} = \frac{\eta_s \times E_s}{T_D} = \frac{\eta_s \times T_C \times P^{Ch}}{T_D} \quad (6)$$

Based on aforementioned phrases, active power loss of the system during discharging period would be calculated based on the following equation (Eq. 7):

$$P_{LD} = \sum_{\substack{j=1\\ j\neq m}}^{N}\sum_{\substack{i=1\\ i\neq m}}^{N}\left[\alpha_{Dij}\left(P_{Di}P_{Dj}+Q_{Di}Q_{Dj}\right)+\beta_{Dij}\left(Q_{Di}P_{Dj}-P_{Di}Q_{Dj}\right)\right]$$
$$+\sum_{\substack{j=1\\ j\neq m}}^{N}\left[\begin{array}{l}\alpha_{Dmj}\left(\left(P_{Dm}+\dfrac{\eta_s \times T_C \times P^{Ch}}{T_D}\right)P_{Dj}+Q_{Di}Q_{Dj}\right)\\ +\beta_{Dij}\left(Q_{Di}P_{Dj}-\left(P_{Dm}+\dfrac{\eta_s \times T_C \times P^{Ch}}{T_D}\right)Q_{Dj}\right)\end{array}\right]$$
$$+\alpha_{Dmm}\left(\left(P_{Dm}+\frac{\eta_s \times T_C \times P^{Ch}}{T_D}\right)^2+Q_{Dm}^2\right) \quad (7)$$

Now, by substituting Eqs. (5) and (7) in Eq. (1), the cost function can be rewritten as follows (Eq. 8):

$$Cost = T_C \times \Pr_C \times \left\{\begin{array}{r}\displaystyle\sum_{\substack{j=1\\ j\neq m}}^{N}\sum_{\substack{i=1\\ i\neq m}}^{N}\left[\alpha_{Cij}\left(P_{Ci}P_{Cj}+Q_{Ci}Q_{Cj}\right)+\beta_{Cij}\left(Q_{Ci}P_{Cj}-P_{Ci}Q_{Cj}\right)\right]\\ +\displaystyle\sum_{\substack{j=1\\ j\neq m}}^{N}\left[\begin{array}{l}\alpha_{Cmj}\left((P_{Cm}+P^{Ch})P_{Cj}+Q_{Ci}Q_{Cj}\right)\\ +\beta_{Cij}\left(Q_{Ci}P_{Cj}-(P_{Cm}+P^{Ch})Q_{Cj}\right)\end{array}\right]\\ +\alpha_{Cmm}\left((P_{Cm}+P^{Ch})^2+Q_{Cm}^2\right)\end{array}\right\}$$

$$+T_D \times Pr_D \times \left\{ \begin{array}{r} \left(\sum_{\substack{j=1 \\ j\neq m}}^{N} \sum_{\substack{i=1 \\ i\neq m}}^{N} \left[\begin{array}{r} \alpha_{Dij}\left(P_{Di}P_{Dj} + Q_{Di}Q_{Dj}\right) \\ +\beta_{Dij}\left(Q_{Di}P_{Dj} - P_{Di}Q_{Dj}\right) \end{array} \right] \right) \\ + \sum_{\substack{j=1 \\ j\neq m}}^{N} \left[\begin{array}{r} \alpha_{Dmj}\left(\left(P_{Dm} - \dfrac{\eta_s \times T_C \times P^{Ch}}{T_D}\right)P_{Dj} + Q_{Di}Q_{Dj}\right) \\ +\beta_{Dij}\left(Q_{Di}P_{Dj} - \left(P_{Dm} - \dfrac{\eta_s \times T_C \times P_{ch}}{T_D}\right)Q_{Dj}\right) \end{array} \right] \\ +\alpha_{Dmm}\left(\left(P_{Dm} - \dfrac{\eta_s \times T_C \times P^{Ch}}{T_D}\right)^2 + Q_{Dm}^2\right) \end{array} \right\}$$

$$+T_C \times \underset{C}{\Pr} \times P^{Ch} - T_C \times \underset{D}{\Pr} \times \frac{\eta_s \times T_C \times P^{Ch}}{T_D} \tag{8}$$

As can be seen, cost is a function of charging power, and the other parameters are constant. It should be noticed that cost is a function of coefficients α and β which depend on voltage magnitude and angle. Updating values of α and β requires calculating load flow. However, numerical results show that variation of α and β by different sizes of ESS is small and negligible [59]. With this assumption, optimum size of ESS for each bus, given by the aforementioned relations can be calculated from the base case load flow (i.e., without ESS). Thus, by setting derivative of the Cost function with respect to charging power equal to zero, the optimal amount of power charging can be achieved (Eqs. 9–11).

$$\begin{aligned} \frac{\partial Cost(P^{Ch})}{\partial P^{Ch}} =& T_C \times \underset{C}{\Pr} \times \frac{\partial P_{LC}}{\partial P^{Ch}} + T_D \times \underset{D}{\Pr} \times \frac{P_{LD}}{\partial P^{Ch}} \\ &+ T_C \times \underset{C}{\Pr} - T_D \times \underset{D}{\Pr} \times \frac{\eta_s \times T_C}{T_D} = 0 \end{aligned} \tag{9}$$

$$\frac{\partial P_{LC}}{\partial P^{Ch}} = \underbrace{\sum_{\substack{j=1 \\ j\neq m}}^{N} \left[\alpha_{Cmj} \times P_{Cj} - \beta_{Cmj} \times Q_{Cj}\right]}_{A_1} + 2 \times \alpha_{Cmm} \times (P_{Cm} + P^{Ch}) \tag{10}$$

$$\begin{aligned} \frac{\partial P_{LD}}{\partial P_{ch}} =& \underbrace{\sum_{\substack{j=1 \\ j\neq m}}^{N} \left[\alpha_{Dmj} \times \left(-\eta_s \times \frac{T_C}{T_D}\right) \times P_{Dj} - \beta_{Dmj} \times \left(-\eta_s \times \frac{T_C}{T_D}\right) \times Q_{Dj}\right]}_{A_2} \\ &+ 2 \times \alpha_{Dmm} \times \left(-\eta_s \times \frac{T_C}{T_D}\right) \times \left(P_{Dm} + \left(-\eta_s \times \frac{T_C}{T_D}\right) \times P^{Ch}\right) \end{aligned} \tag{11}$$

Finally, by substituting both derivatives of loss functions ($\frac{\partial P_{LC}}{\partial P^{Ch}}$ and $\frac{\partial P_{LD}}{\partial P^{Ch}}$) in cost function derivate ($\frac{\partial Cost(P^{Ch})}{\partial P^{Ch}}$), the optimal amount of charging power is achieved as follows (Eq. 12):

$$P^{Ch} = \frac{\begin{bmatrix} -T_C \times \underset{C}{\Pr} \times (1 + A_C + 2 \times \alpha_{Cmm} \times P_{Cm}) \\ -T_D \times \underset{D}{\Pr} \times \left(-\eta_s \times \frac{T_C}{T_D} + A_D - \alpha_{Dmm} \times P_{Dm} \times \left(\frac{2 \times \eta_s \times T_C}{T_D} \right) \right) \end{bmatrix}}{T_C \times \Pr_C \times 2 \times \alpha_{Cmm} + 2 \times T_D \times \Pr_D \times \left(\frac{\eta_s \times T_C}{T_D} \right)^2 \times \alpha_{Dmm}} \tag{12}$$

Based on the obtained closed form equation, determining time duration of both charging and discharging periods, power price at both periods, ESS efficiency, and information of load flow, the exact optimal power charging of the ESS can be obtained.

- Technical assessment (Effect of ESS efficiency)

One of the important parameters in ESS management is its efficiency. Regarding Eq. (12), when $P^{Ch} \geq 0$, charging ESS is economical. Thus, it can be said that there is a critical efficiency for which P^{Ch} is zero. Whenever, efficiency of ESS is greater than this critical value, charging and discharging of ESS is economical. To obtain the critical efficiency, let P^{Ch} to be positive. The critical efficiency is then obtained as follows (Eq. 13):

$$\eta_{critical} = \frac{\Pr_C \times (1 + A_1 + 2 \times \alpha_{Cmm} \times P_{Cm})}{\Pr_D \times \left\{ 1 + \sum_{\substack{j=1 \\ j \neq m}}^{N} \left[\alpha_{Dmj} \times P_{Dj} - \beta_{Dmj} \times Q_{Dj} \right] + 2 \times \alpha_{Dmm} \times P_{Dm} \right\}} \tag{13}$$

As can be seen, the critical efficiency is related to different parameters such as duration, power price, amount of load in each load level and the network characteristics. By calculating the critical value of ESS efficiency, an appropriate ESS can be selected to be installed in the studied system.

- Constraints

In solving the ESS management in this study, there are two kinds of constraints including those related to ESS and those related to DN.

The constraints related to the network are voltage limit and thermal limit of the line which can be mathematically presented as follows (Eqs. 14 and 15):

$$V^i_{\min} \leq V_i \leq V^i_{\max} \tag{14}$$

$$S_k \leq \lim it_k \tag{15}$$

where V is magnitude of the bus voltage and indices max and min denote the upper and lower bounds. $Limit_k$ is the thermal limit of line k, and S_k is the power flow through line k.

The constraints related to ESS include limited capacity, charging rate and discharging rate.

Capacity constraint: State of Charge (SOC) must lie in a predefined range as follows (Eq. 16):

$$Min_{\text{Cap.}} \leq SOC(t) \leq Max_{\text{Cap.}} \quad t = 1:24 \tag{16}$$

where

$$SOC(t) = SOC(t-1) + P_{storage}(t) \tag{17}$$

$$Min_{\text{Cap.}} = (1 - DOD) \times Max_{\text{Cap.}} \tag{18}$$

$$Availble_{Cap} = Max_{\text{Cap.}} - Min_{\text{Cap.}} \tag{19}$$

In these equations, t is time, *MaxCap* is maximum capacity of ESS, SOC is state of charge, DOD is depth of discharge, *MinCap* is the minimum capacity of ESS, *AvailableCap* is the ESS capacity which can be used, and $P_{storage}$ is the charge or discharge power of ESS.

Rate Constraint: the amount of power for charging or discharging (i.e. $P_{storage}$) cannot exceed limitation in an hour. In other words, charging or discharging rate must be lower than a maximum value as follows (Eq. 16.20):

$$\left|P_{storage}(t)\right| \leq Max_{Rate} \tag{20}$$

To satisfy this constraint, $P_{storage}$ in charging and discharging mode can be obtained as follows (Eqs. 21 and 22):

In charging mode:

$$P_{storage}(t) = min\left\{P_{ch}(t), Max_{cap} - SOC(t-1), Max_{Ch.Rate}\right\} \tag{21}$$

In discharging mode:

$$P_{storage}(t) = \max\left\{P_{Disch.}(t), Min_{cap} - SOC(t-1), -Max_{DisCh.\text{Rate}}\right\} \tag{22}$$

2.1 Optimal Placement and ESS Scheduling

According to the previous section, optimal management of ESS can be achieved analytically. However, the place of ESS installation might also influence performance of DN. Procedure of optimal placement and scheduling of multi ESS are illustrated in the flowchart depicted in Fig. 1. It should be noticed that in this flowchart, N_{load_level} is the number of levels considered for daily load profile, and N_{bus} is the number of buses in the studied network. For each load level, the ESS is placed in all buses and optimal management of charging/discharging of the ESS is obtained. The best algorithm for charging/discharging ESS is obtained among different scenarios and the best place for ESS is obtained based on benefit of ESS for the network according to its location (Tables 1 and 2).

3 Simulation and Results

The proposed approach is applied to two standard test cases including 33-bus and 69-bus radial DNs in which the line data and bus data of these two systems can be found in [61], and [59], respectively. The economic and technical data related to load level, power price, power loss and system cost for both test systems are also extracted from [62]. This information is briefly presented in Table 3. Technical information of ESS is also illustrated in Table 4.

Simulations are performed in 4 different scenarios. In the first scenario, optimal placement and ESS scheduling are investigated. In the second scenario, optimal placement and scheduling of two ESSs are studied. In the third scenario, the ESS is placed in the system and only ESS management is considered. In the fourth scenario, effect of ESSE is investigated. All the scenarios are performed for both test systems.

3.1 Scenario 1: Optimal Placement and Scheduling of 1 ESS

In this scenario, optimal location and scheduling of an ESS are obtained. In order to investigate the effect of ESS efficiency, 5 different values are considered for ESS efficiency. The obtained results for both 33-bus and 69-bus test systems are shown in Tables 5, and 6, respectively.

A quick look reveals that the best place for ESS is almost near the first bus of the system, i.e., close to the upper grid. The reason is that in this location, power loss of the system in charging period does not vary dramatically. In other words, ESS can be considered as a load during the charging period. Thus, by installing ESS near the main grid, power required for charging ESS does not flow through the lines; this results in lower power loss. Moreover, lower voltage drops may occur in the system.

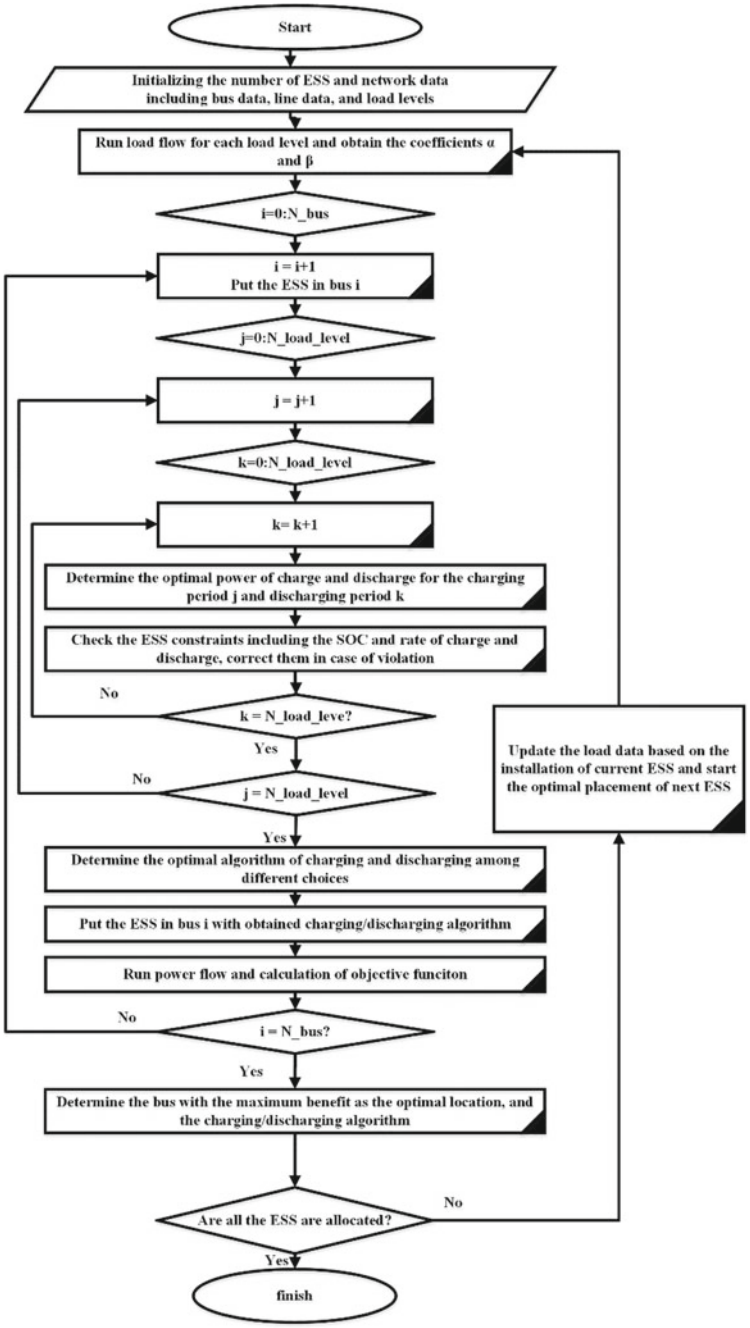


Fig. 1 The flowchart of optimal placement and ESS scheduling

Table 1 Technical and commercial information of 33 and 69 bus test systems

	LF	0.55	0.85	0.7	1	
	Load duration	6	6	6	6	
	Energy price	0.55	0.75	0.65	1	
Loss	69 bus	62.909	158.11	104.48	224.87	Net cost
	33 bus	59.45	148.73	98.52	210.97	
Cost	69 bus	207.6	711.5	407.48	1349.2	2675.7
	33 bus	196.22	669.3	384.25	1265.8	2518.3

Table 2 ESS technical information

	Maximum rate	Maximum capacity
33 bus	2500	10000
69 bus	2500	16000

Table 3 Optimal placement and scheduling of an ESS for different values of ESSE for 33-bus test case

η	Bus	LF	Status	Ch./D.Ch	Loss	Cost	Income	Benefit
1	5	0.55	Ch	1666	130	5929		3002
		0.85	D.Ch	−1666	101		7042	
		0.7	D.Ch	1666	183	7216		
		1	D.Ch	−1666	149		9105	
0.97	5	0.55	Ch	1692	131	6017		2541
		0.85	D.Ch	−1641	185		6928	
		0.7	Ch	1692	101	7321		
		1	D.Ch	−1641	149		8952	
0.94	5	0.55	Ch	1718	133	6109		2073
		0.85	D.Ch	−1616	102		8798	
		0.7	Ch	1718	187	7430		
		1	D.Ch	−1616	150		6814	
0.9	5	0.55	Ch	1757	179	6391		1288
		0.85	–		148	384		
		0.7	–		98	669		
		1	D.Ch	−1581	125		8733	
0.85	5	0.55	Ch	1807	138	6422		831
		0.85	–	–	148	669		
		0.7	–	–	98	384		
		1	D.Ch	−1536	152		8307	

Table 4 Optimal placement and scheduling of an ESS for different values of ESSE for 69-bus test case

η	Bus	LF	Status	Ch./D.Ch	Loss	Cost	Income	Ben.
1	47	0.55	Ch	2500	63	8459		5572
		0.85	D.Ch	−2500	158		10,539	
		0.7	Ch	2500	104	10,159		
		1	D.Ch	−2500	224		13,652	
0.97	47	0.55	Ch	2538	63	8584		4904
		0.85	D.Ch	−2462	158		10,370	
		0.7	Ch	2538	105	10,308		
		1	D.Ch	−2462	224		13,427	
0.94	47	0.55	Ch	2500	63	8459		4021
		0.85	D.Ch	−2352	104		9874	
		0.7	Ch	2500	157	10,159		
		1	D.Ch	−2352	224		12,765	
0.9	9	0.55	Ch	2500	148	8739		2713
		0.85	D.Ch	−2246	121		9564	
		0.7	Ch	2500	204	10,549		
		1	D.Ch	−2246	174		12,436	
0.85	7	0.55	Ch	2500	63	8694	–	1788
		0.85	–	–	158	711	–	
		0.7	–	–	104	407	–	
		1	D.Ch	−2116	224		11,601	

Thus, the ESS is placed in the buses close to the upper grid and the optimal strategy is to charge the ESS in low power price period and discharge in high price time.

It can also be inferred from the tables that ESSE has a significant influence on total benefit of optimal ESS scheduling. As can be seen, a mere 3% increase in ESSE increases the total benefit by 15% and 13% for 33-bus and 69 bus test cases, respectively. The differences between total benefit of both test systems in the case of ESSE of 100 and 85% are staggering. Moreover, the load factor (*LF*) does not influence charging or discharging rate of ESS. However, income of the network for higher LFs is bigger.

3.2 Scenario 2: Optimal Placement and Scheduling of 2 ESSs

In this scenario, optimal placement and scheduling of multi ESSs (2 ESSs) are investigated. Similar to the previous scenario, the optimizations are performed for different values of ESS efficiency. The optimization results are listed in Tables 7 and 8 for

Table 5 Optimal placement and scheduling of two ESSs for different values of ESSE for 33-bus test case

η	Bus	LF	Status	Ch./D.Ch	Loss	Cost	Income	Ben.
1	0.55	5	Ch	1666	142	11,472		7491
		19	Ch	1666				
	0.85	5	D.Ch	−1666	101		14,542	
		19	D.Ch	−1666				
	0.7	5	Ch	1666	197	13,772		
		19	Ch	1666				
	1	5	D.Ch	−1666	147		19,113	
		19	D.Ch	−1666				
0.97	0.55	5	Ch	1692	144	11,648		6397
		19	Ch	1692				
	0.85	5	D.Ch	−1641	101		14,313	
		19	D.Ch	−1641				
	0.7	5	Ch	1692	200	13,981		
		19	Ch	1692				
	1	5	D.Ch	−1641	148		18,807	
		19	D.Ch	−1641				
0.94	0.55	5	Ch	1723	147	11,861		5341
		19	Ch	1723				
	0.85	5	D.Ch	−1611	102		14,045	
		19	D.Ch	−1611				
	0.7	5	Ch	1723	202	14,234		
		19	Ch	1723				
	1	5	D.Ch	−1611	148		18,447	
		19	D.Ch	−1611				
0.9	0.55	5	Ch	1757	149	12,089		5248
		19	Ch	1757				
	0.85	5	D.Ch	−1581	102		13,768	
		19	D.Ch	−1581				
	0.7	5	Ch	1757	205	14,506		
		19	Ch	1757				
	1	5	D.Ch	−1581	149		18,076	
		19	D.Ch	−1581				
0.85	0.55	5	Ch	1807	153	12,436		4048
		19	Ch	1807				
	0.85	5	–		148	669.3		

(continued)

Table 5 (continued)

η	Bus	LF	Status	Ch./D.Ch	Loss	Cost	Income	Ben.
		19	–					
	0.7	5	–		98	384.2		
		19	–					
	1	5	D.Ch	−1536	150		17,538	
		19	D.Ch	−1536				

33-bus and 69-bus test systems, respectively. As expected, the best places for ESSs are determined to be close to the main grid. The results show that power loss of the system during charging period is nearly 10% higher than the previous scenario. In other words, increase in ESS capacity increases the power for charging ESS which results in higher losses. Power loss of the system during discharging period is approximately the same for both cases. However, total benefit of the system in the case of two ESSs is much higher than its value for one ESS. In other words, increase in ESS capacity enhances performance of the system. A higher ESS capacity means a higher load during low power price and higher generation in high power price. This surely increases benefit of system's power supply.

3.3 Scenario 3: ESS Scheduling

This scenario investigates the situation of scheduling one ESS. In other words, in this situation, ESS is placed in the system and the proposed approach is only performed for optimal ESS scheduling. In this scenario, the ESS efficiencies of 85 and 95% are considered. The predefined place for ESS is considered to be on buses 30 and 33 for 33-bus and 69-bus test systems, respectively. The results of optimal ESS scheduling for both 33-bus and 69-bus test systems are listed in Table 9. As can be seen, since location of ESS is not optimal, power loss is higher than its value compared to the case where the ESS is placed in the optimal location. For instance, the power loss of 33-bus test case for ESSE of 85% during charging mode is 242.65 for the case of non-optimal location and 138 for optimal placement of ESS. Total benefit of the system for optimal location is nearly 9 times higher than benefit of the case with non-optimal location for ESSE of 85%. Thus, optimal placement of ESS can affect system performance significantly. It is also obvious that as ESSE improves from 85 to 95%, the total benefit increases drastically.

The SOC, charging and discharging rate of ESS after optimal management for two different efficiencies including 95% (i.e. case 1), and 85% (i.e. case 2) are depicted in Fig. 2 for IEEE-33 bus test system and 69 bus test system. As can be seen, ESS with efficiency of 85% is not charged and discharged between 7 A.M to 18 P.M. In other words, during these hours, utilizing ESS is not economical. The reason is that

Table 6 Optimal placement and scheduling **of** two ESSs for different values of ESSE for 69-bus test case

η	LF	Bus	Status	Ch/D.Ch	Loss	Cost	Income	Benefit
1	0.55	36	Ch	2500	63.2	16,710		13818
		47	Ch	2500				
	0.85	36	D.Ch	−2500	158		21,788	
		47	D.Ch	−2500				
	0.7	36	Ch	2500	105	19,911		
		47	Ch	2500				
	1	36	D.Ch	−2500	224		28,651	
		47	D.Ch	−2500				
0.97	0.55	36	Ch	2500	63.7	16,710		12,255
		47	Ch	2500				
	0.85	36	D.Ch	−2425	158		21,118	
		47	D.Ch	−2425				
	0.7	36	Ch	2500	105	19,911		
		47	Ch	2500				
	1	36	D.Ch	−2425	224		27,758	
		47	D.Ch	−2425				
0.94	0.55	36	Ch	2500	63	16,710		10,411
		47	Ch	2500				
	0.85	36	D.Ch	−2337	158		20,328	
		47	D.Ch	−2337				
	0.7	36	Ch	2500	105	19,911		
		47	Ch	2500				
	1	36	D.Ch	−2337	224		26,704	
		47	D.Ch	−2337				
0.9	0.55	36	Ch	2500	63.7	16,710		8560
		47	Ch	2500				
	0.85	36	D.Ch	−2249	158		19,535	
		47	D.Ch	−2249				
	0.7	36	Ch	2500	105	19,911		
		47	Ch	2500				
	1	36	D.Ch	−2249	224		25,647	
		47	D.Ch	−2249				
0.85	0.55	36	Ch	2500	135	16,946		6342
		47	Ch	2500				
	0.85	36	–		104	407		

(continued)

Table 6 (continued)

η	LF	Bus	Status	Ch/D.Ch	Loss	Cost	Income	Benefit
		47	–					
	0.7	36	–		158	711		
		47	–					
	1	36	D.Ch	−2125	182		24,407	
		47	D.Ch	−2125				

Table 7 Optimal scheduling of an ESS for different values of ESSE

	η	Bus	LF	Status	Ch/D.Ch	Loss	Cost	Inc.	Ben.
33	0.95	30	0.55	Ch	1513	244	5802		1144
			0.85	D.Ch	−1438	125		7875	
			0.7	Ch	1150	244	5441		
			1	D.Ch	−1093	90		4512	
	0.85	30	Ch	Ch	1505	242	5769		92.2
			0.85	–	–	127	1053		
			0.7	–	–	98.5	669		
			1	D.Ch	−1280	148		6915	
69	0.95	33	0.55	Ch	2294	123	7980		1539
			0.85	D.Ch	−2179	128		−8882	
			0.7	Ch	1444	243	6136		
			1	D.Ch	−1372	205		6774	
	0.85	33	0.55	Ch	1837	101	6399		359
			0.85	–	–	158	711		
			0.7	–	–	104	407		
			1	D.Ch	−1562	248		7878	

Table 8 Critical values of efficiency

Discharging period→ Charging period↓	1	2	3	4
1	–	0.77066	0.83285	0.73491
2	1.2987	–	1.1719	0.83863
3	1.2007	0.85329	–	0.79026
4	1.3607	1.1925	1.2654	–

Table 9 Critical values of efficiency

Discharging period / Charging period	1	2	3	4
1	–	0.83625	0.78757	0.7125
2	1.2012	–	1.2378	0.79994
3	1.2697	0.85786	–	0.89546
4	1.4035	1.2514	1.1167	–

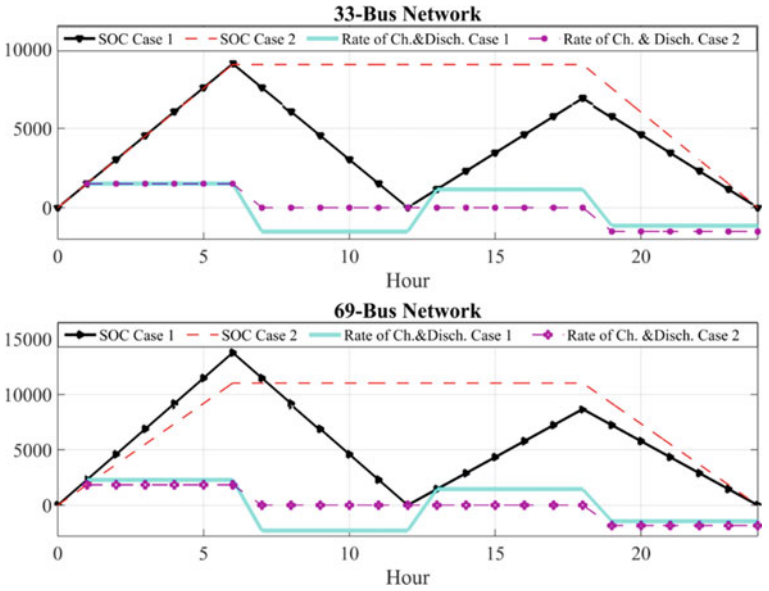


Fig. 2 SOC and rate of charge and discharge of ESS after management for case 1: efficiency of 95% and case 2: efficiency of 85% for both test systems

the critical efficiency (i.e. η critical) is higher than 85%; thus, any ESS with lower efficiency is not reasonable to be used.

3.4 Scenario 4: Investigating Effect of ESSE on System Cost Reduction

In this scenario, the effect of ESSE on system cost is investigated. This investigation is performed for both test systems.

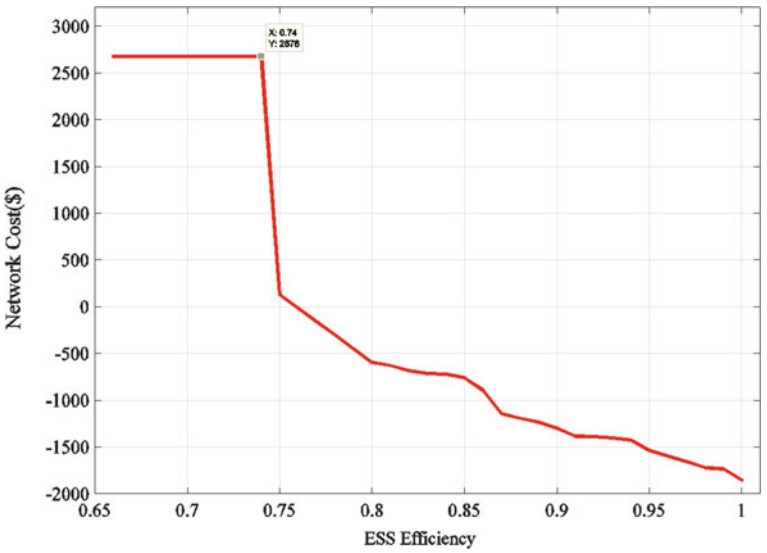


Fig. 3 Cost variation regarding ESSE for 69-bus test system

3.4.1 69-Bus Test Case

In this system, it is assumed that an ESS is placed at bus 33. The critical efficiency for charging in a period and discharging in another period is shown in Table 8. As expected, the optimal performance occurs when efficiency is higher than critical efficiency. For instance, when ESSE is 85%, the best situation occurs when the critical efficiency is minimum. The critical efficiencies which are shown in Tables 8 and 9 reveal that charging and discharging are economical, when ESSE is higher than critical value.

Variations of system cost regarding ESSE are depicted in Fig. 3. It is obvious that the critical efficiency is 73.49%. Any efficiency lower than this value does not affect system costs. Efficiencies higher than this value up to nearly 75% decrease the system cost dramatically. Any efficiency higher than 75% would be economically beneficial for the network.

3.4.2 33-Bus Test Case

In this system, ESS is placed on bus 30. Critical efficiencies of the ESS for charging and discharging in the mentioned bus are listed in Table 9. The cost variation regarding ESSE is presented in Fig. 4. As can be seen, the critical efficiency is 71.25%. Any efficiency higher than this value decreases system cost. Moreover, efficiencies higher than 84% not only decrease the cost, but they are also beneficial.

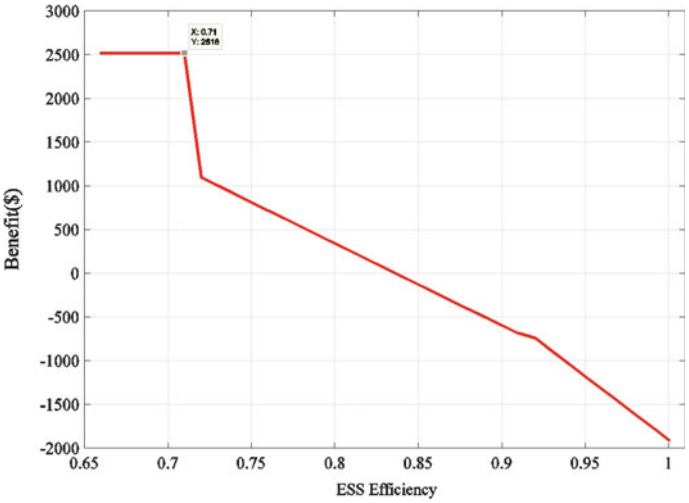


Fig. 4 Cost variation with regards to ESSE for 33-bus test system

4 Conclusion

Regarding increase of renewable energy resources integration in distribution networks, utilization of energy storage system experienced a dramatic increase. Thus, optimal management of these apparatuses especially in conventional distribution networks, is so important. In this chapter, a novel analytical approach for optimal management of energy storage system is proposed. In the proposed approach, the objective function of optimal placement and ESS scheduling is reducing system operation cost; while system constraints are satisfied. In the proposed approach, power loss, power price and load level are considered. The exact optimal operation point is obtained by analytical method and a closed-form equation is achieved. Advantage of the proposed approach compared to previous ones is that the exact operation point is obtained. Moreover, because of closed-form equation, procedure of calculating the proposed approach is much lower than the previous methods. Furthermore, in this chapter, effect of energy storage system efficiency on optimal performance of the ESS is investigated. Critical efficiency is also obtained analytically. Using this equation, minimum energy storage system efficiency can be obtained based on system condition. In other words, the technical requirement of ESS needed for a system can be obtained by the proposed approach. The proposed approach can be applied to different storage systems.

References

1. Nazir H, Batool M, Osorio FJB, Isaza-Ruiz M, Xu X, Vignarooban K, Phelan P, Kannan AM (2019) Recent developments in phase change materials for energy storage applications: a review. Int J Heat Mass Transf 129:491–523
2. Umair MM, Zhang Y, Iqbal K, Zhang S, Tang B (2019) Novel strategies and supporting materials applied to shape-stabilize organic phase change materials for thermal energy storage: a review. Appl Energy 235:846–873
3. Palacios A, Cong L, Navarro ME, Ding Y, Barreneche C (2019) Thermal conductivity measurement techniques for characterizing thermal energy storage materials: a review. Renew Sustain Energy Rev 108:32–52
4. Keck F, Lenzen M, Vassallo A, Li M (2019) The impact of battery energy storage for renewable energy power grids in Australia. Energy 173:647–657
5. Bayon A, Bader R, Jafarian M, Fedunik-Hofman L, Sun Y, Hinkley J, Miller S, Lipiński W (2018) Techno-economic assessment of solid–gas thermochemical energy storage systems for solar thermal power applications. Energy 149:473–484
6. Cutter E, Haley B, Hargreaves J, Williams J (2014) Utility scale energy storage and the need for flexible capacity metrics. Appl Energy 124:274–282
7. Solomon AA, Kammen DM, Callaway D (2014) The role of large-scale energy storage design and dispatch in the power grid: a study of very high grid penetration of variable renewable resources. Appl Energy 134:75–89
8. Johnston L, Díaz-González F, Gomis-Bellmunt O, Corchero-García C, Cruz-Zambrano M (2015) Methodology for the economic optimisation of energy storage systems for frequency support in wind power plants. Appl Energy 137:660–669
9. Santos JM, Moura PS, de Almeida AT (2014) Technical and economic impact of residential electricity storage at local and grid level for Portugal. Appl Energy 128:254–264
10. Lyons PF, Wade NS, Jiang T, Taylor PC, Hashiesh F, Michel M, Miller D (2015) Design and analysis of electrical energy storage demonstration projects on UK distribution networks. Appl Energy 137:677–691
11. Luo X, Wang J, Dooner M, Clarke J (2015) Overview of current development in electrical energy storage technologies and the application potential in power system operation. Appl Energy 137:511–536
12. Das T, Krishnan V, McCalley JD (2015) Assessing the benefits and economics of bulk energy storage technologies in the power grid. Appl Energy 139:104–118
13. Strbac G (2008) Demand side management: Benefits and challenges. Energy Policy 36:4419–4426
14. Lawder MT, Suthar B, Northrop PWC, De S, Hoff CM, Leitermann O, Crow ML, Santhanagopalan S, Subramanian VR (2014) Battery energy storage system (BESS) and battery management system (BMS) for grid-scale applications. Proc IEEE 102:1014–1030
15. Li Z, Guo Q, Sun H, Wang J (2015) Storage-like devices in load leveling: Complementarity constraints and a new and exact relaxation method. Appl Energy 151:13–22
16. Bradbury K, Pratson L, Patiño-Echeverri D (2014) Economic viability of energy storage systems based on price arbitrage potential in real-time US electricity markets. Appl Energy 114:512–519
17. Arghandeh R, Woyak J, Onen A, Jung J, Broadwater RP (2014) Economic optimal operation of Community Energy Storage systems in competitive energy markets. Appl Energy 135:71–80
18. Pearre NS, Swan LG (2015) Technoeconomic feasibility of grid storage: mapping electrical services and energy storage technologies. Appl Energy 137:501–510
19. Cho J, Kleit AN (2015) Energy storage systems in energy and ancillary markets: A backwards induction approach. Appl Energy 147:176–183
20. Li C-H, Zhu X-J, Cao G-Y, Sui S, Hu M-R (2009) Dynamic modeling and sizing optimization of stand-alone photovoltaic power systems using hybrid energy storage technology. Renew Energy 34:815–826

21. Rahbar K, Chai CC, Zhang R (2016) Energy cooperation optimization in microgrids with renewable energy integration. IEEE Trans Smart Grid
22. Sechilariu M, Wang B, Locment F (2013) Building integrated photovoltaic system with energy storage and smart grid communication. IEEE Trans Indus Electron 60:1607–1618
23. Ma T, Yang H, Lu L (2014) A feasibility study of a stand-alone hybrid solar–wind–battery system for a remote island. Appl Energy 121:149–158
24. Belmili H, Haddadi M, Bacha S, Almi MF, Bendib B (2014) Sizing stand-alone photovoltaic–wind hybrid system: Techno-economic analysis and optimization. Renew Sustain Energy Rev 30:821–832
25. Hosseinalizadeh R, Shakouri H, Amalnick MS, Taghipour P (2016) Economic sizing of a hybrid (PV–WT–FC) renewable energy system (HRES) for stand-alone usages by an optimization-simulation model: case study of Iran. Renew Sustain Energy Rev 54:139–150
26. Zhou W, Lou C, Li Z, Lu L, Yang H (2010) Current status of research on optimum sizing of stand-alone hybrid solar–wind power generation systems. Appl Energy 87:380–389
27. Bernal-Agustín JL, Dufo-Lopez R (2009) Simulation and optimization of stand-alone hybrid renewable energy systems. Renew Sustain Energy Rev 13:2111–2118
28. Yang H, Zhou W, Lu L, Fang Z (2008) Optimal sizing method for stand-alone hybrid solar–wind system with LPSP technology by using genetic algorithm. Sol Energy 82:354–367
29. Kellogg WD, Nehrir MH, Venkataramanan G, Gerez V (1998) Generation unit sizing and cost analysis for stand-alone wind, photovoltaic, and hybrid wind/PV systems. IEEE Trans Energy Convers 13:70–75
30. Nelson DB, Nehrir MH, Wang C (2006) Unit sizing and cost analysis of stand-alone hybrid wind/PV/fuel cell power generation systems. Renew Energy 31:1641–1656
31. Vigna KR, Gomathi V, Ekanayake JB, Tiong SK (2019) Modelling and simulation of variable speed pico hydel energy storage system for microgrid applications. J Energy Storage 24:100808
32. Arani AAK, Gharehpetian GB, Abedi M (2019) Review on energy storage systems control methods in microgrids. Int J Electr Power Energy Syst 107:745–757
33. Firouzmakan P, Hooshmand R-A, Bornapour M, Khodabakhshian A (2019) A comprehensive stochastic energy management system of micro-CHP units, renewable energy sources and storage systems in microgrids considering demand response programs. Renew Sustain Energy Rev 108:355–368
34. Brahman F, Honarmand M, Jadid S (2015) Optimal electrical and thermal energy management of a residential energy hub, integrating demand response and energy storage system. Energy Build 90:65–75
35. Nguyen HK, Bin Song J, Han Z (2015) Distributed demand side management with energy storage in smart grid. IEEE Trans Parallel Distributed Syst 26:3346–3357
36. Wu X, Hu X, Moura S, Yin X, Pickert V (2016) Stochastic control of smart home energy management with plug-in electric vehicle battery energy storage and photovoltaic array. J Power Sourc 333:203–212
37. Kumar C, Zhu R, Buticchi G, Liserre M (2018) Sizing and SOC management of a smart-transformer-based energy storage system. IEEE Trans Indus Electron 65:6709–6718
38. Zhou B, Li W, Chan KW, Cao Y, Kuang Y, Liu X, Wang X (2016) Smart home energy management systems: concept, configurations, and scheduling strategies. Renew Sustain Energy Rev 61:30–40
39. Erdinc O, Paterakis NG, Pappi IN, Bakirtzis AG, Catalão JPS (2015) A new perspective for sizing of distributed generation and energy storage for smart households under demand response. Appl Energy 143:26–37
40. Rahbar K, Xu J, Zhang R (2015) Real-time energy storage management for renewable integration in microgrid: an off-line optimization approach. IEEE Trans Smart Grid 6:124–134
41. Anvari-Moghaddam A, Monsef H, Rahimi-Kian A (2015) Optimal smart home energy management considering energy saving and a comfortable lifestyle. IEEE Trans Smart Grid 6:324–332
42. Tummuru NR, Mishra MK, Srinivas S (2015) Dynamic energy management of renewable grid integrated hybrid energy storage system. IEEE Trans Industr Electron 62:7728–7737

43. Song Z, Hofmann H, Li J, Hou J, Han X, Ouyang M (2014) Energy management strategies comparison for electric vehicles with hybrid energy storage system. Appl Energy 134:321–331
44. Shen J, Khaligh A (2015) A supervisory energy management control strategy in a battery/ultracapacitor hybrid energy storage system. IEEE Trans Transp Electri 1:223–231
45. Mendis N, Muttaqi KM, Perera S (2014) Management of battery-supercapacitor hybrid energy storage and synchronous condenser for isolated operation of PMSG based variable-speed wind turbine generating systems. IEEE Trans Smart Grid 5:944–953
46. Hu X, Johannesson L, Murgovski N, Egardt B (2015) Longevity-conscious dimensioning and power management of the hybrid energy storage system in a fuel cell hybrid electric bus. Appl Energy 137:913–924
47. Rahiminejad A, Faramarzi D, Hosseinian SH, Vahidi B (2017) An effective approach for optimal placement of non-dispatchable renewable distributed generation. J Renew Sustain Energy 9:15303
48. Bennett CJ, Stewart RA, Lu JW (2015) Development of a three-phase battery energy storage scheduling and operation system for low voltage distribution networks. Appl Energy 146:122–134
49. Shen X, Zhu S, Zheng J, Han Y, Li Q, Nong J, Shahidehpour M (2015) Active distribution network expansion planning integrated with centralized and distributed Energy Storage System. In: Power and energy society general meeting, 2015 IEEE. IEEE, pp 1–5
50. Levron Y, Guerrero JM, Beck Y (2013) Optimal power flow in microgrids with energy storage. IEEE Trans Power Syst 28:3226–3234
51. Gayme D, Topcu U (2013) Optimal power flow with large-scale storage integration. IEEE Trans Power Syst 28:709–717
52. Bose S, Gayme DF, Topcu U, Chandy KM (2012) Optimal placement of energy storage in the grid. In: Decision and control (CDC), 2012 IEEE 51st annual conference on, IEEE, pp 5605–5612
53. Nick M, Cherkaoui R, Paolone M (2014) Optimal allocation of dispersed energy storage systems in active distribution networks for energy balance and grid support. IEEE Trans Power Syst 29:2300–2310
54. Tang Y, Low SH (2017) Optimal placement of energy storage in distribution networks. IEEE Trans Smart Grid
55. Sousa T, Morais H, Vale Z, Faria P, Soares J (2012) Intelligent energy resource management considering vehicle-to-grid: a simulated annealing approach. IEEE Trans Smart Grid 3:535–542
56. Zhang X, Tan S-C, Li G, Li J, Feng Z (2013) Components sizing of hybrid energy systems via the optimization of power dispatch simulations. Energy 52:165–172
57. Yang P, Nehorai A (2014) Joint optimization of hybrid energy storage and generation capacity with renewable energy. IEEE Trans Smart Grid 5:1566–1574
58. Forooghi Nematollahi A, Dadkhah A, Asgari Gashteroodkhani O, Vahidi B (2016) Optimal sizing and siting of DGs for loss reduction using an iterative-analytical method. J Renew Sustain Energy 8:55301
59. Hung DQ, Mithulananthan N, Lee KY (2014) Optimal placement of dispatchable and nondispatchable renewable DG units in distribution networks for minimizing energy loss. Int J Electr Power Energy Syst 55:179–186
60. Hung DQ, Mithulananthan N (2013) Multiple distributed generator placement in primary distribution networks for loss reduction. IEEE Trans Indus Electron 60:1700–1708
61. Georgilakis PS, Hatziargyriou ND (2013) Optimal distributed generation placement in power distribution networks: models, methods, and future research. IEEE Trans Power Syst 28:3420–3428
62. Khalesi N, Rezaei N, Haghifam M-R (2011) DG allocation with application of dynamic programming for loss reduction and reliability improvement. Int J Electr Power Energy Syst 33:288–295

OPAL-RT Technology Used in Automotive Applications for PEMFC

Maria Simona Raboaca, Mihai Rata, Ioana Manta, and Gabriela Rata

Abstract This chapter aims to present the implementation and real time simulation stages of a PEM fuel cell using OPAL-RT technology. The following topics are presented: OPAL-RT technology, real-time simulation, conditions regarding the implementation of a mathematical model from Simulink/MATLAB in the RT-LAB platform using the OPAL-RT technology. Through the developed mathematical models, the authors can optimize the fuel cells, but also the integration of monitoring and control systems with the purpose of real time visualization of parameters as well as data acquisition using CAN, data storage and processing. The physical mathematical equations of the model were implemented in a programming language, in the form of code or block diagram, in order to be simulated in real-time. The advantages of real-time simulation of the mathematical models developed in the research projects are characterized by three decisive factors: speed of implementation, development flexibility and results predictability. The originality of the method is that the model can be simulated in real time (HIL) using an OPAL-RT architecture.

Keywords Proton exchange membrane fuel cell · PEMFC · OPAL RT technology · Real time simulator · Hardware in the loop · HIL

M. S. Raboaca (✉) · I. Manta
National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenic and Isotopic Technologies, Ramnicu Valcea, Romania
e-mail: simona.raboaca@icsi.ro

I. Manta
e-mail: ioana.manta@icsi.ro

I. Manta
Faculty of Power Engineering, Polytechnic University of Bucharest, Bucharest, Romania

M. Rata · G. Rata
Faculty of Electrical Engineering and Computer Science, Stefan Cel Mare University of Suceava, Suceava, Romania
e-mail: mihair@eed.usv.ro

G. Rata
e-mail: gabrielar@eed.usv.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_17

Nomenclatures

A. Acronyms

FC	Fuel cell
PEMFC	Proton exchange membrane fuel cell
CAN	Controller area network
HIL	Hardware in the loop
PIL	Power in the loop
SIL	Software in the loop
PHIL	Power hardware in the loop
EU	European union
ICE	Internal combustion engine
NCHFC	National center for hydrogen and fuel cell
DC	Direct current
AC	Alternative current
EV	Electric vehicle
PHEV	Plug-in hybrid electric vehicle
FCHV	Fuel cell hybrid vehicle
CFD	Computational fluid dynamics
OCV	Open circuit voltage
RT	Real time
RCP	Rapid control prototyping

B. Symbols/Parameters

H_2	Hydrogen
H^+	Hydrogen proton
e^-	Electron
O_2	Oxygen
H_2O	Water
CO_2	Carbon dioxide
ΔH	Enthalpy
$(h_f)_{H_2O}$	Heat of formation for water
$(h_f)_{H_2}$	Heat of formation for hydrogen
$(h_f)_{O_2}$	Heat of formation for oxygen
ΔG	Gibbs free energy
T	Temperature
ΔS	Entropy
$(s_f)_{H_2O}$	Water entropy
$(s_f)_{H_2}$	Hydrogen entropy
$(s_f)_{O_2}$	Oxygen entropy
W_{el}	Electrical work
q	Charge
E	Theoretical potential

n	Number of electrons per molecule
N_{AVG}	Avogadro's number
q_{el}	Charge of one electron
F	Faraday's constant
E_{th}	Theoretical fuel cell potential
$E_{T,P}$	Fuel cell potential taking into account temperature and pressure
R	Ideal gas constant
P_{H_2}	Partial pressure of hydrogen
P_{O_2}	Partial pressure of oxygen
P_{H_2O}	Partial pressure of water
E_{cell}	Fuel cell potential
α	Transfer coefficient
i	Current density
i_0	Exchange current density
i_L	Limiting current density
R_i	Total internal resistance of the fuel cell
η	Fuel cell efficiency

1 Introduction

On the 10th of November 2018, the European Commission presented the new joint energy strategy of the European Union, which proposes to provide competitive, sustainable and safe energy. The new framework agreed by the European Council sets the European Union's target to at least 27% in terms of renewable energy share in the EU by 2030.

Considering the increase estimates of the number of passenger vehicles by 2050, to 273 million in Europe and to 2.5 billion worldwide, a quasi-total decarbonization cannot be achieved only through development or efficiency increases for the current internal combustion engines (ICE), or only by using alternative fuels. The National Center for Hydrogen and Fuel Cells (NCHFC) has dedicated laboratories for PEM fuel cells. The laboratories are specialized on the development of fuel cells and stacks of high performance and the study of their behavior in different operating conditions with the purpose of improving their functional performances. The activities that are carried out in the laboratories are related to:

- Development of batteries and fuel cell stacks in different construction variants for low and high temperatures;
- Study of the behavior of fuel cells in transient regimes, repeated on-off cycles;
- Studies on the behavior of fuel cells for low temperature operation;
- Improvement of the functional performances of the fuel cells in the case of supplying gas mixture with different concentrations of toxic gases, both on the anode and cathode inlets, which simulate their real operating conditions;

– Optimization of the functional parameters for electricity production during load variations using DC/DC or DC/AC converters with adapters for input impedance.

At present, in the transport sector, the technology of electric or hybrid vehicles with direct power supply from the network is constantly expanding.

The Romanian state facilitates the purchase of Evs and PHEVs by offering support vouchers, so that the prices are similar to those of vehicles with internal combustion. In addition, both state agencies and private traders offer free EV & PHEV charging solutions, all to encourage clean transportation—this free charge is expected to be discontinued as EV & PHEV numbers increase abruptly, thus energy consumption will have to be managed properly, including through pricing.

In this context, the integration of the PEM fuel cell for EV & PHEV aims to replace fossil fuels, with a significant impact on the environment. Fuel cells are electrochemical devices that convert the chemical energy of a fuel (hydrogen) into electricity. Fuel cells are considered to be a viable solution when it comes to alternative energy sources. In recent years, there is an increasing interest for green energy and renewable energy, this topic being the focus of many research centers.

The fuel cell having a proton exchange membrane was invented in the 50 s. Currently, this fuel cell is used in mobile applications, such as the automotive field, for powering electrical portables, as well as for stationary electricity generation systems. The power of these cells varies between 1 and 100 kW.

In the case of proton exchange membrane fuel cells, the electrolyte is a very thin membrane. The most used material for the membrane is Nafion. The electrodes are made of woven or carbonic paper on which fine particles of the catalyst (usually platinum) are deposited.

The following reactions occur at the electrodes:

– At the anode: $2H_2 \rightarrow 4H^+ + 4e^-$
– At the cathode: $O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$

The operating temperature is about 80 °C.

The fuel cells domain is an objective of the big companies in the automotive field, hence, in order to reduce the costs, it is desired to develop fuel cell emulators using HIL and PHIL technology. In the literature, different mathematical models are presented related to integrated automotive applications simulated with Opal RT technology.

This chapter has as main objective the classification of PEM fuel cells, as well as the mathematical modeling and real time simulation of a PEM fuel cell using Opal RT technology.

As the topic of the book suggest, the mathematical modeling of the fuel cell represents a numerical method used for describing the functioning of a fuel cell, which converts chemical energy into electric energy.

The model of the proton exchange membrane fuel cell presented in this chapter aims to achieve a real-time emulator of the fuel cell. The need for a fuel cell emulator is due to the still very high cost of fuel cells. The role of the emulator is to replace a

real fuel cell [1]. It allows the real time simulation of the fuel cell's output voltage in both steady state and transient regime.

Fuel cells are considered as the energy sources of the future, representing a very active research field, but taking into account that there are some aspects to be improved, such as: cost, life time, etc. [2]. The emulator can be used in diagnosis or when it is desired to integrate a fuel cell into a real system, therefore the risk of fuel cell damage or destruction is excluded. The emulator also allows the validation of the fuel cell's auxiliaries as well as the control methods before their implementation in a real system.

This chapter is structured as follows: in *Introduction*, the authors present the main objective of this chapter; in the second part, the research location with its main infrastructure is described; the third part consists of the mathematical modeling of the fuel cell and it is followed by the presentation of the Opal-RT simulator in the fourth part; in the fifth and last part, the conclusions are drawn and future developments are presented.

Based on its experience in the field of hydrogen and fuel cells, the ICSI Energy department aims to develop stationary and mobile applications that make use of hydrogen and fuel cells. In this context, comparing the classical system of electricity production with the system of electricity generation based on fuel cells, the following can be observed:

- If H_2 is obtained through renewable energy sources, the fuel cells produce electricity without generating CO_2 emissions and does not pollute, thus reducing the greenhouse effect;
- The efficiency of a fuel cell-based system (including also the electric power consumed by the auxiliary components) decreases down to 40–45%;
- The fuel cell-based system is small in size. This system is relatively compact, precisely due to the conversion of electrochemical energy into electricity in a single stage.

2 Fuel Cell in ICSI Energy

ICSI Energy initiated in Romania the research activity in the production, storage and applications of hydrogen and fuel cells. At the same time, ICSI Energy represents a research and development facility having the mission to implement, develop and disseminate in Romania the energy technologies based on hydrogen, but also to support the national priorities in the field of energy and environment.

Researchers from the "Production and mobility of hydrogen" department focused on the development of technologies, products and services that compete to achieve a "hydrogen economy" in Romania. Based on the research infrastructure within the department, solutions were developed for hybrid mobility by carrying out experimental-demonstrative research that led to functional models which confirmed the degree of technical and commercial performances of hydrogen-based vehicles.

A special attention was given to PEM fuel cells due to the advantages of having a low operating temperature, a solid electrolyte and a compact shape. PEM fuel cells are used in a wide range of mobile and stationary applications, for example: cars, scooters and bicycle manufacturing industry, golf cars, utility vehicles, aerospace programs, military applications, ships and submarines, fuel cell powered locomotives, production of electrical energy both for domestic consumers and for the national electrical grid, being able to generate in the system powers between 1 and 33 kW. Within the department, there is an undergoing project related to the integration of PEM fuel cells into mobile charging stations for electric vehicles.

This study is part of the research activities within the project won by competition 36PCCDI/2018 with the title "Intelligent conductive charging stations, fixed and MobiLe, for electric propulsion transport" (SMiLE-EV), related to "Methods to simulate the functioning of the PEM fuel cell".

The concept of fuel cell-based vehicle represents a huge leap in the automotive field, for a quasi-total decarbonization that cannot be achieved only through development or efficiency increases for the current propulsion systems based on internal combustion engines or only by using some alternative fuels. The lack of hydrogen supply infrastructure, as well as the cost of hydrogen production that is still high, represent a barrier in the series implementation of fuel cells and hydrogen-based vehicles [3].

Car manufacturers have approached this Green to Green concept by showing the real benefits of using hydrogen as an alternative fuel. Currently, Toyota Motor Corp. developed a series standard hybrid electric car (FC-Bat). Based on the use of innovative simulation techniques [4–6] regarding the integration of fuel cells in the hybrid system, Toyota constantly reports the progress regarding the implementation of this technology [7, 8].

The increased attention of car manufacturers in the development and implementation of a new fuel cell concept derives from the following key factors: fuel cell is not polluting; it has high efficiency and high performance. Of course, worldwide, it has been found that there is a need for an environmentally friendly energy source, and the one having as waste water represents a viable candidate. The energetic performance of fuel cell-based vehicles remains a complex issue as the dynamics of the proton exchange membrane fuel cell is relatively slow for automotive applications. In order to overcome this major disadvantage, it is necessary to store the excess energy in a battery or in an ultracapacitor. This will maintain performance during peak power requests, which are taken over by the ultracapacitor. Using one of the energy storage solutions mentioned above, efficiency problems are also solved by recovering power during the deceleration of the vehicle [9, 10].

When designing the entire complex assembly of the hybrid vehicle with fuel cells, advanced techniques of modeling, design and testing are used, taking into account the behavior of the fuel cell, the battery, the converters and the propulsion system, as a whole unit, in different operating and functioning regimes [11, 12]. In research/design, analysis techniques based on mathematical models are used to develop control systems dedicated to hybrid vehicles. This is due to the fact that the equipment is expensive, the development cycles are relatively short and, last but not

least, due to the fact that the hybrid vehicles cu fuel cell—battery systems represent a new research direction in the automotive field.

The model-based control system requires an exact mathematical model of the different fuel cell-based hybrid vehicles (FCHV) subsystems. The modeling, development and testing of an FCHV is carried out on several levels. It starts from the modeling, dimensioning and control of each subsystem and, finally, the general optimization and control of the FCHV system [13, 14] is achieved through standard simulation or using hardware in the loop (HIL) simulators. In this study, the authors used Opal RT technology dedicated to real-time simulation of the proposed PEM fuel cell model.

3 Mathematical Modeling

Based on the literature study for the fuel cell as an important FCHV component, there are several mathematical models [15, 16]. In some cases, the mathematical models are not applied for the equilibrium state [17, 18], but, in most cases, these are equilibrium state models [13], mainly used for component sizing [19], cumulative fuel consumption [20], optimization of operating points or hybridization studies [21] and models of simulation [22]. The transient model has already been described as being dependent on temperature change. By applying small variations of current or a stepped load to the output terminals of the PEMFC, a dynamic of the behavior of a PEMFC stack (or cells) can be obtained [23, 24].

In recent years, through the use of advanced numerical computing algorithms, it has been possible to model PEMFC systems and individual components with better accuracy.

From the mathematical modeling point of view, for the PEMFC several components must be considered simultaneously: fluids that have multiple aggregation states and multi-dimensional flow, mass and heat transfer and electrochemical reactions. An example of a detailed dynamic model of fuel cells given in [13, 25, 26] includes reactant stoichiometry, hydration and voltage modeling in a single fuel cell and in a stack, but without taking into account the thermal effect on the PEMFC performance [27, 28].

A complete mathematical model is needed in order to characterize all physical, chemical and even mechanical simulations to better understand the complex phenomena that occur in an integrated FC system [11]. Moreover, a complete mathematical model is very important in the design, optimization and realization of a PEMFC.

The development and testing of a complex hybrid system are usually performed at several levels, starting from the individual subsystem sizing and control and ending with the general system optimization and control in standard and HIL simulation. Due to the size of the model and the technique of simulation used, the subsequent optimization and testing steps may take a longer period of time.

The mathematical model must be robust, accurate and provide fast solutions in the event of problems. For a wide range of operating conditions, the model must be able to determine PEMFC performance. The most important physical parameters of the cell to be included in a mathematical model for PEMFC are: the cell (physically—the parameters), the potential of the cell, the temperature, the pressure and the flow rates of the fuel and the oxidant, as well as the stoichiometry of the reactants.

One-dimensional mathematical models for thermal response and water management have been proposed, dynamic models to predetermine internal performances using the electrochemical reaction and the dynamic thermal equation. Two-dimensional models are also developed in order to determine PEMFC performances above and within (y–z axis) or along gas channels (x–z axis), but also three-dimensional mathematical models for evaluating PEMFC performances on all three axes (obviously, with the application of modeling restrictions in two or three dimensions) [29].

3.1 Modeling Criteria for a PEMFC Stack

By type: analytical, empirical or semiempirical. The equations distinguished in a mathematical model of a fuel cell can be: analytical equations, semiempirical equations and empirical equations [30, 49].

Analytical models. An analytical model uses the fundamental physical equations that allow the direct writing of the desired phenomenon. All parameters from the equation have a well-defined physical explanation. These equations are not specific only for a particular type of fuel cell, they also represent basic equations used for describing a phenomenon, which can be found in all types of fuel cells [31]. In the analytical models, the parameters are determined directly, starting from the material's physical characteristics from which the stack is made. For some situations, the characteristics or properties of a certain material cannot be easily measured or determined. In this case, the physical parameters can be obtained based on some laboratory experiments for the modeled stack. Thus, for each modeled cell, the parameters of the governing equations could be simplified and extracted according to experimental data. This type of model is the most general and easiest to understand for fuel cell modeling [15].

Semiempirical models. In the case of semiempirical models, the fundamental equations are maintained for given physical phenomena, but some of them are modelled starting from experimental tests. These empirical equations are obtained for a given material and well-known test conditions. However, they cannot be used outside the experimental validated conditions, so the model loses some of its generality. However, in this type of model, the analytical equations represent the majority of the equations used.

Empirical models. An empirical model mainly uses empirical equations determined by experiments. The conditions used for the validation of the model are quite restrictive. However, the empirical equations are considered to be simpler when using

this model type. For some situations, the empirical equation is represented by an interpolation. These empirical equations could also be extracted from the simplification of the base equations, given the specified test conditions.

By spatial dimension: 0-D, 1-D, 2-D, 3-D. A fuel cell can be developed according to needs with different spatial dimensions, described below [30].

0-D. A zero-dimensional model does not contain any equation including spatial dimension (x, y, z in the Cartesian plane). The physical equations of the model allow the description of the scalar variables, such as the voltage of a cell, the total amount of pressure from each channel, but they are not able to give the spatial distribution of a parameter like the temperature distribution for each cell. This type of model is frequently implemented in order to determine the polarization curve of a fuel cell.

1-D/pseudo 2-D. Compared to a zero-dimensional model, a 1-D model is able to describe on a spatial axis the physical phenomena [28]. The spatial axis is considered to be in the gas diffusion direction. With this type of model, the electrical, thermal and fluidic phenomena can be described according to the axis on which the diffusion takes place. For example, the distribution of water in the membrane can be obtained. The thermal effects can be introduced in this type of model in order to predict the temperature profile of each cell. However, the use of only one axis in modeling can limit the fluidic effects in the channels, since the direction of fluid in channels is perpendicular to the direction of gas diffusion. The 1-D and 2-D pseudo models described below are the most found models from the specialized literature. A pseudo 2-D model is similar to a 1-D model, but it also allows the description of fluidic pheromones in channels according to the fluid axis. Even so, the psuedo 2-D model cannot be considered a real 2-D model. We can observe two modeling axes in the model, but in a specific place, like the gas channels, there is only one modeling axis used for the model [28]. Both axes cannot be combined, meaning that the pair of coordinates (x, y) does not make sense.

2-D. A 2-D model comprises two modeling axes in the fuel cell layers. The two axes are chosen to be orthogonal axes in the fluid flow direction in the channels, which allow the clear modeling of the fluid field in 2-D channels. This type of model allows the study of several types of channels (straight, winded, inter-digital, etc.), and cannot be studied using a 1-D model. For the correct modeling of physical phenomena in 2-D through finite elements or finite volumes, fluid dynamic calculation (CFD) methods are applied.

3-D. A 3-D model is a complete model for a fuel cell. This type of model takes into account the three spatial axes for modeling the fuel cell [32]. With this type of model, the phenomena are described more precisely, for example, the convection of gases towards the diffusion layers of gases in channels (diffusion axis) in the same time with the flow of fluids (axis in the sense of fluids flow), which can be modeled in detail. In addition, the distribution of current density in the electric field, the thermal and fluid fields can also be present in 3-D. The CFD method is required for this model and, due to its complexity, the computing time is quite high.

By temporal nature. A stack model can also be described according to its temporal nature: a static model independent of time and a dynamic model which is time dependent.

Static models. A static model allows the description of the phenomena in the stack in a permanent regime (the parameters do not vary in time) [28]. This type of model does not contain the time derivatives of the state parameters in the physical equations. It is implemented for modeling fuel cells in static applications (power plants, uninterruptible power supplies, etc.) when the dynamic of the load is quite slow. It can also be implemented either in simple models of the fuel cell (for example, 1-D) in all application areas or in 2-D, 3-D models that allow the description of static fields of physical sizes.

Dynamic models. A dynamic model of the fuel cell is similar to the physical reality [33, 34]. For this model, the differential equations with respect to time are presented in a single domain or in several physical domains [1]. It allows the description of the transient regime between two operating points of the fuel cell. The dynamics is required for the modeling of stacks in mobile applications (for example in the automotive field) where the load dynamic is relatively high [35, 36]. In general, a dynamic model is often associated with a 1-D model, because dynamic modeling of a 2-D or 3-D model by the CFD method is usually reduced to a cell, or part of a cell.

By modeled species: stack, cell, individual layer. A fuel cell can be broken down into several individual layers [37]. It is not necessary for all layers to be included in some models: modeling the fluid channels in gas layers does not imply the membrane modeling. For a detailed model, each layer is individually modeled according to its physical properties [37]. Individual layers form the main elements for modeling a fuel cell. However, the individual layers and their detailed phenomena can be neglected and only one cell of the fuel cell stack can be considered. In fact, a fuel cell is composed of several cells connected in series, thus forming a stack of cells, in the literature being known as stack. In the model, the stack can be considered in general case without detailing the individual behavior of each cell. Thus, we discuss about the equivalent medium cell, after which, according to the hypothesis, all the other cells behave in an identical way: the model is isotropic. A fuel cell model can be obtained by stacking individual cells, themselves representing individual layers [38].

By modeled phenomena

Physical domains: electrochemical, *fluidic*, *thermal.* A fuel cell is a multi-physical device that covers different physical domains: electrochemical (electric), thermal and fluidic [30]. A model can cover all these fields or a single physical domain depending on the objectives pursued. As the fuel cell is a device that produces electricity, the electric model is generally introduced in all fuel cell models. Fluid phenomena, such as convection or diffusion, have a significant influence on the fuel cell performance. In order to have a more accurate model, the fluid field must also be introduced. If the cell temperature is to be controlled, the temperature variation due to conduction, radiation and convection must be taken into account [39]. In this case, the thermal field must be considered in the model.

Individual layers phenomena. Different physical features can be distinguished in each individual layer of the fuel cell. A complete cell model can take into account all phenomena in equations, but most of the existing models in the literature comprise only a part of these phenomena.

A usual model of a PEM fuel cell is a combination of several criteria presented above. For example, a model can be 1-D, dynamically and analytically, comprising three different physical domains with different phenomena modeled at individual layers.

Mann et al. [40] have introduced a generic model that can be applied for different types of fuel cells with different characteristics and sizes. The active surface and the thickness of the membrane are introduced as generic parameters. However, this model remains an isothermal model under permanent regime. In addition, it has only the electrochemical model. The obtained results are validated with a Ballard fuel cell powered by H_2–air and H_2–O_2.

Baschuck et al. [41] presented a model for studying the flooding phenomenon. A water layer is inserted between the gas diffusion layer and the catalyst.

The phenomenon of water diffusion in this layer is described by Henry's law. The purpose of this model is to predict the polarization curve by considering the flooding phenomenon in the fuel cell. The membrane is considered to be completely hydrated and despite the fact that the model is isothermal and in permanent regime, the results obtained are comparable to the experimental ones. This model shows that if the air is used as a fuel the electrode flooding can be reduced (as compared to the use of pure oxygen) due to a higher gas flow through the cathode channel. The temperature variation is not taken into account in this model, the simulation can be done only under certain operating conditions (permanent regime). In this model, the dynamics of the fuel cell is not taken into account.

Djihali et al. [42] presented another model of the fuel cell, focused on modeling the non-isothermal and non-thermal effects. This model takes into account the diffusion of gas through the porous electrode, the water transport caused by electro-osmosis, the convection, the generation of heat and the transfer of heat into the cell. The model was first validated under isothermal and isobaric conditions, then a non-isothermal and non-isobaric analysis was performed. The distribution of temperature through the cell was higher for several proposed thermal conductivities. The influence of temperature and pressure on cell performance and water transport in the cell were studied. This model shows that temperature and pressure variation are important in a fuel cell. However, it is a permanent regime model, it does not allow the simulation and evolution of the temperature in the transient regime.

Shan et al. [22] presented a complete model of the fuel cell. Both static and dynamic behavior are analyzed considering an uneven distribution of temperature. A stack with 10 cells was simulated, but the simulation results were not experimentally validated.

Haddad et al. [43] proposed a dynamic, nonlinear model of the fuel cell. Their model is an isothermal model for a single cell. For gas diffusion, Fick's second law was used in order to obtain the dynamic diffusion behavior. The electrical phenomenon was modeled through the electric circuit method in order to determine the activation losses, ohmic and double layer capacity. The effect of the variation of the electric load, the pressure and the humidification of the gas on the behavior of the fuel cell are analyzed and simulated. The fluid and cooling channels are not taken into account and the temperature is considered uniform.

Park et al. [1] have developed a dynamic model of a 20-cell stack. The effect of temperature and the effect of water in liquid and vapor state are taken into account. An analysis was made on starting the fuel cell. The start time allows the fuel cell to reach the nominal operating temperature, being proportional to the current required, the flow rate of the coolant and the temperature of the coolant. However, mechanical losses in supply channels and condensation in channels are not considered in this model. The results show a good dynamic response of the fuel cell, but the experimental validation of the dynamic response of the voltage and temperature of the battery was not achieved.

Pukuspran et al. [44] have developed a dynamic model designed for control. The model includes the characteristics of the flow and the dynamics of the air compressor, the dynamics of the volumes (the connections between different subsystems of the cell), the evolution in time of the partial pressure of the reactant and the water content of the membrane.

3.2 Fuel Cell Mathematical Modeling (Electrochemistry)

As mentioned before, the electrochemical reactions taking place in the same time in the fuel cell are the following:

At anode side: $H_2 \rightarrow 2H^+ + 2e^-$

At cathode side: $^1/_2O_2 + 2H^+ + 2e^- \rightarrow H_2O$

Overall: $H_2 + {}^1/_2O_2 \rightarrow H_2O$

For the overall reaction, one can compute the enthalpy (or heat) of reaction, denoted with ΔH, being defined as the difference of the enthalpies of formation for products and reactants:

$$\Delta H = \left(h_f\right)_{H_2O} - \left(h_f\right)_{H_2} - {}^1/_2\left(h_f\right)_{O_2} \tag{1}$$

where, $\left(h_f\right)_{H_2O}$, $\left(h_f\right)_{H_2}$ and $\left(h_f\right)_{O_2}$ represent the heat of formation for water, hydrogen and oxygen. Using the values from Table 1, at 25 °C, the enthalpy is:

$$\Delta H = -286 \text{ kJ mol}^{-1} \tag{2}$$

The amount of enthalpy which can be transformed in electricity is given by the Gibbs free energy, defined by the following equation:

Table 1 Enthalpy and entropy for products and reactants of the fuel cell

Reactant/product	Enthalpy	Entropy s_f (kJ mol^{-1} K^{-1})
Hydrogen (H_2)	0	0.13066
Oxygen (O_2)	0	0.20517
Liquid Water (H_2O)	−286.02	0.06996

$$\Delta G = \Delta H - T\Delta S \tag{3}$$

where ΔS is the entropy, which can be computed in a similar way with the enthalpy:

$$\Delta S = \left(s_f\right)_{H_2O} - \left(s_f\right)_{H_2} - {}^1\!/_2\left(s_f\right)_{O_2} \tag{4}$$

$\left(s_f\right)_{H_2O}$, $\left(s_f\right)_{H_2}$, $\left(s_f\right)_{O_2}$ representing the entropies for water, hydrogen and oxygen. Using the values from Table 1, at 25 °C, the entropy is:

$$\Delta S = -0.163285 \text{ kJ mol}^{-1}K^{-1} \tag{5}$$

Therefore, the Gibbs free energy for the overall reaction, at 25 °C (298.15 K), is:

$$\Delta G = -237.34 \text{ kJ mol}^{-1} \tag{6}$$

In order to obtain the theoretical fuel cell potential, we start from the general equation of the electrical work:

$$W_{el} = qE \tag{7}$$

where W_{el} is the electrical work, q is the charge and E is the theoretical potential. The charge q can be expressed as:

$$q = n * N_{AVG} * q_{el} = nF \tag{8}$$

where n is the number of electrons/molecule (for H_2 there are 2 electrons), N_{AVG} is the Avogadro's number (=6.022 * 10^{23} molecules/mol) and q_{el} is the electric charge of one electron (=1.602 * 10^{-19} C/electron); the product between N_{AVG} and q_{el} is known as Faraday's constant (F = 96485 C/electron mol)

Therefore, taking into account that the Gibbs free energy corresponds to the total amount of energy given by a fuel cell, the electrical work can be expressed as:

$$W_{el} = nFE = -\Delta G \tag{9}$$

By extracting the cell potential from the previous equation, the theoretical fuel cell potential is:

$$E_{th} = \frac{-\Delta G}{nF} = \frac{237340 \text{ J mol}^{-1}}{2 * 96485 \text{ As mol}^{-1}} = 1.23V \tag{10}$$

If the effect of pressure is also taken into account, the theoretical value decreases even further, according to the following equation:

$$E_{T,P} = E_{th} + \frac{RT}{nF} \ln \frac{P_{H_2} P_{O_2}^{0.5}}{P_{H_2O}} \tag{11}$$

where R is the ideal gas constant (=8.314 J mol^{-1}K^{-1}), and P_{H_2}, P_{O_2} and P_{H_2O} are partial pressures. For example, for a fuel with a temperature operating point of 60 °C, the resulted value of the potential will be:

$$E_{T,P} = 1.189V \tag{12}$$

When the fuel cell is operating, but without a closed electrical circuit, one expects to get an open circuit voltage (OCV) approximately equal to the theoretical fuel cell potential previously computed. In reality, there are some voltage losses further discussed which lead to a value actually less than 1 V. If the circuit is closed, the potential decreases even lower. The main voltage losses are related to:

- The activation polarization
- Internal currents and crossover losses
- The ohmic (or resistive) losses
- The concentration polarization

Therefore, a good representation of the fuel cell potential can be given by the following equation:

$$E_{cell} = E_{T,P} - \frac{RT}{\alpha F} \ln \frac{i}{i_0} - \frac{RT}{nF} \ln \frac{i_L}{i_L - i} - iR_i \tag{13}$$

where: α is the transfer coefficient, i is the current density, i_0 is the exchange current density, i_L is the limiting current density and R_i is the total internal resistance of the fuel cell.

The theoretical fuel cell efficiency is given by the equation below:

$$\eta = \frac{\Delta G}{\Delta H} = 83\% \tag{14}$$

In reality, this efficiency is much smaller than the theoretical one due to several factors: heat, electrode kinetics, electric and ionic resistance, mass transport, fuel processor, power conditioning, balance of plant etc.

3.3 Testing and Simulation Infrastructure

For modeling and developing PEMFC single cells or PEMFC stacks, an OPAL-RT simulation system was purchased within the ICSI Energy department. The simulators are widely used, as it has been proven to be effective in developing, optimizing and testing new technical process management solutions for PEMFC. OPAL-RT

includes the integrated software hardware for eMEGAsim, HYPERSIM, eFPGAsim, ePHASORsim [45]. The eMEGAsim software application is flexible and scalable, user-friendly, being a hybrid real-time simulator with analog and digital inputs and outputs and includes RT-LAB, ARTEMiS and RT-EVENT software. RT-LAB software is based on SimPowerSystem from MATLAB (SIMULINK) [46]. The great advantage of this software package is the fact that it meets the requirements of simulating the transient states of PEMFC and the electrical, electromagnetic and automation systems, in which PEMFC are integrated. SIL (software in-the-loop), HIL (hardware in-the-loop) and PIL (power in-the-loop) software is an important factor in PEMFC development and implementation, as it provides greater prediction and accuracy of the PEMFC behavior under different construction and operating conditions.

4 RT-LAB Multicore Simulator

RT-LAB is the real-time simulation software environment, which has revolutionized the model-based design mode.

Being flexible and scalable, RT-LAB can be used in almost any simulation application or control system, therefore adding computing power to the simulations. The RT-LAB software is fully integrated with MATLAB/Simulink.

The use of the OPAL RT Technology simulator is absolutely necessary because it can quickly correct the design errors before the physical implementation of the PEMFC hybrid system. In addition, the design time can be low if the parallelism is used in the implementation of the workflow. It is the indispensable tool for successful real-time simulations, such as: Hardware In-the-Loop (HIL) tests, Power In-the-Loop (PHIL) tests, Rapid Control Prototyping (RCP). The configuration of the OPAL RT simulator can be described as a distributed system in which the master PC manages all communications of both multi-targets and multi-host (Fig. 1).

Fig. 1 OPAL-RT technology system

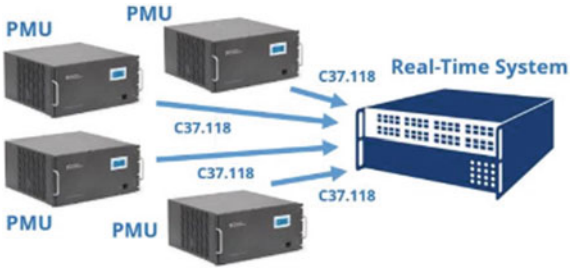


Fig. 2 Distributed architecture using C37.118 standard

The ability to use multi-host allows the grouping of end-users per PC-host, thus facilitating ping of targets. The master PC has total control of the accessed simulator. The host PCs have access only to view the signals obtained from the simulator. The management of the inputs and outputs can be achieved through the designated processors, distributed in several nodes of the network.

Regarding the communication protocols, OPAL-RT has implemented the following standards: IEC 61850 (digital communication interface implementation to non-conventional instrument transformers using GOOSE), IEEE C37.118 (synchronized phasor measurements used in electric power systems Standards, Fig. 2) [47, 48], DNP3 Distributed Network Protocol (IEEE Std 1815TM-2012), OPC, SPECTRACOM, Ethernet communication, CanOpen, Foundation Fieldbus, RS232 serial communication.

CAN represents the communication protocol between OPAL-RT equipment and fuel cell. For this communication predefined Simulink blocks were used, available in *RT-LAB/O/_CANdb_/Softing* library:

- *OpCanAc2Recv*—this block has the role of receiving the current signal;
- *OpCanAc2Send*—through this block the state variables are sent (signals to be viewed on the interface);
- *OpCanAc2Controller*.

Messages sent/received have a specific address, this address being specified in the block setting. Communication between the mathematical model and the CAN bus is made in the master subsystem.

The RT-Lab system is generally used to simulate electrical networks. RT-Lab proposes unique methods and specific solving solutions for real-time simulation in microseconds (μs). This feature is particularly important for models of AC networks (50/60 Hz) and for simulation of power electronics models (1 to 100 kHz). With RT-Lab, each period of an alternating voltage (current) or a PWM wave, can be simulated and observed in real time.

In recent years, the application areas for RT-Lab have started to converge, RT-Lab providing solutions for mobile applications. The real-time processor is the core of a real-time simulation platform. Its performance has a direct, very important impact on the overall performance of the system in real time. The RT-Lab system uses the

Intel *x*86 processor family. This family includes Core2 Duo, Quad, Intel Xeon, etc. Typical tact frequency of processors is between about 2 and 3 GHz.

In most cases, real-time simulation of very complex mathematical models cannot be performed with the help of a single processor. RT-Lab offers simulation solutions distributed on several processors. The critical problem, in a multi-processor system, is the efficient synchronization of multi-core processing during real-time simulation. In a typical configuration, RT-Lab proposes a multi-core processor (up to 6 computing cores, such as Intel Xeon) for multi-core simulation. The model can be sent on different cores, in a single processor. The synchronization and exchange of data between different cores is done through a high speed L1 memory, integrated in the processor.

On a real-time simulation platform, the model of the real-time user program must be stored and loaded into the processor during startup. During the simulation, the model output data could be saved in the platform for post-processing. RT-LAB uses common PC hard drives for data/software storage. Therefore, storage capacity can reach up to several GB.

Real-time simulation platforms are widely used in loop hardware tests. In order to interact with the external environment, the real-time platform should have different input/output (I/O) ports. Those ports include analog and digital inputs/outputs, digital/analog converters, analog/digital converters, PWM wave generator, CAN bus controller (Fig. 3), RS232 port, I2C port.

The device driver package is usually provided by a third-party company. Therefore, in some rare cases, there may be incompatibilities in the MATLAB/Simulink development environment.

RT-LAB uses a third-party Linux-based operating system—QNX. RT-Lab has modified the standard LINUX boot sequence to meet platform requirements over real time. It should be noted that LINUX is a standard *x*86 operating system, which is not just used for real-time simulations. In fact, the user can use standard LINUX commands on an RT-LAB platform like on a PC with LINUX operating system.

Fig. 3 PCI CAN-AC2 card used in the RT-Lab system for CAN bus communication

As it can be seen in the previous sections, RT-Lab uses the *x*86 PC and Linux standard to support their platform. Thus, the boot time of the system is close to that of a regular computer. After performing all the necessary actions: checking the BIOS, starting the different components, and loading the Linux, the boot time of an ordinary RT-Lab system, can take up to a minute.

RT-Lab uses MATLAB/Simulink (software developed by MathWorks) or LabVIEW (software developed by National Instrument) as a user monitoring interface, installed on the computer connected to the platform. RT-Lab does not offer its own interface among their solutions.

Method for implementing and testing the PEMFC mathematical model

The following methodology is a real-time test method with the aim of rapidly simulating a PEMFC under real conditions:

1. Create a new project
2. Build the model
   a. Subgrouping of the mathematical model into subsystems (Fig. 4)
   b. Naming the subsystems
   c. Adding OpComm blocks
   d. Maximizing parallel execution
   e. Setting simulation parameters
3. Load the Model on OPAL RT Simulator
4. Execute the Model
5. Use the Console to Interact with the Simulation
6. Stop the Simulation

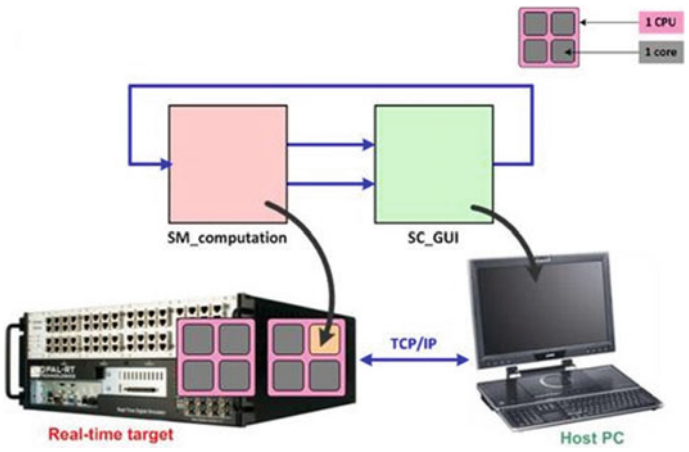


Fig. 4 Grouping the model into subsystems

The models developed, implemented and simulated in the RT-LAB software are dedicated to mobile and stationary applications, being used mainly in areas such as: the automotive industry, the naval industry (fully electric powered ships), the railway industry (electric trains and rail electric network), in aerospace industry (robotics and ship propulsion), energy, robotics, civil engineering, etc.

In order to execute the mathematical models realized in MATLAB/SIMULINK on the RT-LAB platform using the OPAL-RT simulator, the following requirements must be met:

- The mathematical models implemented in MATLAB/SIMULINK must run in SIMULINK without errors;
- The models made must be grouped into three subsystems called: Master, Slave and Console, each subsystem running in a different core of the simulator processor. Master and Slave have a role in the calculation of the elements of the SIMULINK model, and the Console has the role of display and user interface. In Fig. 4 you can see the Master subsystem loading into a real-time target core, connecting to the Host PC via TCP/IP and loading the Console subsystem into the PC-host;
- RT-LAB converts the SIMULINK model into C code;
- In order to save time in executing the implemented routines, the OpComm block is used for all of the input signals in the subsystems. In Fig. 5 an OpComm block with a single entry is represented;
- The time step must be fixed for real-time running of the model. In choosing the step, the hardware capacity of OPAL-RT and the requirements of the model are taken into account;
- Also, through RT-LAB, each subsystem (Master, Slave, Console) is loaded into the simulator cores.

Finally, the models are run with RT-LAB (Fig. 6) and analyzed in order to observe the behavior of the proposed system according to the conditions defined by the user.

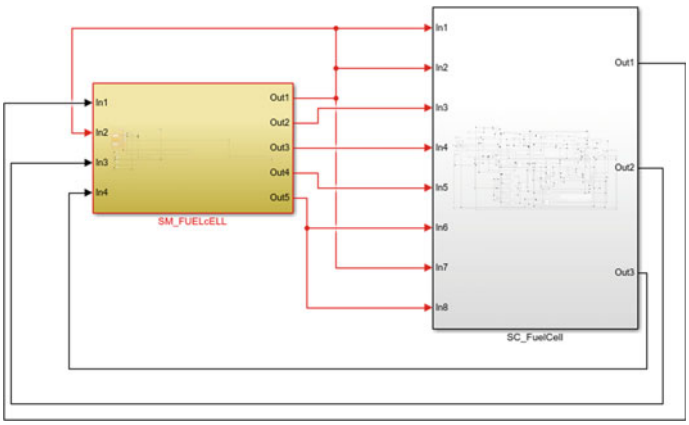


Fig. 5 OpComm block with one entry

Fig. 6 PEMFC RT LAB model

The mathematical model was run in RT-LAB, and the results obtained demonstrate the efficiency of RT-LAB in designing and testing PEMFC.

5 Conclusions

Real-time simulations represent a great advantage in the field of research/design by corroborating three decisive factors: the speed of implementation, the flexibility in development, the predictability of the results. In this chapter, the OPAL-RT simulation system was presented as an alternative in demanding/developing PEM fuel cells.

The most important advantage of this simulator is the connectivity between RT-LAB and MATLAB/Simulink. RT-LAB is a very powerful tool for designing, testing and analyzing power electronics and power systems. Such simulators are used in the defense industry, in the aerospace, automotive and obviously in academic research.

Acknowledgements This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI-UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0776/No. 36 PCCDI/15.03.2018, within PNCDI III.

References

1. Park JSK, Choe SY (2008) Dynamic modeling and analysis of a 20-cell PEM fuel cell stack considering temperature and two-phase effects. J Power Sources 179(2):660–672
2. Tirnovan R, Giurgea S, Miraoui A, Cirrincione M (2008) Proton exchange membrane fuel cell modeling based on a mixed moving least squares technique. Int J Hydrog Energy 33(21):6232–6238
3. Hansson G (2005) A bridge to the hydrogen highway. In: Proceedings of 21st electric vehicle symposium (EVS-21), Monte Carlo, Monaco
4. Barbir F (2005) PEM fuel cells: theory and practice. Elsevier Academic Press, London
5. Ursúa A, Sanchis P (2014) Modeling of PEM fuel cell performance: steady-state and dynamic experimental validation. Energies 7(2):670–700. https://doi.org/10.3390/en7020670
6. Han J, Charpentier JF, Tang T (2014) An energy management system of a fuel cell/battery hybrid boat. Energies 7(5):2799–2820. https://doi.org/10.3390/en7052799
7. Ishikawa T, Hamaguchi S, Shimizu T et al (2005) Development of next generation. In: Proceedings of 21st electric vehicle symposium (EVS-21), Monte Carlo, Monaco
8. Dufour C, Bélanger J, Ishikawa T et al (2005) Advances in real-time simulation of fuel cell hybrid electric vehicles. In: Proceedings of 21st electric vehicle symposium (EVS-21), Monte Carlo, Monaco

9. Matsumoto T, Watanabe N, Sugiura H et al (2001) Development of fuel-cell hybrid vehicle. Paper presented at the 18th international electric vehicle symposium, Berlin
10. Rabbath CA, Desira H, Butts K (2001) Effective modeling and simulation of internal combustion engine control systems. In: Proceedings of the American control conference, Arlington, Virginia
11. Le AD, Zhou B (2008) A general model of proton exchange membrane fuel cell. J Power Sources 182(1):197–223
12. Harakawa M, Yamasaki H, Nagano T et al (2005) Real-time simulation of a complete PMSM drive at 10 us time step. In: Proceedings of the 2005 international power electronics conference (IPEC 2005)—Niigata, Japan
13. Wingelaar PJH (2007) PEM fuel cell model representing steady-state, small-signal and large-signal characterisitcs. J Power Sources 171(1):754–762
14. Dufour C, Abourida S, Bélanger J (2003) Real-time simulation of electrical vehicle motor drives on a PC cluster. In: Proceedings of the 10th European conference on power electronics and applications (EPE 2003), Toulouse
15. Haraldsson K, Wipke K (2004) Evaluating PEM fuel cell system models. J Power Sources 126(1–2):88–97
16. Amphlett JC, Baumert RM, Mann RF et al (1995) Performance modeling of the ballard mark IV solid polymer electrolyte fuel cell. J Electrochem Soc 142(1):9–15
17. Bernardi DM, Verbrugge MW (1992) A mathematical model of the solid polymer-electrolyte fuel cell. J Electrochem Soc 139(9):2477–2491
18. Spiegel C (2008) Mathematical modeling of polymer exchange membrane fuel cells. Elsevier Academic Press, London
19. Chen YS, Lin SM, Hong BS (2013) Experimental study on a passive fuel cell/battery hybrid power system. Energies 6(12):6413–6422. https://doi.org/10.3390/en6126413
20. Yan W, Chen F, Wu H et al (2004) Analysis of thermal and water management with temperature-dependent diffusion effects in membrane of proton exchange membrane fuel cells. J Power Sources 129(2):127–137
21. Xue X, Tang J, Smirnova A et al (2004) System level lumped-parameter dynamic modeling of PEM fuel cell. J Power Sources 133(2):188–204
22. Shan Y, Choe S (2006) Modeling and simulation of a PEM fuel cell stack considering temperature effects. J Power Sources 158(1):274–286
23. Berning T, Djilali N (2003) Three-dimensional computational analysis of transport phenomena in a PEM fuel cell-a parametric study. J Power Sources 124(2):440–452
24. Wingelaar PJH, Duarte, JL, Hendrix MAM (2005) Dynamic characteristics of PEM fuel cells. Published in: 2005 IEEE 36th power electronics specialists conference, pp 1635–1641
25. Zhang G, Jiao K (2018) Multi-phase models for water and thermal management of proton exchange membrane fuel cell: a review. J Power Sources 391(1):120–133
26. Lee JH, Lalk TR, Appleby AJ (1998) Modeling electrochemical performance in large scale proton exchange membrane fuel cell stacks. J Power Sources 70(1):258–268
27. Pukrushpan JT, Stefanopoulou AG, Peng H (2002) Modeling and control for PEM fuel cell stack system. In: Proceedings of the 2002 American Control Conference, Anchorage, AK
28. Springer TE, Zawodzinski TA, Gottesfeld S (1991) Polymer electrolyte fuel cell model. J Electrochem Soc 138(8):2334–2342
29. Aschilean I, Rasoi G, Raboaca M et al (2018) Design and concept of an energy system based on renewable sources for greenhouse sustainable agriculture. Energies 11(5):1201
30. Gao F, Blunier B, Miraoui B (2012) Proton exchange membrane fuel cell modeling. ISTE Ltd and John Wiley & Sons Inc., London
31. Cheddie D, Munroe N (2005) Review and comparison of approaches to proton exchange membrane fuel cell modeling. J Power Sources 147(1–2):72–84
32. Mazumder S, Cole JV (2003) Rigorous 3-D mathematical modeling of PEM fuel cells. J Electrochem Soc 150(11):1510–1517
33. Breaz E, Tirnovan R, Botezan R et al (2011) Dynamic modeling of a proton exchange membrane fuel cell stack for real time simulation. Paper presented at the 4th international conference on modern power systems, Cluj–Napoca, Romania, 17–20 May 2011

34. Oneț O, Tîrnovan R, Breaz E et al (2010) Dynamic model development for PEM fuel cells using electrical circuits. Paper presented at the National Theoretical Electrotechnics Symposium SNET'10, Bucharest, Romania
35. Chrenko D, Gao F, Blunier B et al (2010) Methanol fuel processor and PEM fuel cell modeling for mobile application. Int J Hydrog Energ 35(13):6863–6871
36. Gao F, Blunier B, Bouquain D et al (2010) Emulateur de piles a combustible pour application de Hardware-in-the-Loop. In: Revue de l' electricite et de l'electronique
37. Gao F, Blunier B, Miraoui A et al (2009) Cell layer level generalized dynamic modeling of a PEMFC stack using VHDL-AMS language. Int J Hydrog Energ 34(13):5498–5521
38. Pukrushpan JT, Stefanopoulou AG, Peng H (2002) Simulation and analyses of transient fuel cell system performance based on a dynamic reactant flow model. Proceedings of ASME international mechanical engineering congress and exposition. New Orleans, Louisiana, USA, pp 17–22
39. Breaz E, Tîrnovan R, Oneț O, Vadan I (2010) The heat transfer of a proton exchange membrane hydrogen/oxygen fuel cell. Paper presented at the National Theoretical Electrotechnics Symposium SNET'10, Bucharest, Romania
40. Mann RF, Amphlett JC, Hooper MAI et al (2000) Development and application of a generalized steady state electrochemical model for a PEM fuel cell. J Power Sources 86(1–2):173–180
41. Baschuk JJ, Li X (2000) Modeling of polymer electrolyte membrane fuel cells with variable degrees of water flooding. J Power Sources 86(1–2):181–196
42. Djilali N, Lu D (2002) Influence of heat transfer on gas and water transport in fuel cell. Int J Therm Sci 41(1):29–40
43. Haddad A, Bouyekhf R, Moudni AE, Wack M (2006) Non-linear dynamic modeling of proton exchange membrane fuel cell. J Power Sources 163(1):13
44. Pukrushpan J, Stefanopoulou A, Peng H (2002) Modeling and control for PEM fuel cell stack system. Proceedings of American Control Conference. 8–10 May 2002
45. Spiegel C (2008) PEM fuel cell modeling and simulation using MATLAB. Elsevier Academic Press, Oxford UK
46. Wingelaar PJH, Duarte, JL, Hendrix MAM (2005) Dynamic characteristics of PEM fuel cells. IEEE, pp 1635–1641
47. RT-LAB 7.2, Opal-RT Technologies inc. 1751 Richardson, bureau 2525, Mon-treal Qc H3K 1G6 www.opal-rt.com
48. Vetter R, Schumacher JO (2019) Experimental parameter uncertainty in proton exchange membrane fuel cell modeling. Part II: Sensitivity analysis and importance ranking. J Power Source
49. Gao F, Blunier B, Miraoui A (2012) Proton exchange membrane fuel cells modeling. Wiley, London

Numerical Energy Applications

Theoretical Techniques for the Exploration of Piezoelectric Harvesters

Erol Kurt and Hatice Hilal Kurt

Abstract In this chapter, analytical and numerical techniques for the design and optimization of piezoelectric harvester (PH) systems are handled. In the frame of chapter, initially the approaches on how to start with an initial design will be explained. Then, the techniques to improve the starting design will be described. In the working tasks, especially the applications of finite element analysis (FEA), and time-integration schemes are focused together with the required analytical methods. There exist many tools to implement the time integration, however from the point of engineering, MatLab tool is the most preferable. In the case of FEA, mainly the Maxwell 3D package programme is explained in the present chapter. A route to get the optimized harvester devices is discussed gradually. In the time integration method, the description of dimensionless equations of motion, electricity and magnetic equations are given for the future possible applications. As the most studied sample in the literature—the cantilever structure of piezoelectric materials are considered under single well and double well potential magnetic regions. The time dependent solutions of the systems are given and possible future applications are underlined on the applications. Main interest areas of PH are medicine, automotive industry, space mission and military devices. Following the mechanical design, the ferromagnetic and non-magnetic parts should be clearly identified for the electromagnetic design if the device uses a magnetic component. Each design has its own magnetic flux path, magnetic field density values and magnetic. In order to provide a concrete device, following the analytical description, 2D or/and 3D designs should be drawn under a package programme working with FEA. Magnetostatic and magnetodynamic solutions are required in order to get the voltage and current output from the system. Besides, a time-dependent solution via a MatLab code can be applied. In this chapter, there will be different harvester systems and their analyses to give comprehensive ideas to the

E. Kurt (✉)
Department of Electrical and Electronics Engineering, Technology Faculty, Gazi University, Ankara, Turkey
e-mail: ekurt@gazi.edu.tr

H. H. Kurt
Department of Physics, Faculty of Sciences, Gazi University, Ankara, Turkey
e-mail: hkurt@gazi.edu.tr

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_18

readers for the sake of creating original harvesters. Before the practical applications on the harvesters, theoretical works on the considered harvesters give a chance to the engineers to use the budget and time efficiently.

Keywords Piezoelectric · Mode · Matlab · Harvester · Finite element · Power · Maxwell

Nomenclatures

A. Acronyms

FEA	Finite Element Analysis
FEM	Finite Element Method
PZT	Lead Zirconium Titanate
PM	Permanent magnet
PWEH	Piezoelectric wind energy harvester

B. Symbols/Parameters

B	Magnetic flux density
H	Magnetic field strength
V	Voltage
ε	Dielectric constant
P	Power
f	Frequency
d_{ij}	Tensor element
F	Force
F_m	Magnetic force
d	Distance
u	Displacement of the PZT layer
k	Force constant of the layer
m_p	Inner mass displacement of piezoelectrics
C	Capacitance of piezoelectrics
γ	Damping constant
m	Mass of knob
v	Velocity
τ	Time scale
r	Radial distance
ρ	Density
α	Force constant of piezoelectric
ω	Angular velocity
A	Sectional area
θ	Angular position
R_L	Resistive load

1 Introduction

External excitations in the forms of light, heat, fluid, motion, etc., are known as a good energy source and they can be converted into electricity in a sustainable manner. The environmental conditions of present world enforces the humanity to generate clean energy by annihilating the foot trips of pollutants in energy conversion and generation machines or machine units [1–5]. In the present era, harvesters come to the play for the clean, sustainable, practical and battery-free technological solutions [6–8].

While the energy sector deals with high power scales such as megawatts and gigawatts for intercity and international grids, there is an increasing trend to provide milliwatts and/or microwatts power scales for compact sensors, wireless sensor networks, smoke detectors, thermostats, smart light switches, other low-power equipments, etc. [2, 9, 10]. After the diversity on the technological fields, from automobile industry to space missions, harvesters have taken attention of the engineering and basic science communities. Besides, especially in the medicine sector, the energy requirements of operational robots are considered to be provided by micro-harvesters [11, 12]. Apart from the above-mentioned fields, many other sectors use the harvesters for different goals [13–17]. Some of the fields operates in remote and/or harsh isolated environments. Especially, in these situations, the workforce for a battery replacement becomes unsustainably expensive. Therefore, apart from macroscale power generation process, where the electricity is generally generated in an energy-rich location and is then distributed to end users in another location, technologies on the energy harvesters feature cable-free, battery-free and on-site electricity production. In some cases, in addition to the temporary energy storage such as capacitors, supercapacitors, fast charging/discharging batteries may also be used.

Conventionally, the energy harvesters are designed for unique type of energy source, however there are also hybrid type harvesters [18–22]. For instance, as a single energy source harvester, while a photovoltaic harvester can only harvest energy from electromagnetic waves (i.e. light), a piezoelectric harvester converts the mechanical stress to direct electricity. Similarly, an electromagnetic harvester uses the flux variation by time and converts the magnetic energy to electricity in accordance with the Faraday's Law. Besides, a thermoelectric harvester converts the thermal energy to the electricity. In many engineering applications, especially a hybrid harvester structure is made from the photovoltaic and piezoelectric structures, thereby the mechanical and electromagnetic waves are used at the same time to enable the electrical energy storage.

In Table 1, the power density of various energy sources are summarized roughly [9]. While the outdoor solar panels give better power density, the motion or acoustic noises generate fairly low power densities.

A good technique should be performed to manufacture an efficient harvester. Before any manufacture step, the algorithm in Fig. 1 should be followed. A quick sight gives us the following impression. Initially the external strength(s) should be determined for any device.

Table 1 The surface and volume power densities of various energy sources (from Ref. [9])

Energy type	Power density in order
Solar (direct/indoor)	10's mW/cm^2/10's μW/cm^2
Mechanical vibration	100's μW/cm^3
Human motion	10's–1000's μW/cm^3
Thermoelectric	10's μW/cm^2
Radio-frequency (RF)	100's μW/cm^3
Airflow	100's μW/cm^3
Acoustic noise	10's μW/cm^3

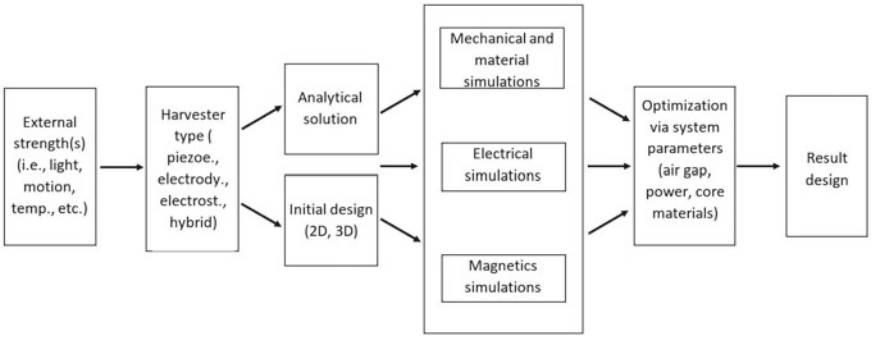


Fig. 1 An algorithm for an efficient harvester manufacture

In this phase, one should decide either a simple source device or a multi-source device will be design. Then, in the further step, harvester type should be decided. In this step, the main concern can be given to the power density. While the power densities of piezoelectric devices reach to 100 s μW/cm^2, that ratio may decrease to the quarter of that value for electrodynamic applications in practice. Following the initial design in a finite element analysis package (i.e. COMSOL Multiphysics, Maxwell 2D/3D, etc.) and an analytical confirmation of the design, one can study on the optimization of the device by adjusting different sets of system parameters.

In the end design, the most optimized device is obtained for the manufacturing process. One should also note that the external circuit load is another vital parameter, since the optimized order of power from the output terminals should be conveyed to the electrical load just for the same impedance value. For many piezoelectric layers, Kilo Ohms or Mega Ohms are required for the optimal load [23–25].

In the present chapter, a closer look on the theoretical exploration of piezoelectric based harvesters has been given. The chapter mainly focuses on the novel design, algebraic, and simulational works in space and/or time. MatLab, Maxwell 3D and COMSOL packages have been used to explore the harvester system and the optimized parameters have been defined. The chapter contains the following sections: Sect. 2 gives an introductory information on the piezoelectric materials and their electricity generation properties. Then, a detailed description on the mechanical modeling of

a piezoelectric layer structure is given in the next section. The modes and natural frequencies depending on the layer geometry are also clarified in the same section. Section 4 presents the techniques on the time-dependent solutions of the algebraic model equations. Section 5 gives the optimization process by using the algebraic formula of the harvester system. Finally, the concluding remarks are given in the conclusions section.

2 Piezoelectric Harvesters (PHs)

Piezoelectric materials are electrically charged under any physical stress in correct direction. How much they are charged is depending on applied force, geometrical features, nano-scale atomic orientation and electromechanical features of piezoelectric material. Table 2 briefly describes the properties of most common PZT materials. The energy density of the piezoelectrics are given by the multiplication of the parameters, namely piezoelectric voltage V and the strain coefficient d. Indeed, the maximum energy density depends on the square of dielectric constant (i.e. ε^2) in this regard (Table 2). Piezoelectric devices are generally used in two modes, namely 33-mode and 31-mode. The 33-mode gives the direction of the external stress and produce a certain voltage in the same direction, however, in the 31-mode, the applied stress is axial direction and that generates voltage in perpendicular direction (see in Fig. 2).

Table 2 Common piezoelectric layers and their structural properties (from Ref. [8])

Component	$\varepsilon_{33}/\varepsilon_0$	d_{33}	V_{33}	$d_{33\ x}\ V_{33}$ (m^2/N)
PZT 701	425	153	0.041	6273×10^{-15}
PZT 703	1100	340	0.03	10200×10^{-15}
PZT 502	1950	450	0.025	11250×10^{-15}
PZT 507	3900	700	0.02	14000×10^{-15}

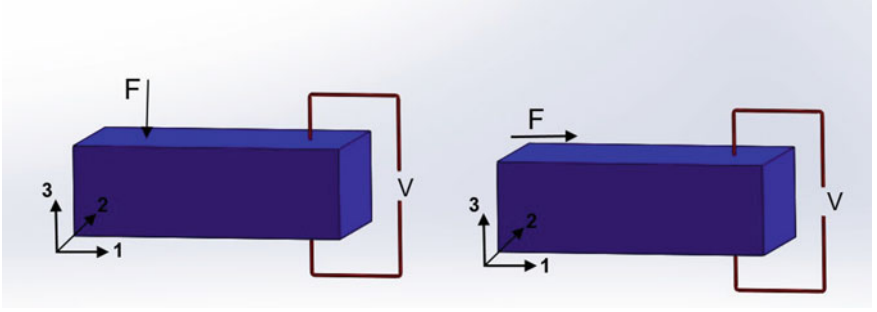


Fig. 2 Modes of piezoelectric energy harvesting

Piezoelectricity occurs from the electrical dipoles, which are naturally produced or artificially made in the crystalline or molecular structures of these special materials. When the tensors are considered as their structural characteristic features, these materials are classified into four main branches: Ceramics, polymers, composites, and single crystals. Among them, a single crystal material has positive and negative ions, which are formed in a periodic fashion inside the material, if we neglect the occasional defects occurred in the crystalline structure. One of the piezoelectric single crystal structure is called as "PMN-PT" as a solid solution. On the one hand, the Ceramics are polycrystalline samples and they are consisted of many single crystal "grains". These grains form the same chemical composition. The ions in these grains of the ceramic can orient in different directions from one another and the gaps among the ions become slightly different. In the case of Polymers, the materials are made of carbon-based structures. Indeed, they have long polymer chains and they repeat the structural units so-called "monomers". These materials exhibit more flexibility than the single crystals and ceramics, therefore they can be used for high speed rotational or linear applications [8]. In addition, for some applications, these three types of materials can also be combined together to comprise composites.

3 Finite Element Analysis Applications of PHs

Finite element analysis (FEA) is frequently used in engineering problems from mechanical issues to the electromagnetic ones. Especially, the material design of a piezoelectric layer can initially be made by mechanical FEA packages of relevant numerical codes such as Ansys Mechanics, Abacus, COMSOL Multiphysics, Solid Works, etc. Figure 3 shows such a piezoelectric layer with a meshed form. The upper and lower parts of the beam are piezoelectric material such as PZT 5 and

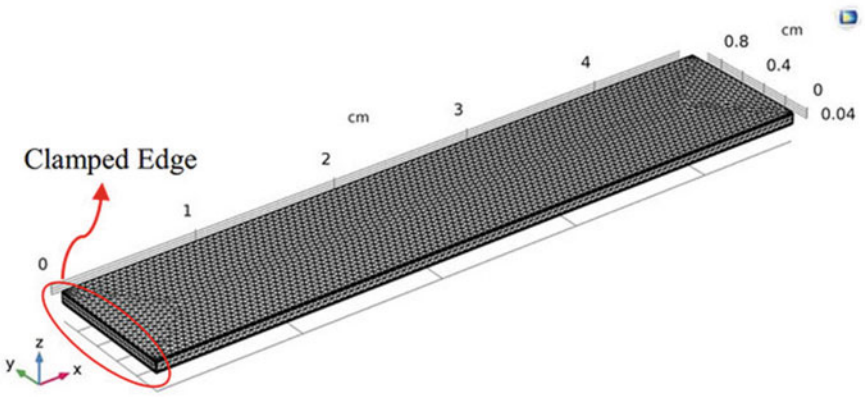


Fig. 3 A sample design for a piezoelectric cantilever with a sufficient mesh structure [26]

a copper layer is sandwiched between those [26]. Note that the solid layer structure should be characterized in terms of modes for various excitation frequencies f. Indeed, a cantilever always have a natural frequency f_0, which gives the maximal vibration amplitude at this natural frequency. At the vicinity of f_0, there is still a certain amount of amplitude to harvest energy, however if the excitation frequency f decreases/increases from f_0, the amplitude decays drastically. When the vibration amplitude decays, the buckling rate will be decreased too. Therefore, the harvested energy will be lowered.

That reality enforces one to design the optimized frequency in accordance with the excitation frequency. A sample excitation frequency information is given in Table 3. It should be noted that these f values are rough and change from one device to other or one condition to the other one.

After the drawing of the solid and meshing it with a adequate number of elements, one becomes ready for the simulations of FEA. One also should keep in mind that there should be an optimal mesh value, neither too high nor too low. Because high number of meshing elements causes to increase the computation time. That is vital to keep the computation time lower especially for industrial applications since the production scheme would not be delayed. On the one hand, determination of low mesh causes to have wrong results especially for lateral regions of the geometry.

Figure 4 represents some modal solutions from the system for frequency characteristics. The highest vibration amplitudes are obtained for lower modes. More exotic

Table 3 Frequency and acceleration values of sample systems [10, 27]

Source	Frequency (Hz)	Acceleration (m/s^2)
Human walk	2–3	0.4–0.6
Door closing on door frame	125	3
Blender	216	6.32
Microwave oven	200	1.12
Washing machine	121	3.5
Automobile	200	12
Office building ventilation	60	0.2–1.5
Refrigerator	100	0.02
Windows due to traffic noise	100	0.7
Office floor	100	0.2
Laptop	90	0.1
Vacuum cleaner	100	1.55
Laptop desktop	120	0.04
Water heater	100	0.06
Ventilating fan	214	0.29
Bridge	2–30	0.098–1.27

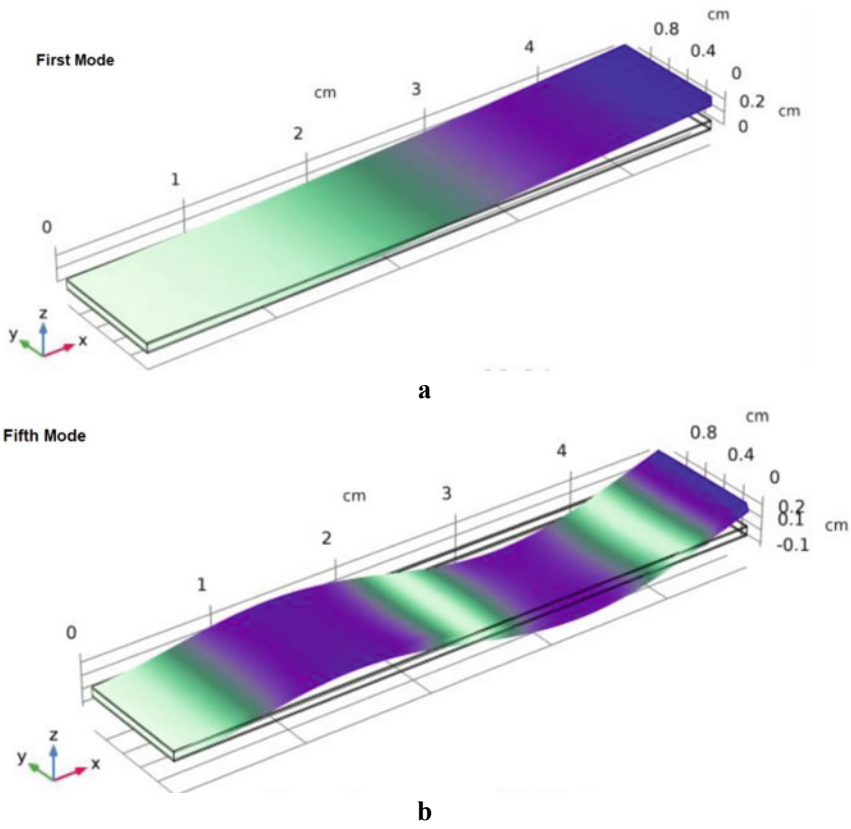


Fig. 4 Sample two modes of a cantilever beam from COMSOL Multiphysics: **a** First mode, **b** fifth mode [26]

geometries such as triangles, circles, etc., can also be designed and the eigenvalue solutions can be determined.

The cantilever beam can then be explored in terms of electrostatic and electrical circuits after the solid mechanics exploration. In that phase of the study, the output voltage from the terminals is considered. Since the solid mechanics determines the eigenvalues, these are nothing else than the optimal operation frequencies. Strictly speaking, the output voltage would be maximum at eigenfrequencies. The density of the beam varies with thickness, because there are two materials, namely a PZT material and copper. In order to consider the variation of two densities while applying a physical load in FEA code, the body load expression should be clearly determined. Indeed, the body load should be applied to every single mesh element with its particular density. Besides, isotropic loss factor can also be included as damping. For further study, the simulations should be directed to the surface of von Mises stress and the output voltages versus frequencies (i.e. eigenvalues) can be obtained. A sample result from Ref. [26] is represented in Fig. 5a, b.

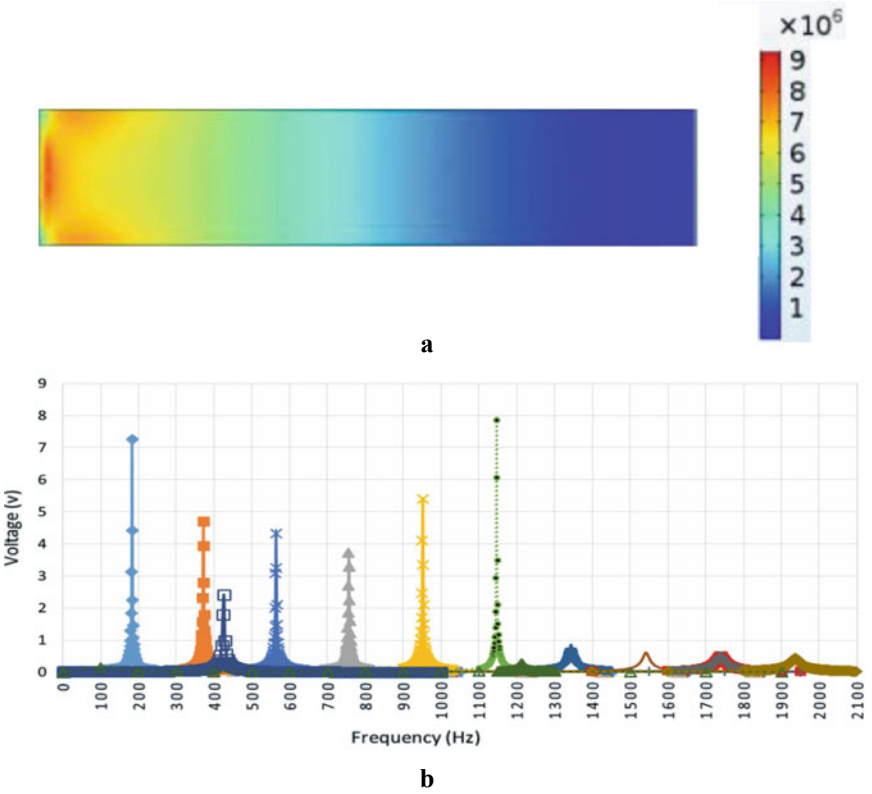
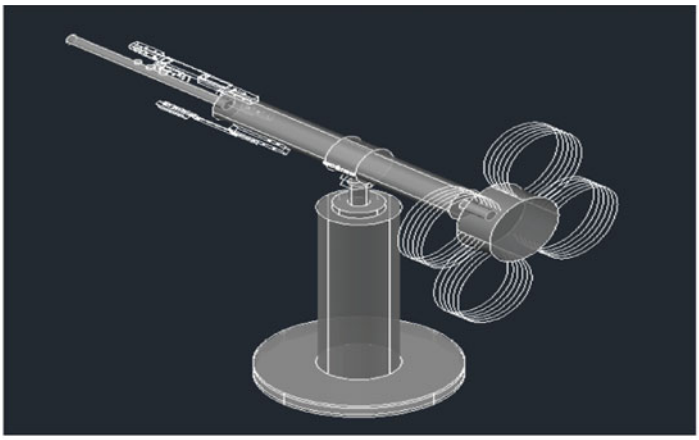


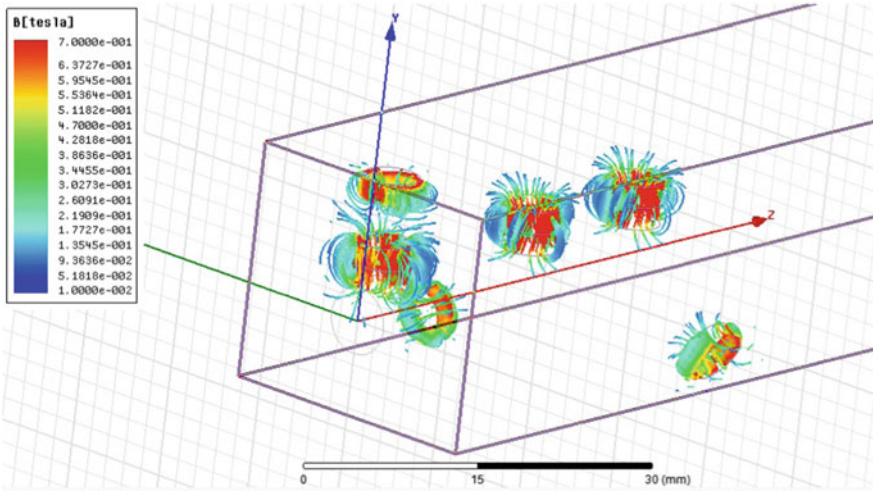
Fig. 5 **a** The von Mises stress and, **b** the voltage outputs for various frequencies. Ref. [26]

Have a much closer look at the spectrum in Fig. 5b, one can easily understand that the ratios of width to length of the beam play important role to determine the natural frequency. By calling from Ref. [26], it can be commended that while the low ratios give high voltages at low natural frequencies, the high ratios give relatively low voltages at high natural frequencies.

The piezoelectric harvesters can also be designed in complicated geometries rather than a single basic cantilever structure. Figure 6a shows such a design made in AutoCad. The triple piezoelectric cantilever system is attached by several magnets and different lengths of nonmagnetic materials in order to have optimized electrical power for various natural frequencies. The magnetostatic FEA results are presented in Fig. 6b for such a geometry. These results are taken from Ansys Maxwell 3D package and proves that magnetic field density of $B = 0.17$ T exist between the magnets at the vicinity of the tip of the harvester [28]. In addition Table 4 gives the simulated magnetic force applied to the tip of the PZT layer. These results can be obtained from FEA analyses in the magnetostatic phase.



a



b

Fig. 6 **a** AutoCad design of a triple piezoelectric wind energy harvester (PWEH) and **b** the FEA results of the tip

Table 4 Radial distance from inner magnet to the magnet at the tip of the layer

d (in radial direction as mm)	*F* (Newton/m)
5	1.35
10	0.70
15	0.40
20	0.28
25	0.20
30	0.15
35	0.1

According to the design, high magnetic densities are not desired due to the healthy vibration of piezoelectric layers. Since the magnets produce a contactless media, it gives a secure design in terms of mechanics [29].

After the FEA studies, Fig. 7 shows the tip of the manufactured PWEH. Since the lengths of the layers are different the natural frequency varies from one layer to other. That gives a superiority for this device to create a wide band operational channel for the aim of harvesting. Such an experimental result is presented in Fig. 8.

According to designed machine, there is a power harvesting capability between frequencies of 1 and 14 Hz over 40 mW, which is a superiority for such a piezoelectric

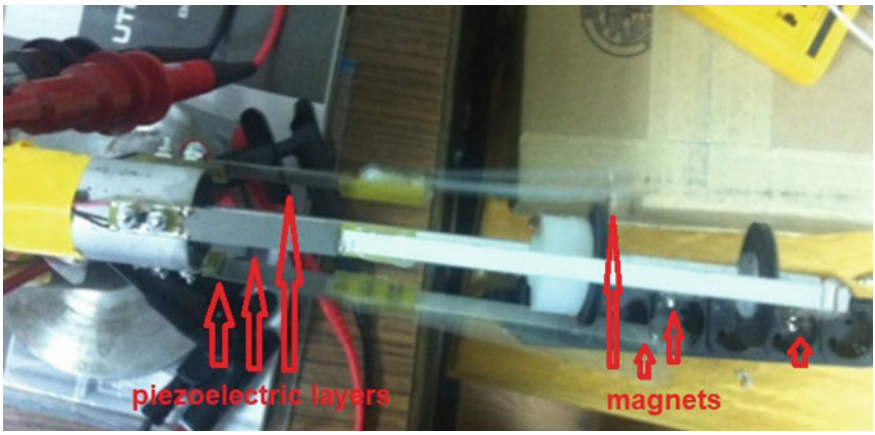


Fig. 7 The manufactured harvester device (i.e. PWEH) [29]

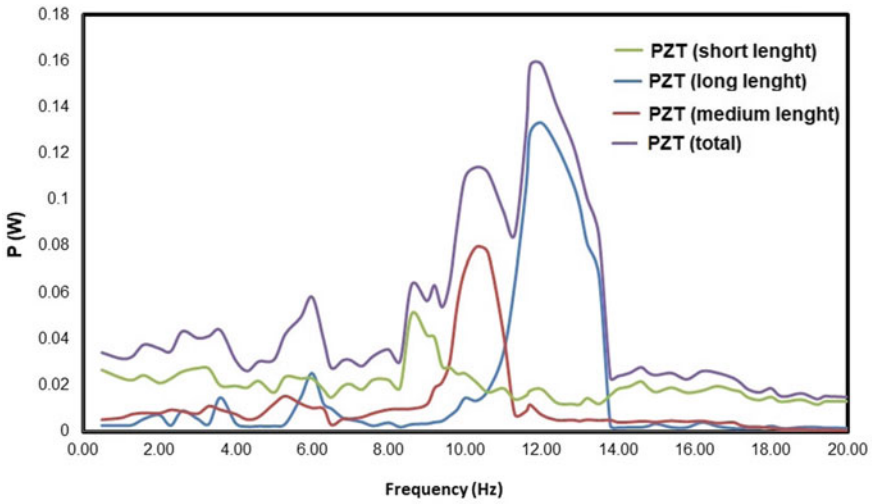


Fig. 8 The experimental wide band energy harvesting via the PWEH device

system. Note also that for a single PZT layer, one obtains only a narrow band width of 2 Hz in maximum, which is ineffective for any application. Besides, the maximum power reaches to only 140 mW for this narrow frequency gap. By using the FEA approach, one can design the full length of the PZT layer to adjust the desired natural frequency and more effective power can be obtained by eliminating the non-magnetic tip in that system. It is clear that the increasing surface yields higher power in PZT formation.

4 Time—Dependent Analyses via MatLab

Let us consider a single PZT layer attached to a stable point and a non-magnetic beam is also attached at the end of PZT layer. Such a model is given in Fig. 9. This situation indeed occurs for each layer of PWEH given in Fig. 7 The only difference in this model is that the permanent magnets are replaced by a ferromagnetic knob and an electromagnet. The response of the PZT layer for excitation frequency and amplitude is important for the harvesting capability of the energy.

According to our earlier study, the magnetic force equation can be described as follows [30]:

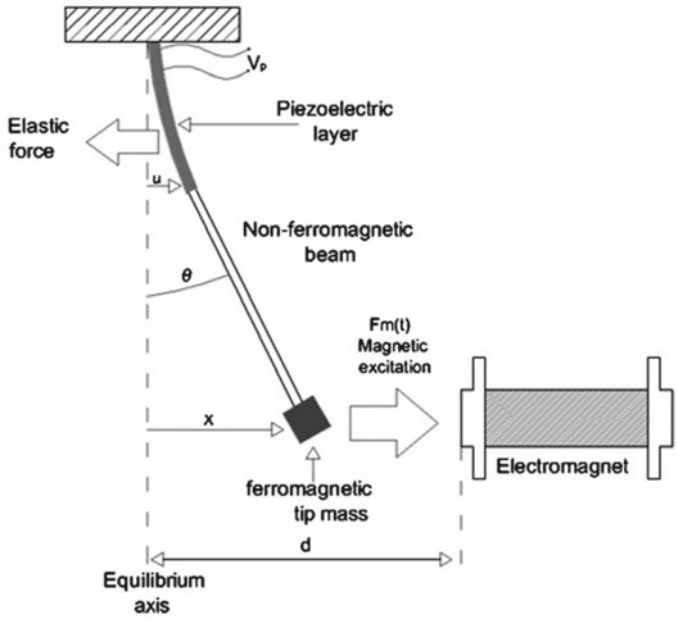


Fig. 9 A design for the exploration of a PZT layer under a periodic magnetic excitation [30]

$$F_m(t) = (1 - 0.7056/d) + 0.0623(1 - 3u) + 28026d(1 - 3u)^2 - 10^6 d^2(1 - 3u)^3, \tag{1}$$

Here, note that u denotes the displacement of the PZT layer, whereas x is nothing else than the displacement of non-magnetic knob. The numerical values in Eq. 1 are defined after a comprehensive fitting study to the experiments above [30, 31]. Therefore this force equation can be clarified after an empirical study.

The entire model of a piezoelectric layer under a periodic magnetic excitation is sketched in Fig. 10. Here k, m_p, C and γ are force constant of the layer, inner mass displacement of piezoelectrics, capacitance and damping constant, respectively. The system driven by an external periodic excitation would produce an electrical voltage v between the terminals of the layer. If the voltage/force ratio is defined by a characteristic ratio α for a PZT layer, one reads as,

$$F(t) = ku + \alpha V + F_m, \tag{2}$$

$$i = \alpha \frac{du}{dt} - c\frac{dv}{dt}, \tag{3}$$

While the first equation gives force relation, the second determines the electrical current harvested from the piezoelectric layer. By applying the masses of layer and knob as m_p and m, the set of the equations can be stated as follows:

$$\frac{du}{dt} = y, \tag{4}$$

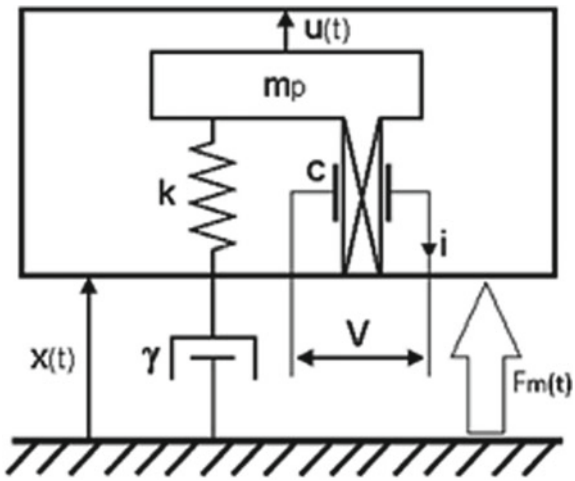


Fig. 10 Sketch of electromechanical model [2]

$$\frac{dy}{dt} = -\frac{\gamma\tau}{3m+4m_p}\,y - \frac{k\tau^2}{3m+4m_p}u - \frac{(m+m_p)\tau^2 F_m(t)}{(3m+4m_p)d} - V\frac{\alpha\tau^2 V_0}{(3m+4m_p)d}, \tag{5}$$

$$\frac{dV}{dt} = \frac{\alpha d}{CV_0}\frac{du}{dt} - \frac{\tau\ V}{V_0 C R_L}. \tag{6}$$

Note that this system of equations is a dimensionless form of a single layer problem under the periodic excitation. Here τ represents the independent variable "time". From the material catalog, one can easily learn capacitance and voltage/force ratio for any layer.

In the case of force constant, the damping curve of the system can be obtained and the constant can be easily determined form the envelope of its curve. After that, one can determine the equations of motion and the electrical equations for triple piezoelectric layers. For this, the most important parameter is the angular velocity ω, which defines the rotation speed of the mill in Fig. 7. The angular position of each magnet is then determined by θ_{n0} as in Eq. 4. Notifying the lengths L_n and the currents I_n, the entire equation system for the triple PWEH is reached [29]:

$$\frac{d\theta}{dt} = \omega, \tag{7}$$

$$L_1\frac{d^2r_1}{dt^2} = \frac{1}{\rho A}(ku_1(t) + \alpha V_1(t) + F_m\delta(\theta - \theta_{10})), \tag{8}$$

$$L_2\frac{d^2r_2}{dt^2} = \frac{1}{\rho A}(ku_2(t) + \alpha V_2(t) + F_m\delta(\theta - \theta_{20})), \tag{9}$$

$$L_3\frac{d^2r_3}{dt^2} = \frac{1}{\rho A}(ku_3(t) + \alpha V_3(t) + F_m\delta(\theta - \theta_{30})), \tag{10}$$

$$I_1(t) = \alpha\frac{du_1(t)}{dt} - C\frac{dV_1(t)}{dt}, \tag{11}$$

$$I_2(t) = \alpha\frac{du_2(t)}{dt} - C\frac{dV_2(t)}{dt}, \tag{12}$$

$$I_3(t) = \alpha\frac{du_3(t)}{dt} - C\frac{dV_3(t)}{dt}, \tag{13}$$

Here r_n is nothing else than the radial distances by time. If Eq. 3 are applied to the equations above, the resulting equation set is obtained for the numerical study in MatLab:

$$\frac{d\theta}{dt} = \omega, \tag{14}$$

$$\frac{dr_1}{dt} = y_1, \tag{15}$$

$$\frac{dr_2}{dt} = y_2, \tag{16}$$

$$\frac{dr_3}{dt} = y_3 \tag{17}$$

$$\frac{dy_1}{dt} = \frac{1}{\rho A L_1}(kr_1(t) + \alpha V_1(t) + F_m \delta(\theta - \theta_{10})) \tag{18}$$

$$\frac{dy_2}{dt} = \frac{1}{\rho A L_2}(kr_2(t) + \alpha V_2(t) + F_m \delta(\theta - \theta_{20})) \tag{19}$$

$$\frac{dy_3}{dt} = \frac{1}{\rho A L_3}(kr_3(t) + \alpha V_3(t) + F_m \delta(\theta - \theta_{30})) \tag{20}$$

$$I_1(t) = \alpha \frac{dr_1(t)}{dt} - C\frac{dV_1(t)}{dt} \tag{21}$$

$$I_2(t) = \alpha \frac{dr_2(t)}{dt} - C\frac{dV_2(t)}{dt} \tag{22}$$

$$I_3(t) = \alpha \frac{dr_3(t)}{dt} - C\frac{dV_3(t)}{dt} \tag{23}$$

The MatLab algorithm for the time integrated solution can be written as in Fig. 11. The maximum iteration number is defined by the user in accordance with the phase portrait of the solution.

Using this algorithm, the numerical solution can be performed via the MatLab code. Figure 11 shows the multiple windows structure of the findings for a sample parameter set. This sample multiple plot findings are obtained for only one PZT layer, whereas the number of windows can be tripled for the entire PWEH device. In Fig. 12, from upper left to below right, the plots of displacement—velocity phase space portray, voltage—current phase space portray, voltage—time, displacement—time, velocity—time and magnetic force—time have been presented, respectively. A more detailed representation is given in Fig. 13a–f. The units of the solutions are in mm, mm/s, Volt, and N/m for displacement, velocity, voltage, and magnetic force, respectively. Time step is 0.002 for a data point of 10^4. However, the time step and data point can be expended in accordance with the users' opinion.

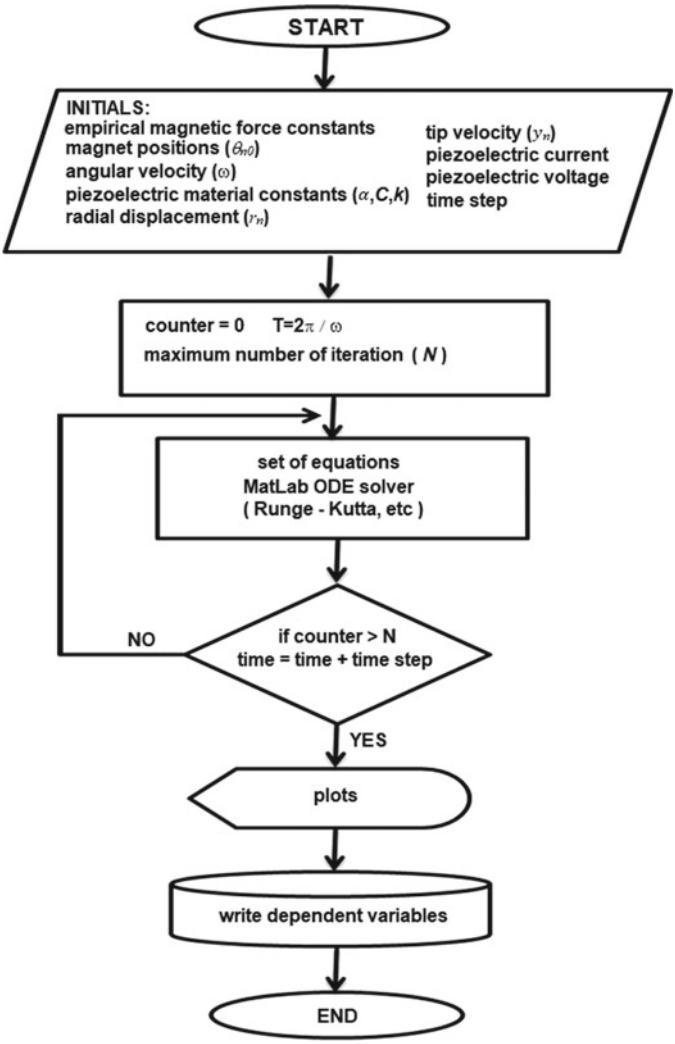


Fig. 11 MatLab solution algorithm for PWEH

5 Optimization of Design

The optimization of a designed system is one vital point. Indeed, if one does not have the optimized working conditions, the power harvested will not be high. For the optimization procedure, a detailed theoretical and experimental study should be performed. As some have also been stated in the second block from right hand side of Fig. 1, airgap, winding turn, material electrical and magnetic parameters, electrical load, core, and length of layer plays an important role. Note that the optimization

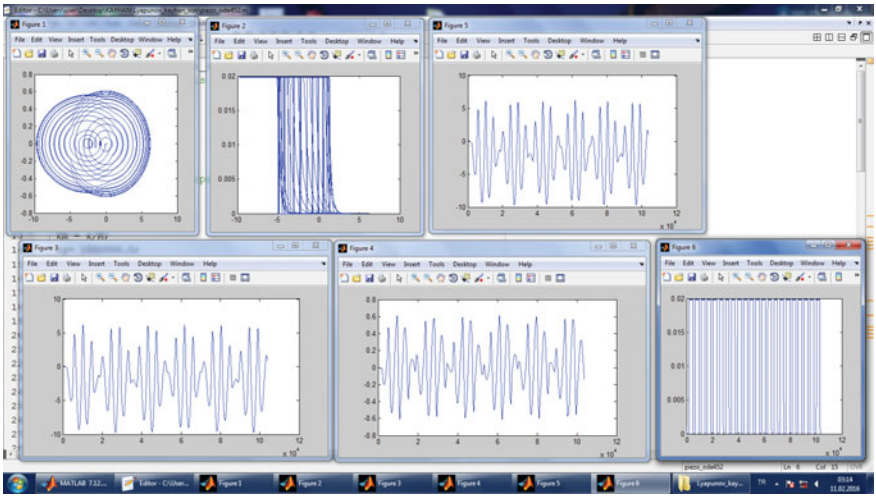


Fig. 12 The time-dependent solution of the equation set given by Eq. 5

block in Fig. 1 is just beside the manufacturing process. Therefore, one should give the required effort to determine the optimized device structure. Since all the optimization process is too comprehensive to discuss in such a short section, we will only enlighten some of the optimization schemes.

One of the optimization steps is the determination of analytical power relation of the system. Because this kind of a formulation helps us to understand the parametrical relation to access the power (i.e. P) harvested. As previously described in one of our works (i.e. Refs. [29, 32]), one can use the frequency domain in order to represent the power relation. For that, the amplitude equation below is the initial point. The voltage equations for three piezoelectric layers are given as,

$$V_1 = \frac{jR_L\alpha\omega u_1}{1 + jCR_L\omega}, \tag{24}$$

$$V_2 = \frac{jR_L\alpha\omega u_2}{1 + jCR_L\omega}, \tag{25}$$

$$V_3 = \frac{jR_L\alpha\omega u_3}{1 + jCR_L\omega}, \tag{26}$$

These equations are easy to obtain by using Eq. 5h–j in frequency domain. Note that all the parameters above are the same with the ones in Eq. 5. By also considering Eq. 5e–g, one arrives at,

$$V_1 = \frac{-\omega^2\alpha y_1 R_L\rho AL_1 - j\alpha\omega R_L F\delta(\theta - \theta_{10})}{1 + j\alpha^2 R_L\omega}, \tag{27}$$

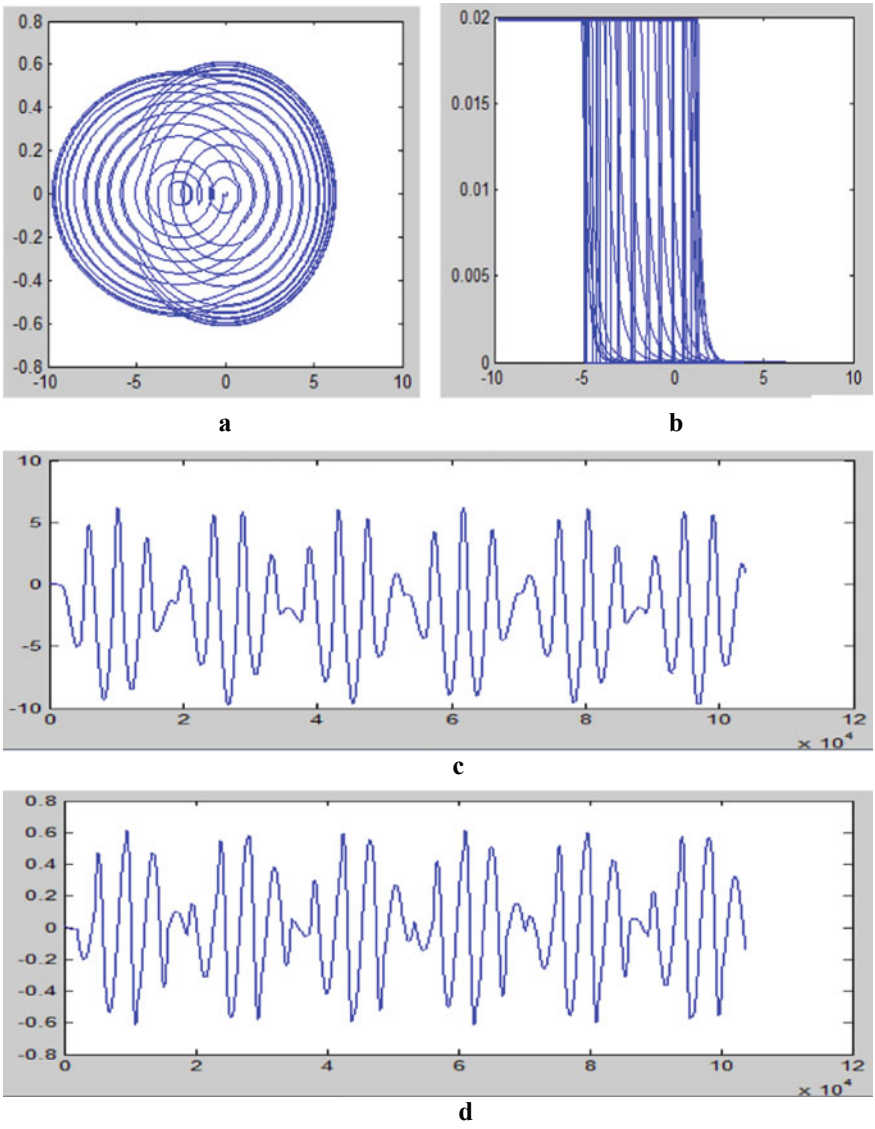
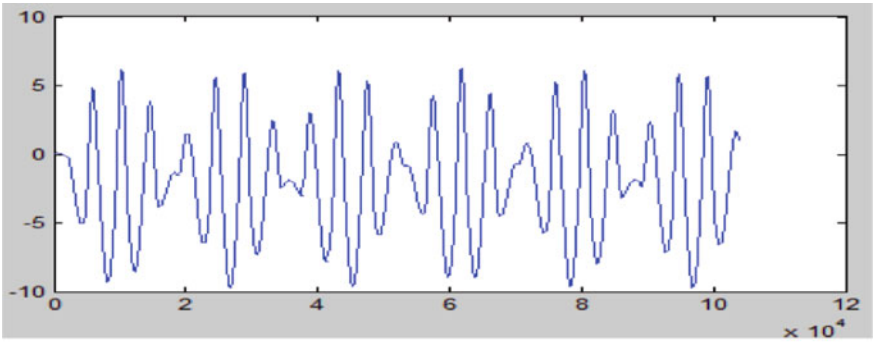
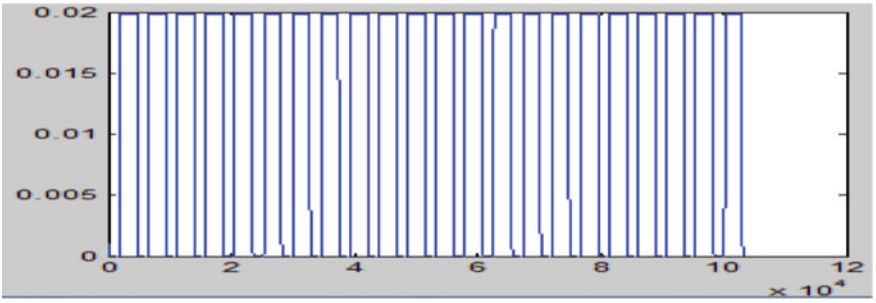


Fig. 13 A close look to MatLab solutions in Fig. 11: **a** Displacement—velocity phase space portray, **b** voltage—current phase space portray, **c** displacement—time, **d** velocity—time, **e** voltage—time and magnetic force (i.e. F_m)—time



e



f

Fig. 13 (continued)

$$V_2 = \frac{-\omega^2 \alpha y_2 R_L \rho A L_2 - j\alpha\omega R_L F \delta(\theta - \theta_{20})}{1 + j\alpha^2 R_L \omega}, \tag{28}$$

$$V_3 = \frac{-\omega^2 \alpha y_3 R_L \rho A L_3 - j\alpha\omega R_L F \delta(\theta - \theta_{30})}{1 + j\alpha^2 R_L \omega}, \tag{29}$$

Note that R_L is nothing else than the electrical load for the use of electrical power. If one considers averaged power as $\langle P \rangle = \langle V V^* \rangle / R_L$, the power—parameter relation is clearly obtained. It should be kept in mind that the sign "*" in the power relation is the conjugate. Then one arrives at,

$$\langle P_1 \rangle = \frac{\omega^4 \alpha^2 y_1^2 R_L \rho^2 A^2 L_1^2 - \alpha^2 \omega^2 R_L F^2 \delta(\theta - \theta_{10})^2}{1 + \alpha^4 R_L^2 \omega^2}, \tag{30}$$

$$\langle P_2 \rangle = \frac{\omega^4 \alpha^2 y_2^2 R_L \rho^2 A^2 L_2^2 - \alpha^2 \omega^2 R_L F^2 \delta(\theta - \theta_{20})^2}{1 + \alpha^4 R_L^2 \omega^2}, \tag{31}$$

$$\langle P_3 \rangle = \frac{\omega^4 \alpha^2 y_3^2 R_L \rho^2 A^2 L_3^2 - \alpha^2 \omega^2 R_L F^2 \delta(\theta - \theta_{30})^2}{1 + \alpha^4 R_L^2 \omega^2}, \tag{32}$$

for three layers. The total averaged power is then written as [32],

$$\langle P \rangle = \frac{\omega^4 \alpha^2 R_L \rho^2 A^2 (L_1^2 y_1^2 + L_2^2 y_2^2 + L_3^2 y_3^2)}{1 + \alpha^4 R_L^2 \omega^2} - \frac{\alpha^2 \omega^2 R_L F^2 (\delta(\theta - \theta_{10})^2 + \delta(\theta - \theta_{20})^2 + \delta(\theta - \theta_{30})^2)}{1 + \alpha^4 R_L^2 \omega^2}, \tag{33}$$

for the entire system. If the magnetic statement $F^2(\delta(\theta - \theta_{10})^2 + \delta(\theta - \theta_{20})^2 + \delta(\theta - \theta_{30})^2)$ is taken into account in parallel with the experiments for the magnetic force strength F and the magnet angular positions θ_{10}, θ_{20}, θ_{30}, as 0, 120, and 240, respectively; power depends on the square of wind speed. In addition, electrical load, force constant of piezoelectrics, and magnetic force are the main optimization parameters. It is obvious that the piezoelectric material features cannot be changed, that enforces us not to change electrical load and force constant. However other parameters such as wind speed and magnetic force can be maximized by considering structural aerodynamic designs and electromagnetic simulations. Another vital point is that one should keep in mind the durability of the materials. Since the layer will always be vibrated up and down especially in high speeds, an optimized magnetic field density is also required. Therefore, this formulation does not include all the optimization procedure.

6 Conclusions

In this chapter, the determination of a theoretical tool from the design point to the time-dependent simulations has been described. The system explained here is a piezoelectric harvester. The methodology for the design and implementation of a beam-shaped piezoelectric layers has been clarified initially. Then, the algebraic and numerical tricks for the solution of the harvester system are given. Especially, the contact-free design of the device has been introduced in such a triple layer system, which operates efficiently in a large frequency band due to different magnet masses attached to the shaft. The power relation between the device parameters has a certain superiority of the model in order to estimate the harvested power in an optimized way. To conclude, the chapter has described the Maxwell Electromagnetic 3D, MatLab, and COMSOL package applications for the relevant parts of the harvester.

Acknowledgements The present research is partially granted by The Scientific and Technological Research Council of Turkey (TUBITAK) under grant no. EEEAG-114E017.

References

1. Akizu O, Bueno G, Barcena I, Kurt E, Topaloğlu N, Lopez-Guede J (2018) Contributions of bottom-up energy transitions in Germany: a case study analysis. Energies 11(4):849. https://doi.org/10.3390/en11040849
2. Bizon N, Tabatabaei NM, Blaabjerg F, Kurt E (eds) (2017) Energy harvesting and energy efficiency: technology, methods and applications. Springer, London Limited, UK
3. Santos-Herrero JM, Lopez-Guede JM, Flores I (2017) A Short review on the use of renewable energies and model predictive control in buildings. J Energy Syst 1(3):112–120. https://doi.org/10.30521/jes.346653
4. Saracoglu BO (2017) An experimental fuzzy inference system for the third core module of the first console on the global grid peak power prediction system & its forecasting accuracy measures' comparisons with the first and the second core modules. J Energy Syst 1(2):75–101. https://doi.org/10.30521/jes.338575
5. Lee J, Yang J-S (2019) Global energy transitions and political systems. Renew Sustain Energy Rev 115:109370. https://doi.org/10.1016/j.rser.2019.109370
6. Mutlu B, Kurt E, Bizon N, Lopez Guede JM (2019) Design and fabrication of a new micro-power scaled electromagnetic harvester. J Energy Syst 3(2):51–66. https://doi.org/10.30521/jes.554900
7. Li H, Tian C, Deng ZD (2014) Energy harvesting from low frequency applications using piezoelectric materials. Appl Phys Rev 1:041301. https://doi.org/10.1063/1.4900845
8. Xu X, Cao D, Yang H, He M (2018) Application of piezoelectric transducer in energy harvesting in pavement. Int J Pave Res Technol 11:388–395. https://doi.org/10.1016/j.ijprt.2017.09.011
9. Matak M, Šolek P (2013) Harvesting the vibration energy. Am J Mechan Eng 1(7):438–442. https://doi.org/10.12691/ajme-1-7-57
10. Varadha E, Rajakumar S (2018) Performance improvement of piezoelectric materials in energy harvesting in recent days—a review. J Vibroeng 20(7):2632–2650. https://doi.org/10.21595/jve.2018.19434
11. Abdelmageed MG, Fath El-Bab AMR, Abouelsoud AA (2019) Design and simulation of pulsatile blood flow energy harvester for powering medical devices. Microelectron J 86:105–113. https://doi.org/10.1016/j.mejo.2019.02.021
12. Chorsi MT, Curry EJ, Chorsi HT, Das R, Baroody J, Purohit PK, Ilies H, Nguyen TD, Piezoelectric biomaterials for sensors and actuators. Adv Mater 31:1802084. https://doi.org/10.1002/adma.201802084
13. Pan Y, Liu F, Jiang R, Tu Z, Zuo L (2019) Modeling and onboard test of an electromagnetic energy harvester for railway cars. Appl Energy 250:568–581. https://doi.org/10.1016/j.apenergy.2019.04.182
14. Seo J, Jhang K-Y, Lee H, Kim Y-C (2019) Vibration energy harvesting technology for smart tire monitoring. J Mech Sci Technol 33(8):3725–3732. https://doi.org/10.1007/s12206-019-0714-2
15. Hande A, Shah P, Falasco JN, Weiner D (2010) Autonomous energy harvesting embedded sensors for border security applications. Proc SPIE Int Soc Opt Eng 7666:766624. https://doi.org/10.1117/12.850162
16. Koszewnik A, Oldziej D (2019) Performance assessment of an energy harvesting system located on a copter. Eur Phys J Spec Top 228:1677. https://doi.org/10.1140/epjst/e2019-800128-3
17. Song Y (2019) Finite-element implementation of piezoelectric energy harvesting system from vibrations of railway bridge. J Energy Eng 145(2):04018076. https://doi.org/10.1061/(ASCE)EY.1943-7897.0000595
18. Li Z, Li T, Yang Z, Naguib HE (2019) Toward a 0.33 W piezoelectric and electromagnetic hybrid energy harvester: design, experimental studies and self-powered applications. Appl Energy 255:113805. https://doi.org/10.1016/j.apenergy.2019.113805
19. Bolat FC, Basaran S, Sivrioglu S (2019) Piezoelectric and electromagnetic hybrid energy harvesting with low-frequency vibrations of an aerodynamic profile under the air effect. Mech Syst Signal Process 133:106246. https://doi.org/10.1016/j.ymssp.2019.106246

20. Oh Y, Kwon D-S, Eun Y, …, Kang SG, Kim J (2019) Flexible energy harvester with piezoelectric and thermoelectric hybrid mechanisms for sustainable harvesting. Int J Precis Eng Manuf Green Technol 6(4):691–698. https://doi.org/10.1007/s40684-019-00132-2
21. Rahman MS, Sarker P, Chakravarty UK (2019) A hybrid energy harvester based on solar radiation and mechanical vibration. AIP Conf Proc 2121:120008. https://doi.org/10.1063/1.5115945
22. Chamanian S, Ciftci B, Ulusan H, Muhtaroglu A, Kulah H (2019) Power-efficient hybrid energy harvesting system for harnessing ambient vibrations. IEEE Trans Circuits Syst I: Reg Papers 66(7):8667860, 2784–2793. https://doi.org/10.1109/TCSI.2019.2900574
23. Çelik E, Kurt E, Öztürk N (2019) Wind turbine speed control of a contactless piezoelectric wind energy harvester. Int J Electron. https://doi.org/10.1080/00207217.2019.1643039
24. Çelik K, Kurt E (2018) A novel meander line integrated E-shaped rectenna for energy harvesting applications. Int J RF Microwave Comput Aided Eng 29:e21627. https://doi.org/10.1002/mmce.21627
25. Uzun Y, Demirbas S, Kurt E (2014) Implementation of a new contactless piezoelectric wind energy harvester to a wireless weather station. Elektronika Ir Elektrotechnika 20(10):35–39. https://doi.org/10.5755/j01.eee.20.10.8871
26. Foroughi J, Chabok HR, Abbaspour M, Tabeshpour MR (2018) Geometry optimization of bimorph piezoelectric cantilever energy harvester for maximum output voltage. In: The 8th international conference on acoustics & vibration (ISAV2018), Shahid Beheshti University, Iran, 4–5 Dec. https://doi.org/10.13140/RG.2.2.23899.41767
27. Maurya D, Peddigari M, Kang M-G, Geng LD, Sharpes N, Annapureddy V, Palneedi H, Sriramdas R, Yan Y, Song H-C, Wang YU, Ryu J, Priya S (2018) Lead-free piezoelectric materials and composites for high power density energy harvesting. J Mater Res 33(16). https://doi.org/10.1557/jmr.2018.172
28. Kurt E, Uzun Y, Durmus C (2015) Power characteristics of a new contactless piezoelectric harvester. In: 4th international conference on electric power and energy conversion systems (EPECS), 24–26 Nov. 2015, Sharjah, UAE. https://doi.org/10.1109/EPECS.2015.7368495
29. Kurt E, Cottone F, Uzun Y, Orfei F, Mattarelli M, Özhan D (2017) Design and implementation of a new contactless triple piezoelectrics wind energy harvester. Int J Hydrogen Energy 42(28):17813–17822. https://doi.org/10.1016/j.ijhydene.2017.02.157
30. Uzun Y, Kurt E (2013) The effect of periodic magnetic force on a piezoelectric energy harvester. Sens Actuators A Phys 192:58–68. https://doi.org/10.1016/j.sna.2012.12.017
31. Uzun Y, Kurt E, Kurt HH (2015) Explorations of displacement and velocity nonlinearities and their effects to power of a magnetically-excited piezoelectric pendulum. Sens Actuators A Phys 224:119–130. https://doi.org/10.1016/j.sna.2015.01.033
32. Uzun Y, Kurt E (2012) Implementation and modeling of a piezoelastic pendulum under a harmonic magnetic excitation. In: 11th international conference on applications of electrical engineering, 7–9 March 2012, Athens, Greece, pp 184–189

Numerical Analysis of Electromagnetic Fields

Javier Bilbao, Eugenio Bravo, Olatz Garcia, Carolina Rebollar, and Concepcion Varela

Abstract Classically, the solution to contour problems in electromagnetism was based on analytical techniques, looking for closed solutions. The solution, whether computational or analytical, of electromagnetic problems is extremely important for analyzing the interactions of wave emitting and receiving devices among themselves and with their environment, including both inanimate dispersing objects and living beings. There are many applications in various areas: radio frequency antennas, radar, optics, wireless communications, imaging in bioengineering, nanotechnology and metamaterials, electrical substations, etc. Such analytical or computational solutions are particularly useful to increase productivity in all these well-established areas, to provide procedures to improve existing designs before actual implementations and to facilitate the design of new processes and devices. Typically, electromagnetism problems can be formulated using Maxwell equations. However, the Maxwell equations only admit an analytical solution for some dispersing or emitting objects with canonical geometric shapes, such as the sphere, the infinite plane, elemental antennas, etc. Numerical methods broaden the spectrum of known solutions which, while to be considered approximate, in many cases can be selected to what level of precision the calculated results describe the physical reality being analyzed. In recent decades,

J. Bilbao (✉) · E. Bravo · O. Garcia · C. Rebollar · C. Varela
Applied Mathematics Department, University of the Basque Country, Bilbao, Spain
e-mail: javier.bilbao@ehu.eus

E. Bravo
e-mail: eugenio.bravo@ehu.eus

O. Garcia
e-mail: olatz.garcia@ehu.eus

C. Rebollar
e-mail: carolina.rebollar@ehu.eus

C. Varela
e-mail: concepcion.varela@ehu.eus

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_19

driven by the availability of increasingly powerful computers, the area of computational electromagnetics (CEM) has experienced a remarkable increment as an area of research. Mathematical formulations of physical electromagnetic problems produce systems of equations that can now be solved numerically by computers. Thanks to advances in computational technology and increasingly sophisticated mathematical algorithms of electromagnetic modeling, it is a reality to simulate radiation or scattering problems containing arbitrary and complex structures for which there is no analytical solution to the Maxwell equations. There are various methods of computational electromagnetism and various classifications. Depending on the geometric model used by their formulations to characterize the dispersers, they can be classified into three types: ray tracing, surface discretizations, and volume discretizations. Depending on the precision achieved in the results and the field of application, they are classified into full-wave and asymptotic methods, also called low and high-frequency methods. Methods based on volumetric discretizations, such as finite-difference time-domain (FDTD) and frequency domain finite-element method (FEM), have the advantages that they allow for easy modeling of non-homogeneous media, and their associated 3D mathematical formulations are relatively simple. However, they suffer from the fact that the resulting system of linear equations has a number of unknowns proportional to the simulated volume, so the computational demand grows very rapidly as the electrical dimensions considered in the simulation increase. The methods based on discretizations of surfaces present characteristics that make them computationally more efficient than the volumetric ones. The formulations used in surface methods are based on surface integral equations (SIE) which, unlike volumetric formulations, are mathematically more difficult to implement in a computational code, partly due to the various types of singularities of the Green function. Another disadvantage of this type of methods is the impossibility of simulating general non-homogeneous means, although they have the great advantage that they only require discretizing the interfaces, that is to say, the two-dimensional surfaces that delimit the dispersing objects. Among the surface methods, the method-of-moments (MoM) and its computational optimizations stand out, in exchange for introducing a controllable numerical error on the results of the pure MoM, known as fast multipole method (FMM) and multilevel fast multipole algorithm (MLFMA). The physical optics (PO) is also considered as a surface method based on SIEs since it is based on surface discretizations, although using approximations valid only for electrically large objects. The PO supports a correction method to include diffraction, called physical theory of diffraction (PTD), although this correction is only applicable to perfect electric conductors (PEC). In this chapter, we will analyze some of the numerical methods used in electromagnetism.

Keywords Numerical analysis · Electromagnetic fields · Electromagnetism · Computational electromagnetics · Surface integral equations

Abbreviation/Acronyms

CEM	Computational electro magnetics
CFIE	Combined field integral equation
EFIE	Electric field integral equation
FDTD	Finite-difference time-domain
FE	Finite elements
FEM	Finite-element method
FMM	Fast multipole method
GMRES	Generalized minimal residual algorithm
GTD	Geometrical theory of diffraction
HF	High frequency
MECA	Modified equivalent current approximation
MFIE	Magnetic field integral equation
MLFMA	Multi level fast multipole algorithm
MoM	Method of moments
PEC	Perfect electric conductors
PO	Physical optics
PTD	Physical theory of diffraction
RWG	Rao Wilton Glisson
SIE	Surface integral equations
TE	Transverse-electric wave
TM	Transverse-magnetic wave
ε	Permittivity
σ	Conductivity
μ	Permeability
$\vec{E}$	Electric field
$\vec{H}$	Magnetic field
$\vec{J}$	Electric surface current
$\vec{M}$	Magnetic surface current
η	Impedance
L, K	Integral-differential operators
$G(\vec{r}, \vec{r}')$	Green scalar function

1 Introduction

Electromagnetic problems have a multitude of applications in our lives, in industry and particularly in engineering and they can be very different in nature. This wide range of possibilities in which to apply computational electromagnetics (CEM) and the great variety of types of problems that we can find have led scientists to create a large number of different algorithms to deal with this type of problems. However, nowadays, there is no algorithm that stands out from the rest for any situation and

problem that we are going to face. In other words, there are algorithms that are more suitable for one type of problem or conditioning and others are more suitable for another type [1–3].

The classification of these algorithms is usually done into the low frequency algorithms (or accurate algorithms) and high frequency algorithms (or approximate algorithms). Often, electromagnetic problems are also classified on the basis of the working domain: time domain or frequency domain.

Briefly, the main or most used algorithms to solve electromagnetic problems are the following [4], remembering that they are not the only ones and that they can exist, and exist, other algorithms that for certain particular problems can have an advantage of calculation on the ones mentioned here.

1.1 Low Frequency Methods

Some algorithms solve Maxwell's equations without hidden approximations and are generally applied to small electrical problems due to calculation times and system memory limitations: these algorithms are the low frequency methods. Although computers are becoming increasingly powerful and solving more and more problems, it is likely that this concept of limitation, related to the computers, can become obsolete in the medium future.

Within this type of methods, we will cite the three most used, without the order in which they appear presupposes their better or worse applicability.

1.1.1 The Finite Difference Time Domain Method

The Finite Difference Time Domain method (FDTD) uses the finite difference method in order to solve Maxwell's equations in the time domain. The implementation of the FDTD method is generally quite simple [5–8]: a solution domain is usually subdivided in small rectangular or curvilinear elements, with a "jump" in the time used to calculate the electric and magnetic fields.

FDTD works normally very well in the analysis of non-homogeneous and nonlinear media, but it requires very high quantity of dedicated memory in the computer. It is due to the discretization process of the solution of the entire domain. Usually, it is not recommended for dispersion or scattering problems. FDTD is used in waveguide packaging techniques and issues, as well as in wave propagation studies.

1.1.2 Finite Element Method

The Finite Element Method (FEM) is a method used to solve the problems of electromagnetic, with boundary values, in the frequency domain [9–12]. As FDTD method, it tries to solve Maxwell's equations in a differential way.

Although the name of the FEM has been established in the last decades of the last century, the concept has been used for several centuries. The use of temporal and spatial discretizing methods, and also numerical approximation procedures, to obtain solutions to engineering or physical problems has been known since ancient times. The concept of "finite elements" is based on this idea.

The development of finite elements as they are known today has been linked to structural calculation primarily in the aerospace field. In the 1940s, Courant [13] proposed the use of polynomial functions for the formulation of elastic problems in triangular sub-regions, as a special method of the Rayleigh-Ritz variational method to approximate solutions. It was Turner, Clough, Martin and Topp [14] who presented FEM in the form accepted today. In their work, they introduced the implementation of simple finite elements (bars and triangular plates with loads in their plane) to the analysis of aeronautical structures [15], using the concepts of discretization and functions of form.

The books by Przemieniecki [16] and Zienkiewicz and Holister [17] present the MEF in its application to structural analysis. The book by Zienkiewicz and Cheung [18] or Zienkiewicz and Taylor [19] presents a broad interpretation of FEM and its application to any field problem. It demonstrates that FE equations can be obtained using a residual weight approximation method, such as the Galerkin method or the least squares method.

It is considered as a frequency domain algorithm.

1.1.3 Method of Moments

The publication in 1968 of the work "Field Computation by Moment Methods" [20], by Harrington, allowed the systematic formulation of the existing numerical methods by means of a very general concept denoted by the Method of Moments. The method of moments is one of the most widely used numerical techniques today to determine the fields emitted or received by radiant structures.

The method of moments allows the solution of the problem of Poisson in its integral version and, in particular, to find the distribution of load on the surface of a system of conductors, known the potential to which each one of them is found. From the load distribution is obtained directly, the field and potential at any point in space.

The Method of Moments (MoM) is a technique used in the frequency domain.

1.2 High Frequency Methods

Large electromagnetic problems have set out long before the existence of computers and also some one or two decades ago, when these machines currently could not solve them. Common examples of larger problems are the prediction of the radar cross section and the calculation of the radiation pattern of an antenna when mounted on a large structure (typical use for telephony). Many approximations have been made to

the radiation and scattering equations to make these problems manageable. Most of these treat the fields at the asymptotic or high frequency (HF) boundary and employ ray optics and edge diffraction. When the problem is very large from the electrical point of view, many asymptotic methods produce results that are sufficiently accurate by themselves or can be used as a "first or previous step" before applying a more precise but computationally demanding method.

1.2.1 Theory of Geometric Diffraction and Physical Theory of Diffraction

One of the first methods for the calculation of electromagnetic diffraction was the Geometrical Theory of Diffraction (GTD) [21, 22] introduced by Keller. This method is also based on ray tracing, such as geometrical optics, but introduces diffracted rays at the edges.

When a ray hits a conductive wedge, a diffraction is observed forming a reflection angle with the same edge as the incidence. In this case, unlike what happens in flat surfaces where there is only one direction of reflection, infinite directions are observed that form with the edge an angle equal to the angle of incidence.

These directions form the so-called Keller cone [23].

The physical theory of diffraction (PTD) [24, 25], developed in parallel with Keller's theory, obtains equivalent results avoiding some problems. The result of PTD is finite and contains only the diffraction of the edge, so the fields reflected in the superficies, calculated for example by the approximation of Physical Optics (PO), must be added to it [26]. The physical theory of diffraction (PTD) is a means of complementing the PO solution by adding the effects of non-uniform currents at the diffraction edges of an object [27, 28]. PTD is commonly used in high-frequency radar cross section and scatter analysis [29].

1.2.2 Physical Optics

The method of Physical Optics (PO) consists in the fact that the currents induced on the parts of the object not illuminated, for example by a radar, are very small compared with those produced in the illuminated areas [30, 31]. In fact, the approximation made by this method consists in annulling them. For the other surfaces, the calculation will be made by obtaining the equivalent currents that would exist in a tangent infinite plane at each point of the surface.

1.2.3 Shooting and Bouncing Beams

The ray tracing is a method similar to the previous one that also takes into account the possible multiple reflections of the field reflected by the object, but its complexity is greater [32, 33].

The first reflection of the form commented in the PO method is obtained, later it is calculated if the direction of the reflection returns to intersect with some surface of the object. If so, the diffracted field that produces this new reflection is recalculated.

The calculation ends when there are no more reflections accumulating the value of all reflections. To carry it out, a large number of rays are released. These are reflected in the object according to geometric optics (Snell Law). Finally, the contributions of the rays that return to the initial position are added [34, 35].

2 Surface Integral Equations

We have focused on computational electromagnetics (CEM) based on surface integral equations (SIEs) because it provides great versatility when analyzing homogeneous and isotropic objects that occupy electrically large volumes [36]. In addition, these methods have been developed mainly since 1990.

The present chapter deals with different computational methods based on SIEs applicable to homogeneous and isotropic general media, i.e. not limited only to perfect electrical conductors (PEC) as is the case in an important part of the previous literature on EMF. Physical optics (PO) belongs to the so-called high frequency methods and allows fast predictions with a limited level of accuracy proportional to the electrical size of the objects. On the contrary, the so-called full wave methods such as the method-of-moments (MoM) allow very precise predictions of dispersed field, but their computational cost makes their application in volumes whose electrical sizes extend several wavelengths totally unfeasible. Among the techniques for accelerating MoM, the fast multipole method (FMM) [37] and its multilevel extension, based on a hierarchical multilevel partition of geometries, known as MLFMA (multilevel fast multipole algorithm) [38], stand out. In this chapter we have opted for the development of both full wave techniques (MoM, FMM, MLFMA) and high frequency (PO for penetrable media). Even hybridization between full wave and high frequency techniques could be implemented in other fields, such as radiation [39] and wave propagation [40].

This fully realistic approach, not limited to a few CEM techniques, is justified by the fact that many real problems are not fully addressable—not even using current supercomputers—but by high-frequency techniques, or by hybridizations where these operate.

3 Method of Moments

The Method of Moments (MoM) is a full wave method, introduced in CEM by R. F. Harrington in 1967. His book, currently re-published by the IEEE [20], remains a fundamental reference. The complexity of this method is of the order O(N2) in

memory and O(N3) in time in case of using to solve the system of resulting equations a direct method such as LU decomposition.

We remember that the order of a method is given by the number p which means the number of terms used in the weighted average used in that method.

Formally, the definition of the order of a method can be enunciated as follows:

Let's be p ≥ 0. It is said that a method is of order ≥p if for all sufficiently regular solution of a problem of initial values, we have

$$\max_{0 \le n \le N-1} |\sigma_n| \le C\, h^p$$

for some constant C (where C can depend on the solution x). Note that if a method has order $\geq p$ with $p > 0$, then it is consistent.

This last complexity can easily be reduced to O(N2) by replacing the direct method with an iterative method, according to Kim et al. [41], as the Generalized Minimal Residual algorithm, GMRES method [42].

MoM is a numerical method that allows solving a discretization of a surface integral equation (SIE). As a previous step to explain MoM's own discretization, we will briefly introduce the concept of SIE for a single dispersing object.

3.1 Using Surface Integral Equations for MoM

Let's be a penetrable homogeneous dispersant surrounded by an unlimited homogeneous medium. Let us denote by R_1 the region corresponding to the unlimited medium from which the incident wave proceeds. Next, we will denote by R_2 the limited region related to the dispersing object. From now on, we will associate a subscript $i = 1$ for all the quantities related to R_1 and another subscript $i = 2$ for the amounts of R_2. Each medium, for $i = 1, 2$, is characterized by its constitutive parameters, that we can resume as the following: the complex permittivity $\varepsilon_i = \varepsilon_{r,i} \cdot \varepsilon_0$ (which includes the effects of conductivity σ_i) and the complex permeability $\mu_i = \mu_{r,i} \cdot \mu_0$. Respectively, $\varepsilon_{r,i} \in \mathbb{C}$ and $\mu_{r,i} \in \mathbb{C}$ are the complex relative permittivity and the complex relative permeability of the medium in the region R_1. ε_0 and μ_0 are the constitutive parameters of the vacuum. We assume a harmonic time dependence e^{jwt} that we will omit in the use of the SIE. The process starts with an incident field ($\vec{E}_{inc}$, $\vec{H}_{inc}$), and we want to calculate a scattered field ($\vec{E}_{1,scatt}$, $\vec{H}_{1,scatt}$) for the region R_1 external to the disperser, and another scattered field ($\vec{E}_{2,scatt}$, $\vec{H}_{2,scatt}$) for the internal region R_2.

Applying the first principle of equivalence, or Love's equivalence principle (according to Medgyesi-Mitschang et al. [43]), it is possible to formulate an equivalent problem in R_1, where we will have $\vec{E}_{2,scatt} = 0$ and $\vec{H}_{2,scatt} = 0$ and it is necessary to impose electric surface currents $\vec{J}_1 = \hat{n}_1 \times \vec{H}_{1,scatt}\Big|_S$ and magnetic surface currents $\vec{M}_1 = -\hat{n}_1 \times \vec{E}_{1,scatt}\Big|_S$ on the surface S of the dispersant. This is

an equivalent problem, valid only for region R_1. In the same way, it is possible to set out an equivalent problem for R_2.

In order to deduce the generic form of an SIE, we start from the Maxwell equations for $\nabla \times \vec{E}_{1,scatt}$ and for $\nabla \times \vec{H}_{1,scatt}$, and we apply again the nabla operator in order to get two new equations of the following form: $\nabla \times \nabla \times \vec{E}_{1,scatt}$, and $\nabla \times \nabla \times \vec{H}_{1,scatt}$. Finally, we use mathematical tools. First, the vector Green theorem in these two equations on the surface S of the disperser (Poggio and Miller developed this vector theorem in a rigorous description [44]). Second, together with the previous theorem, these following four equals (two equals for fields and other two equals for currents), deduced, on the one hand, with the currents of the Love principle and, on the other hand, imposing tangential contour conditions:

$$\begin{aligned} \hat{n}_1 \times (\vec{E}_{1,scatt} - \vec{E}_{2,scatt})\Big|_S = 0 \quad &\Rightarrow \quad \underbrace{-\hat{n}_1 \times \vec{E}_{1,scatt}\Big|_S}_{\vec{M}_1} = \underbrace{\hat{n}_2 \times \vec{E}_{2,scatt}\Big|_S}_{-\vec{M}_2} \\ \hat{n}_1 \times (\vec{H}_{1,scatt} - \vec{H}_{2,scatt})\Big|_S = 0 \quad &\Rightarrow \quad \underbrace{\hat{n}_1 \times \vec{H}_{1,scatt}\Big|_S}_{\vec{J}_1} = \underbrace{-\hat{n}_2 \times \vec{H}_{2,scatt}\Big|_S}_{-\vec{J}_2} \end{aligned} \tag{1}$$

If the reader wants to deepen in the mathematical development of the previous procedure, he or she can find more information in the references of Medgyesi-Mitschang [43] and Ylä-Oijala et al. [45].

This procedure produces a tangential integral equation for the electric field and for each medium (T-EFIE, tangential electric field integral equation) and, by applying a cross product by the normal towards R_1, a normal integral equation for the electric field and for each medium (N-EFIE, normal electric field integral equation). Similarly, we can obtain the T-MFIE and the N-MFIE for the magnetic field. The eight tangential (T) and normal (N) equations of EFIE and MFIE for each medium depend on the surface currents, normal currents and the incident field, and are summarized as the following:

T-EFIE$_1$, medium 1:

T-EFIE$_2$, medium 2:

$$\vec{E}^{inc}(\vec{r})\Big|_{\tan} = L_1 \cdot \vec{J}(\vec{r})\Big|_{\tan} - K_1 \cdot \vec{M}(\vec{r})\Big|_{\tan} - \frac{1}{2}\vec{M}(\vec{r}) \times \hat{n}(\vec{r}) \tag{2}$$

$$\vec{0} = L_2 \cdot \vec{J}(\vec{r})\Big|_{\tan} - K_2 \cdot \vec{M}(\vec{r})\Big|_{\tan} + \frac{1}{2}\vec{M}(\vec{r}) \times \hat{n}(\vec{r}) \tag{3}$$

T-MFIE$_1$, medium 1:

T-MFIE$_2$, medium 2:

$$\vec{H}^{inc}(\vec{r})\Big|_{\tan} = K_1 \cdot \vec{J}(\vec{r})\Big|_{\tan} + \frac{1}{\eta_1^2} L_1 \cdot \vec{M}(\vec{r})\Big|_{\tan} + \frac{1}{2}\vec{J}(\vec{r}) \times \hat{n}(\vec{r}) \tag{4}$$

$$\vec{0} = K_2 \cdot \vec{J}(\vec{r})\Big|_{\tan} + \frac{1}{\eta_2^2} K_2 \cdot \vec{M}(\vec{r})\Big|_{\tan} - \frac{1}{2}\vec{J}(\vec{r}) \times \hat{n}(\vec{r}) \tag{5}$$

N-EFIE$_1$, medium 1:
N-EFIE$_2$, medium 2:

$$\hat{n}(\vec{r}) \times \vec{E}^{inc}(\vec{r}) = \hat{n}(\vec{r}) \times \left[L_1 \cdot \vec{J}(\vec{r}) - K_1 \cdot \vec{M}(\vec{r}) \right] - \frac{1}{2}\vec{M}(\vec{r}) \tag{6}$$

$$\vec{0} = \hat{n}(\vec{r}) \times \left[L_2 \cdot \vec{J}(\vec{r}) - K_2 \cdot \vec{M}(\vec{r}) \right] + \frac{1}{2}\vec{M}(\vec{r}) \tag{7}$$

N-MFIE$_1$, medium 1:
N-MFIE$_2$, medium 2:

$$\hat{n}(\vec{r}) \times \vec{H}^{inc}(\vec{r}) = \hat{n}(\vec{r}) \times \left[K_1 \cdot \vec{J}(\vec{r}) + \frac{1}{\eta_1^2} L_1 \cdot \vec{M}(\vec{r}) \right] + \frac{1}{2}\vec{J}(\vec{r}) \tag{8}$$

$$\vec{0} = \hat{n}(\vec{r}) \times \left[K_2 \cdot \vec{J}(\vec{r}) + \frac{1}{\eta_2^2} L_2 \cdot \vec{M}(\vec{r}) \right] - \frac{1}{2}\vec{J}(\vec{r}) \tag{9}$$

In these eight equations, $\vec{J}(\vec{r}) \equiv \vec{J}_1$ and $\vec{M}(\vec{r}) \equiv \vec{M}_1$ denote the surface equivalent currents for region R_1, a priori unknown. $\vec{J}(\vec{r})$ and $\vec{M}(\vec{r})$ are vector functions of an arbitrary point $\vec{r}$ on the surface of the dispersant. The vector $\hat{n}(\vec{r}) \equiv \hat{n}_1$ corresponds to the normal to the surface pointing to the outer region R_1. Moreover, $\eta_i = \left(\mu_i / \varepsilon_i\right)^{1/2}$ is the impedance intrinsic of the medium in R_i. The integral-differential operators L_i and K_i in these equations are defined as:

$$L_i \cdot \vec{X}(\vec{r}) = \iint_S \left[jw\mu_i \cdot \vec{X}(\vec{r}') + \frac{j}{w\varepsilon_i} \nabla \left(\nabla' \cdot \vec{X}(\vec{r}') \right) \right] \cdot G_i(\vec{r}, \vec{r}') \cdot dS'$$

$$K_i \cdot \vec{X}(\vec{r}) = P.V. \iint_S \vec{X}(\vec{r}') \times \nabla G_i(\vec{r}, \vec{r}') \cdot dS' \tag{10}$$

The *P.V.* notation is used in the definition of the K_i operator to indicate that the integration is taken as a Cauchy principal value integral (that is, when we have an improper integral, normally, a contour integral of a complex-valued function). The integration surface S alludes to the separation interface between R_1 and R_2. The term in $G_i(\vec{r}, \vec{r}')$ in the Eq. (10) is the Green scalar function for the region R_i, defined as:

$$G_i(\vec{r}, \vec{r}') = \frac{e^{(-jk_i|\vec{r}-\vec{r}'|)}}{4\pi|\vec{r} - \vec{r}'|} \tag{11}$$

In the last two equations, (10) and (11), the vector $\vec{r}'$ refers to a source point and $\vec{r}$ denotes an observation point. In addition, $k_i = w(\varepsilon_i \mu_i)^{1/2}$ is the wave number in R_i.

A generic SIE formulation can be established by combining different EFIEs and MFIEs in (2)–(9).

Table 1 Parameters to obtain the SIE formulations

SIE	a_i for $i = 1, 2$	b_i for $i = 1, 2$	c_i for $i = 1, 2$	d_i for $i = 1, 2$
PMCHWT	η_i	0	0	$1/\eta_i$
JMCFIE	1	1	1	1
CTF	1	0	0	1
CNF	0	1	1	0
MNMF	0	$\mu_i/(\mu_1 + \mu_2)$	$\varepsilon_i/(\varepsilon_1 + \varepsilon_2)$	0

We perform a general combination of these eight equations using the same sign criteria as Ergül in [46]:

$$\begin{aligned} &\frac{a_1}{\eta_1}(T - EFIE_1) + \frac{a_2}{\eta_2}(T - EFIE_2) + \\ &\quad + b_1(N - MFIE_1) - b_2(N - MFIE_2) = \vec{0} \\ &\quad - c_1(N - EFIE_1) + c_2(N - EFIE_2) + \\ &\quad + d_1\eta_1(T - MFIE_1) + d_2\eta_2(T - MFIE_2) = \vec{0} \end{aligned} \tag{12}$$

There are infinite values that can be assigned to the complex scalar parameters $a_i, b_i, c_i, d_i \in \mathbb{C}$ for $i = 1, 2$ in order to obtain valid and stable formulations. The parameters in (12) that allow to obtain some types of SIEs can be found in Table 1.

The SIE formulations in the table are well known and well reported for the case of dispersers isolated in free space. These formulations are known as Poggio-Miller-Chang-Harrington-Wu-Tsai (called by the acronym PMCHWT) [44, 47–49], combined tangential formulation (CTF) [42, 46], combined normal formulation (CNF) [45, 49], modified normal Müller formulation (MNMF) [45], and electric and magnetic current combined-field integral equation (JMCFIE) [50–54]. Other references incorporating other known stable formulations are [55, 56].

3.2 Discretization of SIE Formulations

To resolve such a SIE formulation of the form (12), the current densities $\vec{J}(\vec{r})$ and $\vec{M}(\vec{r})$, in our case unknown variables, are approximated in terms of linear combinations of known vector base functions $\vec{f}_n$, with $n = 1, 2, \ldots, N$, such as

$$\vec{J}(\vec{r}) = \sum_{n=1}^{N} J_n \cdot \vec{f}_n \,, \quad \vec{M}(\vec{r}) = \sum_{n=1}^{N} M_n \cdot \vec{f}_n \tag{13}$$

where J_n and M_n, with $n = 1, 2, \ldots, N$, are the unknown complex coefficients in the expansions of the Eq. (13). These coefficients are the unknowns determined in MoM

by solving a system of linear equations. Usually, in MoM we choose some functions $\vec{f}_n$, with $n = 1, 2, \ldots, N$, known as RWG bases (Rao-Wilton-Glisson) [57, 58], due basically to their simplicity to implement them in code.

From this point forward, we will always use the base functions of type RWG to represent $\vec{f}_n(\vec{r}') \equiv \vec{f}_n$. Each RWG base is associated with the side $n = 1, 2, 3$ of a triangle and is defined for the two triangles of the discretization that share that side. One of the triangles will be assigned with a " + " label and a "−"label will be assigned to the other side/triangle. These RWG functions fulfill the following:

$$\vec{f}_n(\vec{r}') = \frac{\vec{\rho}'^{\pm} l_n}{2A_n^{\pm}} \text{ and } \nabla' \cdot \vec{f}_n(\vec{r}') = \pm \frac{l_n}{A_n^{\pm}} \tag{14}$$

where the value A_n^+ (and/or A_n^-) denotes the area of the triangle + (or −), and l_n is the length of the side. In addition, $\vec{\rho}'^{+} = \vec{r}' - \vec{p}^{+}$ ($\vec{\rho}'^{-} = \vec{p}^{-} - \vec{r}'$) represents the vector that joins the node $\vec{p}^{+}$ ($\vec{p}^{-}$) opposite to the considered side of the triangle up to (from) the source point $\vec{r}'$ on the triangle + (or −).

Applying the so-called Galerkin procedure, each side of the triangle is assigned a weighting function (also called a test function) denoted as $\vec{f}_m(\vec{r}) \equiv \vec{f}_m$, which has the same vector expression as the corresponding base function $\vec{f}_n(\vec{r}') \equiv \vec{f}_n$ assigned to the same side. For simplicity, the weighting operation, or test operation, using a function $\vec{f}_m(\vec{r}) \equiv \vec{f}_m$ to weight a generic vector $\vec{v}_n(\vec{r})$, we will denote as:

$$\left\langle \vec{f}_m, \vec{v}_n \right\rangle = \iint_{S_m} \vec{f}_m(\vec{r}) \cdot \vec{v}_n(\vec{r}') \cdot dS \tag{15}$$

where the dot operator within the integral represents an inner product, and S_m is the area of integration over which $\vec{f}_m$ is defined.

Substituting (14) in (13) and weighting with functions $\vec{f}_m$, with $m = 1, \ldots, N$, finally we find a linear equation system $\bar{\bar{Z}} \cdot \vec{I} = \vec{V}$ of dimensions 2 N *x* 2 N:

$$\bar{\bar{Z}} \cdot \vec{I} = \vec{V}\ , \quad \bar{\bar{Z}} = \begin{bmatrix} \bar{\bar{Z}}^{J,(T-EFIE,N-MFIE)} & \bar{\bar{Z}}^{M,(T-EFIE,N-MFIE)} \\ \bar{\bar{Z}}^{J,(T-MFIE,N-EFIE)} & \bar{\bar{Z}}^{M,(T-MFIE,N-EFIE)} \end{bmatrix} \tag{16}$$

The inputs for the five sub-matrixes N × N are given by the following expressions for $m = 1, \ldots, N$ and $n = 1, \ldots, N$:

$$\begin{aligned} \bar{\bar{Z}}^{J,(T-EFIE,N-MFIE)}_{m,n} = & \left\langle \vec{f}_m, \left(\frac{a_1}{\eta_1} L_1 + \frac{a_2}{\eta_2} L_2 \right) \vec{f}_n \right\rangle + \\ & + \left\langle \vec{f}_m, \hat{n}_m \times (b_1 K_1 - b_2 K_2) \vec{f}_n \right\rangle + \\ & + \frac{b_1 + b_2}{2} \left\langle \vec{f}_m, \vec{f}_n \right\rangle \end{aligned} \tag{17}$$

$$\begin{aligned}\bar{\bar{Z}}_{m,n}^{M,(T-EFIE,N-MFIE)} &= -\left\langle \vec{f}_m, \left(\frac{a_1}{\eta_1}K_1 + \frac{a_2}{\eta_2}K_2\right)\vec{f}_n \right\rangle + \\ &+ \frac{1}{2}\left(\frac{a_1}{\eta_1} - \frac{a_2}{\eta_2}\right)\left\langle \vec{f}_m, \hat{n}_m \times \vec{f}_n \right\rangle + \\ &+ \left\langle \vec{f}_m, \hat{n}_m \times \left(\frac{b_1}{\eta_1^2}L_1 - \frac{b_2}{\eta_2^2}L_2\right)\vec{f}_n \right\rangle \end{aligned} \tag{18}$$

$$\begin{aligned}\bar{\bar{Z}}_{m,n}^{J,(T-MFIE,N-EFIE)} &= \left\langle \vec{f}_m, \hat{n}_m \times (-c_1 L_1 + c_2 L_2)\vec{f}_n \right\rangle + \\ &+ \left\langle \vec{f}_m, (d_1\eta_1 K_1 + d_2\eta_2 K_2)\vec{f}_n \right\rangle - \\ &- \frac{d_1\eta_1 - d_2\eta_2}{2}\left\langle \vec{f}_m, \hat{n}_m \times \vec{f}_n \right\rangle \end{aligned} \tag{19}$$

$$\begin{aligned}\bar{\bar{Z}}_{m,n}^{M,(T-MFIE,N-EFIE)} &= \left\langle \vec{f}_m, \hat{n}_m \times (c_1 K_1 - c_2 K_2)\vec{f}_n \right\rangle + \frac{c_1 + c_2}{2}\left\langle \vec{f}_m, \vec{f}_n \right\rangle + \\ &+ \left\langle \vec{f}_m, \left(\frac{d_1}{\eta_1}L_1 + \frac{d_2}{\eta_2}L_2\right)\vec{f}_n \right\rangle \end{aligned} \tag{20}$$

We can write the $\vec{I}$ vector, which contains the unknowns (coefficients of the RWG bases) of the linear system, as the following:

$$\vec{I} = (J_1, J_2, \ldots, J_N, M_1, M_2, \ldots, M_N)^T \tag{21}$$

And the $\vec{V}$ excitation vector of the lineal system is:

$$\begin{aligned}\vec{V} &= \begin{bmatrix} \vec{V}^{(T-EFIE,N-MFIE)} \\ \vec{V}^{(T-MFIE,N-EFIE)} \end{bmatrix} = \\ &= (V_1^{(T-EFIE,N-MFIE)}, \ldots, V_N^{(T-EFIE,N-MFIE)}, V_1^{(T-MFIE,N-EFIE)}, \ldots, V_N^{(T-MFIE,N-EFIE)})^T \end{aligned} \tag{22}$$

where the coefficients are the following;

$$\begin{aligned} V_m^{(T-EFIE,N-MFIE)} &= \frac{a_1}{\eta_1}\left\langle \vec{f}_m, \vec{E}^{inc}(\vec{r}) \right\rangle + b_1\left\langle \vec{f}_m, \hat{n}_m \times \vec{H}^{inc}(\vec{r}) \right\rangle \quad \text{para } m = 1, \ldots, N \\ V_m^{(T-MFIE,N-EFIE)} &= -c_1\left\langle \vec{f}_m, \hat{n}_m \times \vec{E}^{inc}(\vec{r}) \right\rangle + d_1\eta_1\left\langle \vec{f}_m, \vec{H}^{inc}(\vec{r}) \right\rangle \quad \text{para } m = 1, \ldots, N \end{aligned} \tag{23}$$

Once the linear system is resolved, the current densities $\vec{J}(\vec{r})$ and $\vec{M}(\vec{r})$, calculated at each point, can be determined with (14). The electric field scattered at any point in space, inside and outside the disperser, can be calculated directly with $\vec{J}(\vec{r})$ and $\vec{M}(\vec{r})$, using the following two expressions:

$$\vec{E}_{1,scatt}(\vec{r}) = -L_1\vec{J}(\vec{r}) + K_1\vec{M}(\vec{r})$$
$$\vec{E}_{2,scatt}(\vec{r}) = L_2\vec{J}(\vec{r}) - K_2\vec{M}(\vec{r}) \quad (24)$$

The above expressions for scattered field can be derived by following the theoretical development to Eq. (1), but without applying the equalities in (1) to the field vectors. It is important to note that the field equations in (24) are valid for any point of the space and they do not impose any restrictions on the size of the dispersers.

Typically, (24) is known as "near field expressions". For far-field, these expressions are also valid, but it should be much more computationally efficient to use a simplified expression.

Sometimes, the implementation of MoM in computers requires some extra procedures. For example, the integrals of the operators L_i and K_i in (10) can be calculated numerically by a rule of Gaussian quadrature consisting of seven points per triangle, as described Gibson [4]. In this same reference, some theoretical procedures to extract the singularities that occur in integrals when the source point and the observation point are close to each other, $\left|\vec{r} - \vec{r}'\right| \to 0$, can be found.

4 Physical Optics

Physical optics (PO) is a computational technique of high frequency used to calculate the electromagnetic dispersion coming from complex and electrically large PEC (perfect electric conductor) structures [59, 60]. Unlike so-called full-wave methods, for example the method of moments (MoM), the PO does not require high amounts of computational resources to solve dispersion problems with acceptable levels of precision and, above all, with a high efficiency. In this way, simulations that normally take hours with the method of moments, typically are resolved in just a few seconds or minutes with the PO.

4.1 MECA Method

The MECA (modified equivalent current approximation) method [61, 62], has extended PO to dielectric materials with losses characterized by complex effective permittivity. In MECA, the equivalent magnetic and electrical currents are calculated based on the incidence of a locally plane wave on the surface of the dispersant. The MECA equations are derived using a decomposition of the incident field into TE (transverse-electric wave) and TM (transverse-magnetic wave) components, relative to the incident direction and to the normal vectors of each triangular facet in which the surface of the dispersant is discretized. Contrary to what happens in other generalizations of the PO for dielectric media [63–65], MECA takes into account the differences between the TE and TM components, with the consequent increase

in accuracy. In addition, unlike previous approaches, the current distribution on each facet has a uniform amplitude and a distribution of phase that is linear. Therefore, the radiation integral can be resolved in an analytic way and, thus, some difficult problems for a full-wave simulation, especially at very high frequencies, are successfully modelled with MECA.

In the MECA method, the equivalent current densities, magnetic and electrical, are calculate at the barycenter of each facet using the following two equations, respectively:

$$\begin{aligned}\vec{M}_{i0} &= E^i_{TE}(1+R_{TE})(\hat{e}_{TE}\times\hat{n}_i) + E^i_{TM}\cos(\theta_{inc})(1+R_{TM})\,\hat{e}_{TE}\big|_{S_i}\\ \vec{J}_{i0} &= \frac{E^i_{TE}}{\eta_1}\cos(\theta_{inc})(1-R_{TE})\hat{e}_{TE} + \frac{E^i_{TM}}{\eta_1}(1-R_{TM})\,(\hat{n}_i\times\hat{e}_{TE})\big|_{S_i}\end{aligned} \tag{25}$$

where η_1 is the impedance in the medium of incidence, and R_{TE} (and R_{TM}) is the TE (TM) reflection coefficient. The expressions of these two coefficients can be found in work of Meana et al. [62].

As shown in Fig. 2.1, $\vec{E}^i_{TE} = E^i_{TE}\hat{e}_{TE}$ and $\vec{E}^i_{TM} = E^i_{TM}\hat{e}_{TM}$ are the TE and TM components of the incident electric field at the barycenter of the triangle S_{i}. In addition, $\hat{p}_i$ is a unit vector pointing in the direction of propagation of the incident wave, θ_{inc} is the angle of incidence, and $\hat{n}_i$ is the unit normal vector with outgoing direction of the S_{i} triangular facet. The first medium is characterized by its constitutive parameters: permittivity ε_1, permeability μ_1 and conductivity σ_1. Similarly, the second medium is characterized by (ε_2, μ_2, σ_2). The reflection coefficients R_{TE} and R_{TM} depend on all these constitutive parameters. A reflection coefficient is defined for an incidence medium and for a dispersing medium (Fig. 1).

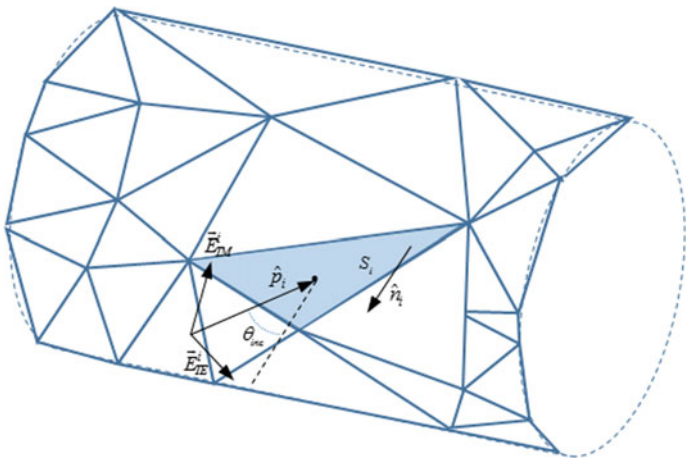


Fig. 1 Oblique incidence on a S_{i} triangular facet

The theoretical calculation of the reflection coefficients R_{TE} and R_{TM} is performed by assuming a locally specular reflection and imposing contour conditions. Ideally, the MECA method would be accurate if the dispersing object were an infinite and homogeneous semi-space. This method, therefore, presents two main sources of error for the TE polarization (analogous for TM, replacing the electric field with the magnetic one). One of the sources of error is in the lack of modeling of diffraction phenomena in the formulation. The other source of error is that multiple reflections are not considered, although this second type of error can be mitigated by iteratively applying coefficients of reflection on discontinuities.

After obtaining the current densities $\vec{M}_{i0}$ and $\vec{J}_{i0}$, an analytical solution can be derived for the radiation integral corresponding to the observation point $\vec{r}_k$, which is located in the far field of each of the triangular facets. The scattered electric field $\vec{E}_k^s$ in $\vec{r}_k$, due to the contribution of all i facets of a given mesh, can be obtained, according to Balanis [66], as:

$$\vec{E}_k^s = \frac{j}{2\lambda} \sum_i \frac{e^{-jk_1 r_{ik}}}{r_{ik}} (\vec{E}_{ik}^a - \eta_1 \vec{H}_{ik}^a \times \vec{r}_{ik}) \tag{26}$$

where λ is the wavelength in the medium of incidence, k_1 is the wave number in the medium of incidence, and $\vec{r}_{ik} = r_{ik}\hat{r}_{ik}$ is the position vector from the barycenter $\vec{r}_i$ of the ith facet to the observation point $\vec{r}_k$. The Eq. (26) is valid when $k_1|\vec{r}_k| \gg 1$. Figure 2.2 summarizes all notation for the position vectors involved in all the scatter calculations that we have used (Fig. 2).

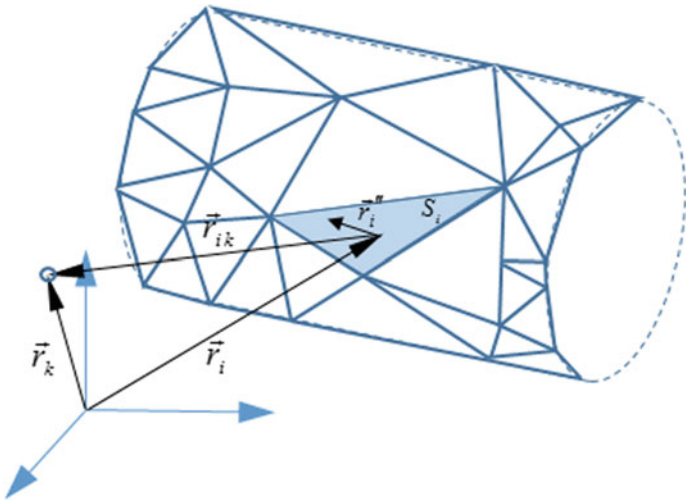


Fig. 2 S_i facet, observation point $\vec{r}_k$ and corresponding position vectors. $\vec{r}_i''$ denotes a variable vector from the barycenter $\vec{r}_i$ to any point on S_i

Assuming that currents have a constant amplitude and a phase variation that is linear and depends on the propagation direction of the incident wave, $\hat{p}_i$, the vector values $\vec{E}^a_{ik}$ and $\vec{H}^a_{ik}$ of the Eq. (26) can be calculated, according to Meana [60], as:

$$\vec{E}^a_{ik} = (\hat{r}_{ik} \times \vec{M}_{i0}) I_i(\hat{r}_{ik}) \tag{27}$$

$$\vec{H}^a_{ik} = (\hat{r}_{ik} \times \vec{J}_{i0}) I_i(\hat{r}_{ik}) \tag{28}$$

where $\vec{M}_{i0}$ and $\vec{J}_{i0}$ are the current densities given in Eqs. (25), and $I_i(\hat{r}_{ik})$ is an integral given by:

$$I_i(\hat{r}) = \iint_{S_i} e^{jk_1(\hat{r}-\hat{p}_i)\cdot\vec{r}''_i}\, dS_i \tag{29}$$

where $\vec{r}''_i$ is a variable vector from the barycenter $\vec{r}_i$ of the ith facet to the points placed on the S_i triangular surface, as it is shown in Fig. 3.

These current distributions make possible to carry out a modelling with facets bigger than those used in other high-frequency methods. This fact implies a computational cost reduction in terms of both time and memory.

The integral of Eq. (29) always admits an analytical solution, according to the procedure described by Arias-Acuña et al. [67]. The method to solve analytically is briefly summarized below.

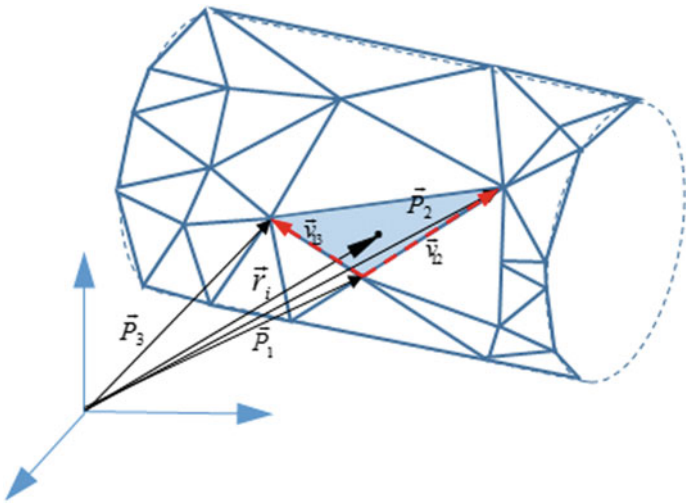


Fig. 3 Triangular facet with barycenter $\vec{r}_i$ and vertices $\vec{P}_1$, $\vec{P}_2$, and $\vec{P}_3$

First, it is considered a plane triangular facet (as shown in Fig. 3). The i triangle is defined by three points: $\vec{P}_1$, $\vec{P}_2$, and $\vec{P}_3$. And $\vec{r}_i$ is a reference point located in the barycenter ($\vec{r}_i = (\vec{P}_1 + \vec{P}_2 + \vec{P}_3)/3$).

We define $\vec{v}_{mn}$ as a vector such that $\vec{v}_{mn} = \vec{P}_n - \vec{P}_m \ , \quad m, n \in \{1, 2, 3\}$. The vector $\hat{n}$, normal to the i triangle, is defined so that $\vec{v}_{12} \times \vec{v}_{13} = 2A_i\hat{n}$, as shown in Fig. 2.3, being A_i the area of that triangle.

Finally, we use a coordinate system with two scalar variables (u, v) such that any point $\vec{r}_i''$ of the surface of the triangle can be described as:

$$\vec{r}_i'' = \vec{P}_1 - \vec{r}_i + u \cdot \vec{v}_{12} + v \cdot \vec{v}_{13} \tag{30}$$

The integral (29) is given by:

$$I_i(\hat{r}) = 2A_i e^{-j\frac{\alpha+\beta}{3}} \int_{u=0}^{u=1} \int_{v=0}^{v=1-u} e^{j(\alpha u+\beta v)} dv du \tag{31}$$

and its solution is:

$$I_i(\hat{r}) = 2A_i e^{-j\frac{\alpha+\beta}{3}} \left[\frac{\alpha e^{j\beta} - \beta e^{j\alpha} + \beta - \alpha}{\alpha\beta(\alpha - \beta)} \right] \tag{32}$$

where

$$\alpha = k_1 \vec{v}_{12}(\hat{r} - \hat{p}_i) \tag{33}$$

$$\beta = k_1 \vec{v}_{13}(\hat{r} - \hat{p}_i) \tag{34}$$

The expression (32) has the following four eigenvalues:

$$\alpha = 0, \beta \neq 0 \Rightarrow I_i(\hat{r}) = 2A_i \ e^{-j\frac{\beta}{3}} \ \frac{1 + j\beta - e^{j\beta}}{\beta^2} \tag{35}$$

$$\alpha \neq 0, \beta = 0 \Rightarrow I_i(\hat{r}) = 2A_i \ e^{-j\frac{\alpha}{3}} \ \frac{1 + j\alpha - e^{j\alpha}}{\alpha^2} \tag{36}$$

$$\alpha = \beta \neq 0 \Rightarrow I_i(\hat{r}) = 2A_i \ e^{j\frac{\alpha}{3}} \ \frac{1 - j\alpha - e^{-j\alpha}}{\alpha^2} \tag{37}$$

$$\alpha = \beta = 0 \Rightarrow I_i(\hat{r}) = A_i$$

5 Comparison Between Method of Moments and Modified Equivalent Current Approximation

In this section, we do to a graphical comparison that allows to easily show the difference between PO and MoM in terms of accuracy level. Figure 2.4 shows the monostatic radar cross section (RCS) of a sphere, defined as the following:

$$RCS_{mono} = \lim_{r\to\infty}\left(4\pi r^2\left|\vec{E}_{1,scatt}(r,\theta_{inc},\phi_{inc})\right|^2/\left|\vec{E}^{inc}\right|^2\right) \tag{38}$$

where a is the radius of the sphere and $\frac{2\pi}{\lambda}$ is the wave number in the medium of incidence. In case, $\vec{E}^{inc}$ is a uniform plane wave. Two cases are included in Fig. 4: PEC sphere and dielectric sphere with losses simulated with MECA for the parameters shown in the Table 2.

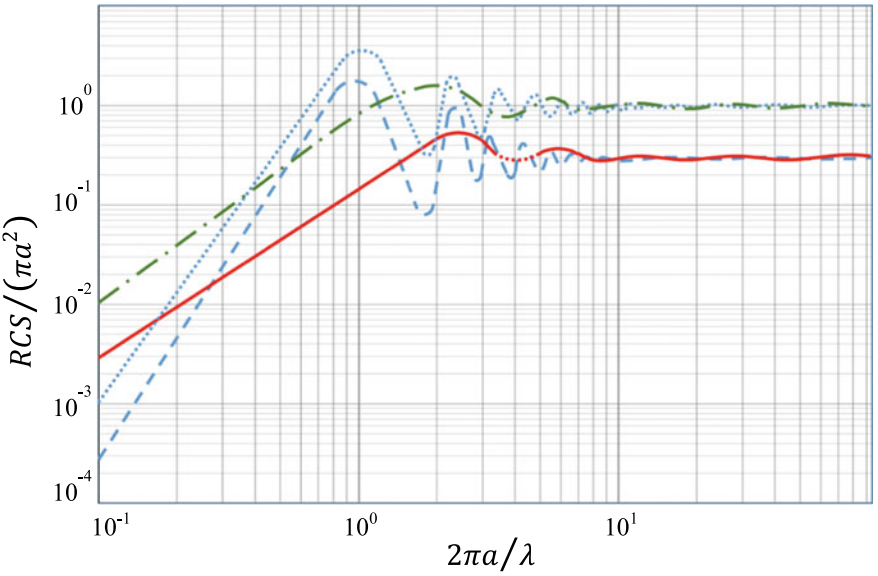


Fig. 4 Comparison between PO (MECA) and MoM solutions for scattering due to a sphere of radius a with different constitutive parameters

Table 2 Parameters used in the simulation of Fig. 4

	ε_r	μ_r	RCS
MECA	1.8	1.5	$10w\varepsilon_0$
MoM	1.8	1.5	$10w\varepsilon_0$

6 Conclusions

Techniques to solve classical electromagnetism problems have evolved over time. At first, the analysis and design of electromagnetic devices and structures was done experimentally, achieving their characterization; this result was used for the development of new technologies. Subsequently, analytical models emerged where closed form solutions are obtained, that is, solutions that model and describe the electromagnetic phenomenon through a simplified algebraic equation under ideal situations.

At present, the numerical solution is arrived at using computational algorithms, in which various numerical analysis techniques are used that describe the phenomena in time and space of electromagnetic problems that previously could not be solved analytically. In fact, many numerical analysis techniques have been developed in recent years, leading to advances in this area, referred to as computational electromagnetism (CEM). Similarly, the wide range of electromagnetic problems has led to the development of different algorithms in computational electromagnetism, each with its advantages and limitations.

Numerical solutions to three-dimensional electromagnetic dispersion problems are generally found on the formulation of surface integral equations, such as the electric field integral equations (EFIE) and the similar magnetic field integral equations (MFIE), or even the less used combined field integral equation (CFIE). The moment method (MoM) is, at present, one of the most commonly used numerical method to solve these type of equations. In the solution of this method, the induced electric current and magnetic current are unknown variables. Moreover, the surface is generally partitioned into small flat patches. In these patches, the currents are approximately calculated by some appropriate basic functions.

These patches have simple shapes to be easier implemented and for doing calculations in also easier way. The most commonly used forms for this partition or division are triangular and rectangular patches. If the size of these subdivisions (patches) is small enough, then we could approximate the induced surface currents by the triangular (or rectangular) functions of the ceiling. The application of these type of functions using the method of moment in order to solve the surface integral equations has the consequence of the evaluation of double integrals with single cores.

The Method of Moments (MoM) is a numerical technique used to convert the integral equation into a linear system, which can be solved using a computer.

The main reasons why researchers select this method are:

- It solves the Maxwell equations without implicit approximations.
- It presents greater numerical stability in the discretization of integrals versus derivatives.
- It allows to exclude the medium that is around the structure and therefore facilitates the analysis of open structures.
- It analyzes the problem in a rigorous and precise way, taking into account most of the physical phenomena that occur in the structures, so that the analysis is valid in principle for any frequency.

This method allows the systematic formulation of the problem through the discretization of the electric field integral equation (EFIE), and calculates by numerical methods the densities of unknown currents. However, it presents a disadvantage, which is given from restricting it to problems of small electrical size due to limitations in memory and time in the computational process.

References

1. Sadiku M (2018) Elements of electromagnetics, 7th edn. Oxford University Press
2. Christiansen D, Jurgen R, Fink D (1996) Electronics engineers handbook (standard handbook of electronics engineering), 4th edn. McGraw-Hill Education
3. Baro I, Bilbao J, Varela C (2007) Linear algebra (*in Basque: Aljebra lineala*). Ed. Centro de Publicaciones de la Escuela Técnica Superior de Ingenieros Industriales y de Telecomunicación de Bilbao, ISBN: 978-84-95809-31-5
4. Gibson WC (2014) The method of moments in electromagnetics, 2nd edn. Chapman & Hall/CRC
5. Taflove A, Hagness SC (2005) Computational electrodynamics: the finite-difference time-domain method, 3rd edn. Artech House
6. Kunz K, Luebbers R (1993) The finite difference time domain method for electromagnetics. CRC Press
7. Thomas JW (1995) Numerical partial differential equations: finite difference methods. Springer Verlag, Berlin, Germany
8. Kermani MH, Ramahi OM (2006) The complementary derivatives method: a second-order accurate interpolation scheme for non-uniform grid in FDTD simulation. IEEE Microw Wireless Compon Lett 16:60–62
9. Jin J (1993) The finite element method in electromagnetics. Wiley
10. Volakis JL, Chatterjee A, Kempel LC (1998) Finite element method for electromagnetics. IEEE Press
11. Bathe KJ (1995) Finite element procedures, 2nd edn. Prentice Hall
12. Polycarpou AC (2006) Introduction to the finite element method in electromagnetics. Synthesis Lect Comput Electromag 1(1):1–126
13. Courant R (1943) Variational methods for the solution of problems of equilibrium and vibrations. Bull Am Math Soc 49:1–43
14. Turner MJ, Clough RW, Martin HC, Topp LJ (1956) Stiffness and deflection analysis of complex structures. J Aeronaut Sci 23:805–824
15. Bilbao J, Bravo E, García O, Varela C, Rodríguez M, González P (2015) Blade aerodynamic design and analysis as first step to achieve the expected power performance of a small wind turbine. Int J Tech Phys Prob Eng 3(7):42–46
16. Przemieniecki JS (1968) Theory of matrix structural analysis. Mc GRaw-Hill, New York
17. Zienkiewicz OC, Holister GS (1965) Stress analysis. Wiley, London
18. Zienkiewicz OC, Cheung YK (1967) The finite element method in structural and continuum mechanics. Mc Graw-Hill, London
19. Zienkiewicz OC, Taylor RL, Zhu JZ (2005) The finite element method: its basis and fundamentals, 6th edn. Elsevier
20. Harrington RF (1993) Field computation by moment methods. IEEE Press. Series on Electromagnetic Waves, pp. 15–64
21. Keller JB (1962) Geometrical theory of diffraction. J Opt Soc Am 52:116–130
22. Kouyoumjian RG, Pathak PH (1974) A uniform geometrical theory of diffraction for and edge in a perfectly conducting surface. Proc IEEE 62:1448–1461
23. Keller J (1960) Backscattering from a finite cone. IRE Trans Antennas Propag 8(2)

24. Ufimtsev PY Method of edge waves in the physical theory of diffraction. Foreign Technology Div Wright-Patterson AFB, Ohio. Retrieved from: https://apps.dtic.mil/dtic/tr/fulltext/u2/733203.pdf
25. Ufimtsev PY (2014) Fundamentals of the physical theory of diffraction, 2nd edn. Wiley
26. Akhiyarov VV, Borzov AB, Likhoedenko KP, Karakulin YV, Seregin GM, Suchkov VB (2018) Mathematical simulation of electromagnetic scattering field from perfectly conducting object with dielectric cover on the base of physical theory of diffraction. In: CSAE '18 proceedings of the 2nd international conference on computer science and application engineering, article No. 67, Hohhot, China, October 22–24
27. Ufimtsev PY (1957) Approximate computation of the diffraction of plane electromagnetic waves at certain metal bodies (i and ii). Sov Phys Tech 27:1708–1718
28. Michaeli A (1984) Equivalent edge currents for arbitrary aspects of observation. IEEE Trans Antennas Propag 23:252–258
29. Lee H, Koh IS (2018) Consideration of diffraction effect in iterative physical optics combining physical theory of diffraction for conducting body. In: 12th European conference on antennas and propagation (EuCAP 2018)
30. Balanis CA (1989) Advanced engineering electromagnetics. Wiley
31. Al-Azzawi A (2018) Physical optics: principles and practices. CRC Press
32. Canta SM, Kipp RA, Carpenter S, Petersson LER (2018) Range-Doppler radar signature prediction of wind turbine using SBR. In: 12th European conference on antennas and propagation (EuCAP 2018)
33. Ghanmi H, Khenchaf A, Pouliguen P (2018) Radar cross section of modified target using gaussian beam methods: experimental validation. In: 2018 international conference on radar (RADAR)
34. Ling SLH, Chou R (1989) Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity. IEEE Trans Antennas Propag 37:194–205
35. Ling SLH, Chou R (1989) High-frequency RCS of open cavities with rectangular and circular cross sections. IEEE Trans Antennas Propag 37:648–652
36. Yilmaz AE, Jin JM, Michielssen E (2004) Time domain adaptive integral method for surface integral equations. IEEE Trans Antennas Propag 52(10)
37. Engheta N, Murphy WD, Rokhlin V, Vassiliou MS (1992) The fast multipole method (FMM) for electromagnetic scattering problems. IEEE Trans Antennas Propag 40(6)
38. Song J, Lu CC, Chew WC (1997) Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects. IEEE Trans Antennas Propag 45(10)
39. Hasanov ER, Hajiyeva VM, Salimi Rikani A, Tabatabaei NM (2019) Radiation doped semiconductors with certain impurities. Int J Tech Phys Prob Eng (IJTPE) 11(2):13–17
40. Akbarli RS (201) Waves propagation in the fluid flowing in an elastic tube, considering viscoelastic friction of surrounding medium. Int J Tech Phys Propag Eng (IJTPE) 10(2):39–42
41. Kim OS, Meincke P, Breinbjerg O (2004) Method of moments solution of volume integral equations using higher-order hierarchical Legendre basis functions. Radio Sci 39:RS5003
42. Saad Y, Schultz M (1986) GMRES: a generalized minimal residual algorithm for solving non-symmetric linear systems. SIAM J Sci Stat Comput 7(15):856–869
43. Medgyesi-Mitschang LN, Putnam JM, Gedera MB (1994) Generalized method of moments for three-dimensional penetrable scatterers. J Opt Soc Am A 11(4):1383–1398
44. Poggio AJ, Miller EK (1973) Integral equation solutions of three-dimensional scattering problems. In: Mittra R (edn) Computer techniques for electromagnetics. Pergamon Press, Oxford
45. Ylä-Oijala P, Taskinen M, Järvenpää S (2005) Surface integral equation formulations for solving electromagnetic scattering problems with iterative methods. Radio Sci 40(6):RS6002
46. Ergül Ö (2009) Accurate and efficient solutions of electromagnetics problems with the multilevel fast multipole algorithm. Ph.D. dissertation, Bilkent University, Ankara, Turkey, 2009. Retrieved from: https://core.ac.uk/download/pdf/52925780.pdf
47. Chang Y, Harrington RF (1977) A surface formulation for characteristic modes of material bodies. IEEE Trans Antennas Propag 25:789–795

48. Wu TK, Tsai LL (1977) Scattering from arbitrarily-shaped lossy dielectric bodies of revolution. Radio Sci 12:709–718
49. Ergül Ö, Gürel L (2008) Novel electromagnetic surface integral equations for highly accurate computations of dielectric bodies with arbitrarily low contrasts. J Comput Phys 227(23):9898–9912
50. Ylä-Oijala P, Taskinen M (2005) Application of combined field integral equation for electromagnetic scattering by dielectric and composite objects. IEEE Trans Antennas Propag 53(3):1168–1173
51. Ylä-Oijala P (2008) Numerical analysis of combined field integral equation formulations for electromagnetic scattering by dielectric and composite objects. Prog Electromagn Res C 3:19–43
52. Cui Z, Han Y, Xu Q, Li M (2010) Parallel MoM solution of JMCFIE for scattering by 3-D electrically large dielectric objects. Prog Electromagn Res M 12:217–228
53. Taboada JM, Rivero J, Obelleiro F, Araujo MG, Landesa L (2011) Method-of-moments formulation for the analysis of plasmonic nano-optical antennas. J Opt Soc Am A 28(7):1341–1348
54. Araujo MG, Taboada JM, Rivero J, Solís DM, Obelleiro F (2012) Solution of large-scale plasmonic problems with the multilevel fast multipole algorithm. Opt Lett 37(3):416–418
55. Ergül Ö, Gürel L (2008) Stabilization of integral-equation formulations for the accurate solution of scattering problems involving low-contrast dielectric objects. Trans Antennas Propag 56(3):799–805
56. Ergül Ö, Gürel L (2009) Comparison of integral-equation formulations for the fast and accurate solution of scattering problems involving dielectric objects with the multilevel fast multipole algorithm. IEEE Trans Antennas Propag 57(1):176–187
57. Rao SM, Wilton DR, Glisson AW (1982) Electromagnetic scattering by surfaces of arbitrary shape. IEEE Trans Antennas Propag 30(3):409–418
58. Li X, Lei L, Chen Y, Jiang M, Nie Z, Hu J (2019) Efficient electromagnetic analysis for complex planar thin-layer composite objects by a hybrid method. IEEE Antennas Wireless Propag Lett 18(9)
59. Uluisik C, Cakir G, Cakir M, Sevgi L (2008) Radar cross section (RCS) modeling and simulation, part 1: a tutorial review of definitions, strategies, and canonical examples. Antennas Propag Mag IEEE 50(1):115–126
60. Martínez-Lorenzo JA, Pino AG, Vega I, Arias M, Rubiños O (2005) ICARA: induced-current analysis of reflector antennas. Antennas Propag Mag IEEE 47(2):92–100
61. Meana JG, Martínez-Lorenzo JA, Las-Heras F, Rappaport C (2009) A PO MoM comparison for electrically large dielectric geometries. In: Antennas and propagation society international symposium, 2009. APSURSI '09, IEEE. 1–5 June 2009
62. Meana JG, Martínez-Lorenzo JA, Las-Heras F, Rappaport C (2010) Wave scattering by dielectric and lossy materials using the modified equivalent current approximation (MECA). IEEE Trans Antennas Propag 58(11):3757–3761
63. Rengarajan SR, Gillespie ES (1988) Asymptotic approximations in radome analysis. IEEE Trans Antennas Propag 36(3):405–414
64. Hodges RE, Rahmat-Samii Y (1993) Evaluation of dielectric physical optics in electromagnetic scattering. In: Proceedings 1993 antennas and propagation society international symposium, USA
65. Sáez de Adana F, Gutierrez O (2010) Practical applications of asymptotic techniques in electromagnetics. Artech House, 2010
66. Balanis CA (1997) Antenna theory: analysis and design. Wiley, New York
67. Arias-Acuña M, Rubiños O, Cuiñas I, Pino AG (2000) Electromagnetic scattering of reflector antennas by fast physical optics algorithms. Recent Res Dev Magn 1:43–63

Optimization Methods for Wireless Power Transfer

Lavinia Bobaru, Mihai Iordache, Marilena Stanculescu, Dragos Niculae, and Sorin Deleanu

Abstract This chapter provides an investigation of the wireless power transfer domain. It follows the description of the analysis and identification of the system's parameters, which consist of two magnetically connected coils utilized in constructing the wireless power transfer system. A section of this chapter focuses on the optimization aspects of the wireless energy transfer. The optimization considers a function of several parameters of the system such as the structure, the number of turns, the WPTS's working frequency.

Keywords Wireless power transfer · Optimization · Magnetic couplings · Parameter identification

Nomenclature

A. Acronyms

WPT	Wireless Power Transfer
WEP	Wireless Energy Transfer
SCORs	Strongly Coupled Resonances

L. Bobaru · M. Iordache · M. Stanculescu (✉) · D. Niculae
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania
e-mail: marilena.stanculescu@upb.ro

L. Bobaru
e-mail: lavinia.bobaru@upb.ro

M. Iordache
e-mail: mihai.iordache@upb.ro

D. Niculae
e-mail: dragos.niculae@upb.ro

S. Deleanu
Northern Alberta Institute of Technology, Edmonton, Canada
e-mail: sorind@nait.ca

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_20

MIT	Massachusetts Institute of Technology
MM	Fmagnetomotive force
EMF	Electromotive force
AC	Alternating current
ss	series-series configuration
pp	parallel-parallel configuration
sp	series-parallel configuration
ps	parallel-series configuration
APL	active power transmitted to the load

B. Symbols/Parameters

$u_{m\Gamma}$	magnetomotive force
i_{S_Γ}	conduction electric current strength
Ψ_{S_Γ}	electric flux
Θ_{S_Γ}	turn
H	magnetic field strength (intensity)
J	current density
B	magnetic flux density (induction)
D	electric displacement
Φ_{S_Γ}	magnetic flux
E	electric field strength
$\overline{w}$	local speed vector of the medium
u_{Γ}	EMF
u_t	EMF induced by transformation
u_m	EMF induced by movement
R	resistance
L	inductivity
C	capacity
M	mutual inductivity
R	the matrix of resistive coefficients
L	the matrix of inductive coefficients
C	the matrix of capacitive coefficients
r	initial radius of the turn
p	pitch (distance between two consecutive turns)
h	the distance between the two coils along Oz axis
N_1, N_2	the number of turns for coil 1, respectively, for coil 2
f	frequency
$H(f)$	transfer function
g_j	objective function
P	power
η	efficiency
(Γ)	closed curve
(S_Γ)	open surface
(Σ)	closed surface

$\overline{dA}$ area element
$\overline{dl}$ line element

1 Introduction to Wireless Power Transfer

Wireless Power Transfer (WPT) or Wireless Energy Transfer (WEP) represents a new technology, suitable to bring electricity to places where the utilization of system interconnection via cables is either difficult, impractical or even impossible. Although the WPT is ineffective for large distances [1, 2] due to the electromagnetic field weakness [3], there is the possibility to achieve a more efficient power transfer if the emitting coil (emitter) and the receiving coil (receiver) are at resonance [5, 6]. Both the electromagnetism and the electric circuit theory recognize the condition of "resonance" for the best efficiency possible regarding the power transfer.

Despite a total power loss counting for 30% of the transferred power, WPT has numerous applications [7–9] such as:

- *Medical implants*. Technological progress in the field of biomedical allowed the creation of biomedical implants such as pacemakers, cochlear implants, subcutaneous drug supply implants
- *Chargers* used for smartphones, electric and hybrid vehicles, unmanned aerial vehicles.
- *Appliances* such as ironing, vacuum cleaner, TV.

Currently, there are three types of WPTs: *radiant transfer*, *inductive transfer*, and *resonant coupling transfer*. In the case of radiant transfer, most of the generated power is lost in surrounding space. In contrast, the received power comes in small amounts (i.e., in order of mW), making the radiant transfer suitable for the transmission of information. Regarding the inductive coupling [5], the power transfer can be very efficient, although the distance emitter-receiver is only in terms of a few centimeters. The application of the resonant coupling method allows the transfer of significant power over relatively long distances (few meters). The experimental studies carried out at MIT led to a potential breakthrough in WPT. The researchers proposed a new scheme based on the concept named Strongly Coupled Resonances (SCORs) [6, 9]. The fundamental principle of this type of transmission is based on the idea that resonant objects exchange energy, whereas the non-resonant ones do not realize such an energy transfer [7, 8]. The WPT transformation from concept into large scale applications occurred relative recently, although one can track the first research works on the subject back in the 1880s [8]. Energy transmission without galvanic contacts did not become an established technique yet, capable of operating with clear solutions, design methods, and regulated practices. For example, only a few years ago (2012), WPT had a real narrow niche, its applications addressing the cell phones and digital tablets, only. However, later, in 2015, the WPT market considerably expanded, its value reaching 1 billion dollars. There are predictions in

place [10], suggesting that the WPT market would reach a value of 5 billion dollars by 2020. A WPT market dominated by ordinary small customers is in sight for the future, no matter the real big business desire for a significant WPT expansion in the industry in general, and the automotive sector.

The chapter starts by analyzing the current literature and by presenting the different wireless power transfer methods together with their corresponding applications, which varies from domains such as the medical one up to the military one. It follows the description of the analysis and identification of the system's parameters, which consist of two magnetically connected coils utilized in constructing the wireless power transfer system.

A section of this chapter focuses on the optimization aspects of the wireless energy transfer. The optimization requires the consideration of several aspects. It is presented the WPTS's optimization by using the transfer function method. The optimization method for the system consisting of the two magnetically coupled coils and used in WPTS considers a function depending on several parameters of the system, such as the structure, the number of turns, the WPTS's working frequency.

Electromagnetic energy wireless transfer is a developing, emerging technology, resulted from the significant progress in the power electronics domain. The possible applications of this technology present an enormous potential, which may influence the way we use the current applications.

Despite the fact there are vital signs of progress in this domain, we are still far from reaching this objective, due to the significant design challenges. The small efficiency and the limited range of transfer are two of the most crucial aspect which must be improved.

The chapter ends with conclusions and many references in the field of wireless power transfer.

2 Theoretical Considerations

The WPT systems operate according to two fundamental principles:

- The magnetic circuit law (Ampere's Law)
- The electromagnetic induction law (Faraday-Lenz Law).

The magnetic circuit law(Ampère's Law), according to [11], has the initial definition mathematically summarized in the differential form (3.1). The magnetic circuit law consists of the following statement:

The magnetomotive force (MMF), $u_{m\Gamma}$, along any closed curve (Γ) is equal to the sum between the *conduction electric current strength* i_{S_Γ} through an open surface (S_Γ) arbitrary chosen, bordered by the closed curve (Γ) and the *time derivative of the electric flux* Ψ_{S_Γ} over the same surface: (S_Γ)

$$u_{m_\Gamma} = i_{S_\Gamma} + \frac{d\Psi_{S_\Gamma}}{dt} \tag{1}$$

The term i_{S_Γ} is also called turn, and denoted by Θ_{S_Γ}.

By rewriting relation (1), one obtains the integral form of Ampere's law (2):

$$\int\limits_{(\Gamma)} \overline{H} \cdot \overline{dl} = \int\limits_{(S_\Gamma)} \overline{J} \cdot \overline{dA} + \frac{d}{dt} \int\limits_{(S_\Gamma)} \overline{D} \cdot \overline{dA} \tag{2}$$

The integral form is valid only in the context of the application of the corkscrew rule between the reference direction, the oriented line element $\overline{dl}$, and the oriented area element $\overline{dA}$ This observation applies concerning the electromagnetic induction law as well. For domains with continuity, one can modify the relation (2) according to the properties of the operators into (3).

$$curl\overline{H} = \overline{J} + \frac{\partial \overline{D}}{\partial t} + \overline{w} \cdot \rho_v + curl\left(\overline{D} \times \overline{w}\right) \tag{3}$$

The last form is the magnetic circuit law expressed in the local (punctual) form:

For stationary media $\overline{w} = 0$ the local form becomes:

$$curl\overline{H} = \overline{J} + \frac{\partial \overline{D}}{\partial t} \tag{4}$$

According to (4), one can affirm that the closed lines of the magnetic field border the entities responsible for its generation:

- The wires transited by conduction currents.
- The lines of the time variable electric field that generates them.

The electromagnetic induction law (Faraday's law) has the following statement [11]:

The electromotive force (EMF) u_Γ along a closed curve (Γ) is equal to the negative of the time derivative of the magnetic flux Φ_{S_Γ} across any surface (S_Γ) bordered by the closed curve (Γ):

$$u_\Gamma = -\frac{d\Phi_{S_\Gamma}}{dt} \tag{5}$$

Note: The form presented in (5) includes the adjustment represented by the "negative" sign introduced by Lenz.

Considering the definition relations of EMF, respectively of the magnetic flux, one can extrapolate (8) into the integral form of the Faraday-Lenz law (6):

$$\int\limits_{(\Gamma)} \overline{E} \cdot \overline{dl} = -\frac{d}{dt} \int\limits_{(S_\Gamma)} \overline{B} \cdot \overline{dA}. \tag{6}$$

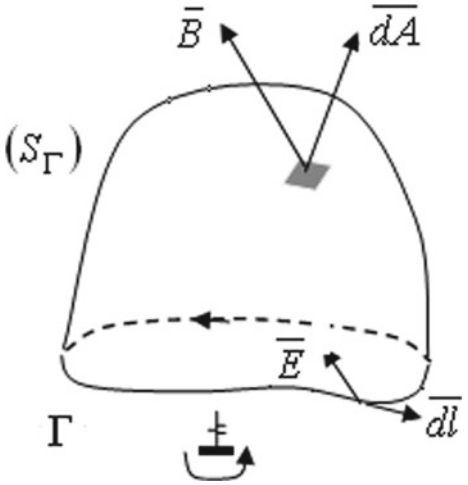


Fig. 1 The Electromagnetic induction law

The integral for of the electromagnetic induction law is valid only when fulfilling the condition: the reference direction of the closed curve (Γ) (i.e., the reference direction of the oriented line element $\overline{dl}$) and the direction of the normal to the surface (S_Γ) (i.e., the oriented area element $\overline{dA}$)) comply with the right corkscrew rule (see Fig. 1).

For mobile media, the integration domains follow the bodies in their movement, and the time derivative of the magnetic flux becomes a substantial derivative and it is computed using the following relation:

$$\frac{d}{dt}\int\limits_{(S_\Gamma)} \overline{B}\cdot\overline{dA} = \int\limits_{(S_\Gamma)} \left[\frac{\partial \overline{B}}{\partial t} + \overline{w}\cdot div\,\overline{B} + curl\left(\overline{B}\times\overline{w}\right)\right]\cdot\overline{dA}, \tag{7}$$

In (7), $\overline{w}$ is the local speed vector of the medium.

Using Stoke relation, results:

$$\int\limits_{(\Gamma)} \overline{E}\cdot\overline{dl} = \int\limits_{(S_\Gamma)} \left(curl\,\overline{E}\right)\cdot\overline{dA} \tag{8}$$

Considering the local form of the magnetic flux law, one obtains a new integral form of the Faraday's law of electromagnetic induction:

$$u_\Gamma = \int\limits_{(\Gamma)} \overline{E}\cdot\overline{dl} = -\int\limits_{(S_\Gamma)} \frac{\partial \overline{B}}{\partial t}\cdot\overline{dA} - \int\limits_{(\Gamma)} \left(\overline{B}\times\overline{w}\right)\cdot\overline{dl} \tag{9}$$

The physical significance of the law: the time-variable magnetic field produces (induces) an electric field through the electromagnetic induction. Therefore, the electromagnetic induction is a physical phenomenon, unlike electric the flux density $\overline{D}$ and magnetic induction $\overline{B}$, which are physical quantities.

Furthermore, the decomposition of the EMF in two components:

$$u_\Gamma = u_t + u_m \tag{10}$$

With

$$u_t = -\int\limits_{(S_\Gamma)} \frac{\partial \overline{B}}{\partial t} \cdot \overline{dA} = \int\limits_{(S_\Gamma)} \left(-\frac{\partial \overline{B}}{\partial t}\right) \cdot \overline{dA} \tag{11}$$

called electromotive force (EMF) induced by transformation and, respectively,

$$u_m = -\int\limits_{(\Gamma)} \left(\overline{B} \times \overline{w}\right) \cdot \overline{dl} = \int\limits_{(\Gamma)} \left(\overline{w} \times \overline{B}\right) \cdot \overline{dl} \tag{12}$$

called EMF induced by movement.

The two components reveal the two forms of electromagnetic induction:

- time variation of the magnetic induction $\overline{B}$, with no movement u_t
- at least a portion of the closed curve (Γ) is mobile in the space containing magnetic field u_m.

The simultaneous application of both laws, the magnetic circuit law, and the electromagnetic induction one, detailed above, provide the proper frame for WPT development. The wireless transfer of the electromagnetic power requires the presence of, at least, two magnetically coupled coils, a fact which labels WPTs as inductive power transfer systems.

A simple explanation of the WPT operation is:

- An AC flowing through a coil (called primary coil, source or transmitter) produces an AC magnetic field, having the same frequency of the current.
- The AC magnetic field due to the primary coil sweeps the closed surface bordered by another coil (called secondary coil, resonator device, or receiver) placed about the primary one and induces an EMF across the receiver transferred inductively, [12–18] (see Fig. 2).

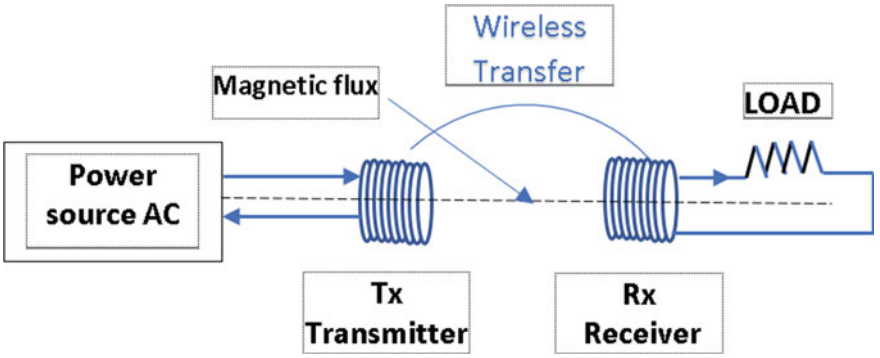


Fig. 2 Schematic view of a WPTS

3 Parameters Identification for the Wireless Power Transfer

Several literature resources identified the pancake type of coils as suitable for applications regarding battery charging solutions [19–24]. Assuming that for every structure emitter-receiver (two coils magnetically coupled), the number of turns, geometrical dimensions, and the fabrication materials remain the same (the frequency is the variable of interest). Following theoretical and experimental tests, one can conclude that the structure with the value of the mutual inductivity represents the optimal coupling solution [19–29]. A method exemplifying the optimization process applied to the configuration with two magnetically coupled coils utilized the ANSYS Q3D Extractor software [30–32]. According to this method, there is a set of matrices $\boldsymbol{R}$, $\boldsymbol{L}$, $\boldsymbol{C}$, and $\boldsymbol{G}$, representing the resistive, inductive, capacitive, respectively coupling coefficient elements. The representation of the last matrix, $\boldsymbol{G}$, becomes a simple number. The assessment of several possible solutions of coupled coils utilized in WPT, concerning configurations, structures, and frequencies guides the designer, finally, to the optimum one. The frequency values of interest in these simulations are 50, 5000, and 10,000 kHz, whereas the four analyzed configurations with their parameters appear in Figs. 3, 4, 5, 6.

3.1 The Assessment of Two Parallel-Bases Helicoidally Pancake-Type Coils

The two coils (see Fig. 3) have the following geometrical parameters:

- Initial radius of the turn $r = 10\ mm$
- Pitch (distance between two consecutive turns) on OY axis $p = 1.3\ mm$
- Section of the conductors $s = 1.2\ mm \times 0.8\ mm = 0.96mm^2$

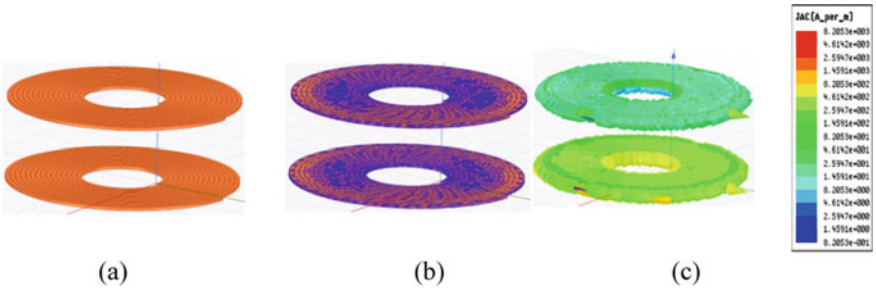


Fig. 3 **a** The structure of the two helicoidally coils with parallel bases; **b** The distribution of the discretization points; **c** Current sheet distribution

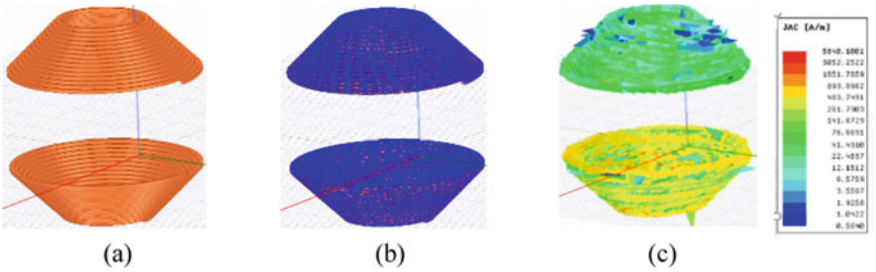


Fig. 4 **a** The structure of the two coils truncated cone-like with parallel bases **b** Distribution of discretization points; **c** Distribution of current sheet

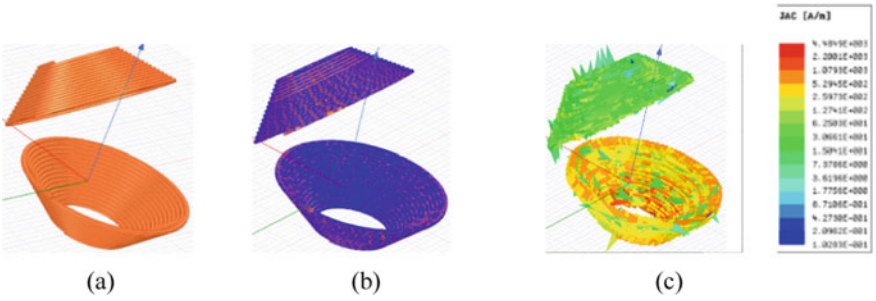


Fig. 5 **a** The structure of the two coils truncated cone-like with non-parallel bases, one rotated 450; **b** Distribution of discretization points; **c** Distribution of current sheet

- Distance between the two coils along Oz axis $h = 20\ mm$
- Number of turns $N_1 = N_2 = 15$.

The identification of electric parameters (i.e., ohmic resistances, parasitic capacitances, self-inductivities, as well as the coupling coefficient) filled in for Table 1.

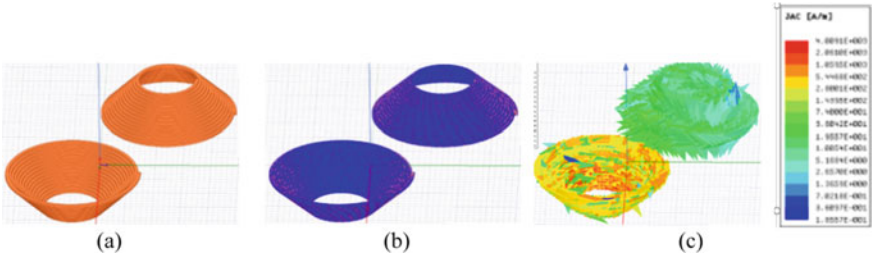


Fig. 6 **a** The structure of the two truncate cone-like coils, with parallel bases, z = 20 mm and y = 15 mm; **b** Distribution of the discretization points; **c** Distribution pf current sheet

Table 1 The electric parameters for two helicoidally pancake coils with parallel basis

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.72029 & 0.26477 \\ 0.26477 & 0.72313 \end{pmatrix}$	$\begin{pmatrix} 0.20907 & 0.02964 \\ 0.02964 & 0.20839 \end{pmatrix}$	$\begin{pmatrix} 2.9425 & 0.41043 \\ 0.41043 & 3.0142 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 8.0489 & 1.5812 \\ 1.5812 & 8.0487 \end{pmatrix}$	$\begin{pmatrix} 6.8438 & 0.26477 \\ 0.63748 & 6.7874 \end{pmatrix}$	$\begin{pmatrix} 6.6993 & 0.62623 \\ 0.62623 & 6.6878 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 3.1308 & -1.7213 \\ -1.7213 & 3.1325 \end{pmatrix}$	$\begin{pmatrix} 3.1308 & -1.7213 \\ -1.7213 & 3.1325 \end{pmatrix}$	$\begin{pmatrix} 3.1308 & -1.7213 \\ 0.2648 & 3.1325 \end{pmatrix}$
Coupling coefficient [μH]	0.19646	0.19646	0.19742

3.2 *The Assessment of Two Truncated Cone-Like Coils, with Parallel Bases, Configuration a*

The two coils (see Fig. 4) have the following geometrical parameters:

- Turn's initial radius $r = 10$ mm
- Pitch on OY axis $p = 0.9$ mm
- Section of the conductors $s = 1.2 \times 0.8 = 0.96$ mm^2
- Distance between the two coils along Oz axis $h = 20$ mm
- Number of turns $N_1 = N_2 = 15$.

The identification of electric parameters (i.e., ohmic resistances, parasitic capacitances, self-inductivities, as well as the coupling coefficient) are filled in for Table 2.

Table 2 The electric parameters for two truncated cone-like coils with parallel basis, configuration A

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.72029 & 0.26477 \\ 0.26477 & 0.72313 \end{pmatrix}$	$\begin{pmatrix} 1.2159 & 0.12714 \\ 0.12714 & 1.4602 \end{pmatrix}$	$\begin{pmatrix} 2.0539 & 0.211247 \\ 0.21124 & 1.7596 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 6.7639 & 0.63251 \\ 0.63251 & 6.7609 \end{pmatrix}$	$\begin{pmatrix} 6.8438 & 0.26477 \\ 0.63748 & 6.7874 \end{pmatrix}$	$\begin{pmatrix} 6.7947 & 0.63513 \\ 0.63513 & 6.8295 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 3.1309 & -1.7213 \\ -1.7213 & 3.1325 \end{pmatrix}$	$\begin{pmatrix} 2.376 & -1.0406 \\ -1.0406 & 2.3759 \end{pmatrix}$	$\begin{pmatrix} 3.1308 & -1.7213 \\ -1.7213 & 3.1325 \end{pmatrix}$
Coupling coefficient [μH]	0.093534	0.093533	0.093236

3.3 The Assessment of Two Truncated Cone-Like Coils, with Non-Parallel Bases with One Coil Rotated 45^0.

The two coils (see Fig. 5) have the following geometrical parameters:

- Turn's initial radius $r = 10$ mm
- Pitch on OY axis $p = 0.9$ mm
- Section of the conductors $s = 1.2 \times 0.8 = 0.96$ mm^2
- Distance between the two coils along Oz axis $h = 20$ mm
- Number of turns $N_1 = N_2 = 15$.

The identification of electric parameters (i.e. ohmic resistances, parasitic capacitances, self-inductivities as well as the coupling coefficient) filled in for Table 3

Table 3 The electric parameters for two truncated cone-like coils with non-parallel bases, one coil rotated 45^0

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.20907 & 0.029635 \\ 0.029636 & 0.20839 \end{pmatrix}$	$\begin{pmatrix} 0.20907 & 0.029635 \\ 0.029636 & 0.20839 \end{pmatrix}$	$\begin{pmatrix} 2.9425 & 0.41043 \\ 0.41043 & 3.0142 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 6.7639 & 0.63251 \\ 0.63251 & 6.7609 \end{pmatrix}$	$\begin{pmatrix} 6.8438 & 0.26477 \\ 0.63748 & 6.7874 \end{pmatrix}$	$\begin{pmatrix} 6.6993 & 0.62623 \\ 0.62623 & 6.6878 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 3.1309 & -1.7213 \\ -1.7213 & 3.1325 \end{pmatrix}$	$\begin{pmatrix} 2.376 & -1.0406 \\ -1.0406 & 2.3759 \end{pmatrix}$	$\begin{pmatrix} 2.376 & -1.0406 \\ -1.0406 & 2.3759 \end{pmatrix}$
Coupling coefficient [μH]	0.092983	0.092983	0.093557

Table 4 The electric parameters of 2 truncate cone-like coils, with parallel bases, configuration B

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.2042 & 0.0056 \\ 0.0056 & 0.2031 \end{pmatrix}$	$\begin{pmatrix} 2.0238 & 0.05731 \\ 0.057309 & 2.0168 \end{pmatrix}$	$\begin{pmatrix} 2.8266 & 0.0802 \\ 0.0802 & 2.8088 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 6.7052 & 0.12503 \\ 0.1250 & 6.7051 \end{pmatrix}$	$\begin{pmatrix} 6.6964 & 0.6213 \\ 0.1244 & 6.6968 \end{pmatrix}$	$\begin{pmatrix} 6.7280 & 0.6262 \\ 0.1260 & 6.7324 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 2.1893 & -0.7505 \\ -0.7505 & 2.1895 \end{pmatrix}$	$\begin{pmatrix} 2.1893 & -0.7505 \\ -0.7504 & 2.1895 \end{pmatrix}$	$\begin{pmatrix} 2.1893 & -0.7505 \\ -0.7505 & 2.1895 \end{pmatrix}$
Coupling coefficient [μH]	0.01865	0.01863	0.01871

3.4 The Assessment of Two Truncated Cone-Like Coils, with Parallel Bases, Configuration B

The two coils (see Fig. 6) have the following geometrical parameters:

- Turn's initial radius $r = 10$ mm
- Pitch on OY axis $p = 0.9$ mm
- Section of the conductors $s = 1.2 \times 0.8 = 0.96$ mm^2
- Distance between the two coils along Oz axis $h = 20$ mm
- Number of turns $N_1 = N_2 = 15$.

From the four configurations of the two magnetically coupled coils under investigation, one can remark the highest mutual inductivity of $M = \mu H$ (see Table 1), a fact which indicates the configuration from 20.2.1 as the optimum. MATLAB software package, as a universal standard for scientific applications, proved very attractive for the elaboration of efficient procedures aiming for the numerical calculation of the mutual inductance corresponding to the two magnetically coupled coils [33–41].

4 The Optimization of Parameters

The most straightforward configuration of a WPT system consists of two magnetically coupled coils (transmitter and receiver). This configuration can be *series – series (ss), parallel – parallel (pp), series–parallel (sp) and parallel – series (ps).* The transmitter releases a non-radiant magnetic field with an oscillation frequency in the full range of *30 kHz – 40 MHz.* In this way, between the receiver and transmitter (*assumed in magnetic resonance*), it is ensured an efficient power transfer. A generally accepted statement from the literature [1–10], claims that the magnetic interaction between

the transmitter (source) and the receiver must be *strong enough*, making it possible to neglect the other magnetic interactions with non-resonant objects.

The result is an efficient channel for WPT. Available studies [4, 6] show that WPT doesn't change noticeably, in the presence of people or various objects placed between the two coils or in their close neighborhood. This statement is also valid for the case when these objects completely cover the direct line between the transmitter and the receiver. Some materials, such as aluminum, can produce changes in the resonance of their resonance frequency, but the correction of this issue is relatively easy by using a reaction circuit. WPT offers the possibility of connecting the electronic devices without wires, a fact that determines a reduction in the dimensions of the equipment, accompanied by a cost reduction. Both the amount of transferred between two coils and the WPT efficiency strongly depend on their parameters.

The *maximum transmitted active power* or the *maximum transmission efficiency* counts as criteria meant to determine the optimum values for the magnetically coupled coils. The next section contains a presentation of two optimization methods, based on the *transfer functions*, respectively, the *output square error.* The MATLAB functions (from MATLAB toolbox) that can be used are the minimization functions "*fminimax*", "*fminunc*" [38].

4.1 Transfer Function and Output Square Error Method Based on "fminunc" Function

The system under study consist of a system composed of two series-series resonators (magnetically coupled, no wires) used in inductive WPT as in Fig. 7.

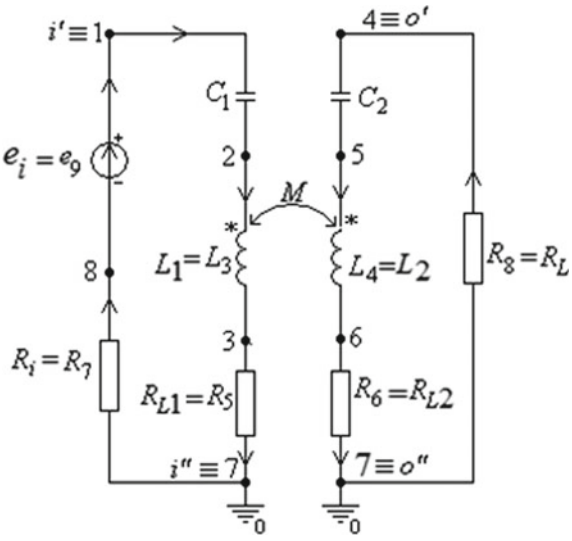


Fig. 7 Two series-series resonators magnetically coupled, powered by an AC voltage source

One can look to the circuit as being a two-port system operating in the frequency domain. There are several possibilities to optimize the WPT for this system, utilizing one of the functions: *useful active power transmitted to the load* (APL), *power transmission efficiency*, *signal transmission efficiency*, or *any other transfer function* [39, 40]. From the equivalent scheme of an analog two-port system in the AC regime, one can achieve any complex transfer function *H(jw)*, expressed either in an entirely *symbolic form*, *partially symbolic* or *numeric*. The measurements are targeting the absolute value of the phase angle imposed by the system energization from a voltage source with variable frequency capabilities. *H(f)* is the transfer function (or the output quantity) in an entirely symbolic form using the software developed and explained in [36, 37]. The system's parameters chosen for optimization or identification are $x_1, x_2, \ldots, x_p$ (p - the number of unknowns). The remained $(n-p)$ parameters come as nominal values taken from the catalog. Whereas considering k frequency samples at which the circuit function is available through measurements or simulations (the output function or any other circuit performance quantity), one can formulate the following objective function as in (13) and (14):

$$g_j(x_1, x_2, \ldots, x_p, f_j) = \left(\left|\underline{H}(f_j)\right| - \left|\underline{H}(f_j, x_1, x_2, \ldots x_p)\right|\right)^2,\ j = \overline{1,k} \tag{13}$$

$$f(x_1, x_2, \ldots, x_p) = \sum_{j=1}^{k} \left(\left|\underline{H}(f_j)\right| - \left|\underline{H}(f_j, x_1, x_2, \ldots, x_p)\right|\right)^2 \tag{14}$$

For a scalar objective function, k can be 1. When the objective function is the useful power P_2 or the efficiency $\eta = (P_2/P_1)100$ $\eta_{21} = \frac{P_2}{P_1} \cdot 100$, then the objective function has one of the following structures (15):

$$f(x_1, x_2, \ldots, x_p) = \sum_{j=1}^{k} \left(P_2(x_1, x_2, \ldots, x_p, f_j) - P_2(f_j)\right)^2 \tag{15a}$$

$$f(x_1, x_2, \ldots, x_p) = \sum_{j=1}^{k} \left(\eta_{21}(x_1, x_2, \ldots, x_p, f_j) - \eta_{21}(f_j)\right)^2, k \le 3 \tag{15b}$$

$$f(x_1, x_2, \ldots, x_p) = \sum_{j=1}^{k} \left(P_2(x_1, x_2, \ldots, x_p, f_j) - P_2(f_j)\right) \tag{15c}$$

$$f(x_1, x_2, ..., x_p) = \sum_{j=1}^{k} \left(\eta_{21}(x_1, x_2, ..., x_p, f_j) - \eta_{21}(f_j)\right), k \leq 3 \quad (15d)$$

In conditions of constant (fixed) frequency, the optimization functions become (16)

$$f(x_1, x_2, ..., x_p) = \left(P_2(x_1, x_2, ..., x_p, f) - P_2(f)\right)^2 \quad (16a)$$

$$f(x_1, x_2, ..., x_p) = \left(\eta_{21}(x_1, x_2, ..., x_p, f) - \eta_{21}(f)\right)^2 \quad (16b)$$

$$f(x_1, x_2, ..., x_p) = \left(P_2(x_1, x_2, ..., x_p, f) - P_2(f)\right) \quad (16c)$$

$$f(x_1, x_2, ..., x_p) = \left(\eta_{21}(x_1, x_2, ..., x_p, f) - \eta_{21}(f)\right) \quad (16d)$$

Usually, the quantity *H(f)* is a rational function which depends on frequency. The coefficients of the polynomials form the numerator and denominator of the transfer function (output quantity) are involved, consisting of multiplicators of parameters. The functions *fminimax* from MATLAB toolbox [38], solve a minimax problem under certain constraints.

5 Case Study: Optimization of Magnetically Coupled Coils

5.1 *Optimization of Useful Active Power $P_2 = P_{RL}$ Depending on the Parameters L_1, L_2, and M*

One can perform the analysis of the circuit from Fig. 7 in the frequency domain. The series-series connection exemplified by the circuit from Fig. 7 is the most efficient with respect to the WPT [6, 7]. The geometric dimensions of the two coils are: $r = 150$ mm, $p = 3$ mm, $w = 2$ mm (the conductor's diameter), $g = 150\ mm$ (the distance between the coils), $N_1 = N_2 = 5$ (the turn's number). Using ANSOFT EXTRACTOR Q3D (ANSYS) [38], the numerical values of the two coils magnetically coupled (Fig. 7) are:

$$
\begin{aligned}
&PLss_f = 61342RLf^0C2^2M^2C2^2Ei^2 \Big/ \Big[F_1 f^8 + (F_2 + F_3) f^6 + F_4 f^4 + F_5 f^2 + F_6\Big] \\
&F_1 = (1 + (-0.48 * 10^7 C1^2 L1 C2^2 L2 M^2 + 0.24 C2^2 M^4 C1^2 + 0.24 C1^2 L1^2 C2^2 L2^2) \\
&F_2 = 122684(RLRL1C1^2M^2 + C1^2RiC2^2RL2M^2 - C1^2L1^2C2^2L2 + ... \\
&C1^2L1^2C2^2RL2RL + C2^2L2M^2C1 + C1^2RiC2^2L2RL1 + C1^2RiC2^2M^2 ... \\
&\quad + RL1C1^2C2^2RL2M^2 - C1L1C2^2L2^2 + C1^2L1C2M^2) \\
&F_3 = 61342(RL1^2C1^2C2^2L2^2 + C1RiC2^2L2^2 + C1^2L1^2C2^2 + C1^2L1^2C2^2RL2^2) \\
&F_4 = 3111(RL2C1^2RiC2^2RL1 - C2^2M2C1 + RLRL1^2C1^2C2^2RL2 - RL2C1L1C2^2 \\
&... + RLC1^2RL2 - C1^2RiC2L2 + RL1C1^2C2^2RL2^2Ri - C1^2RL1^2C2L2 - C2^2RL2^2C1L1) \\
&F_5 = [79(-C1L1 - C2L2 + C1^2RiRL1 + C2RL2RL)] \\
&F_6 = 6222(C1L1C2L2 + RLC1^2RiRL1RL2 - C1^2RiRL1C2L2 - C2^2RL2RLC1L1) ... \\
&\quad + 1555(C2^2L2^2 + RL2RL1^2C1^2C2^2 + C1^2Ri2C2^2RL2^2 + RL1^2C1^2C2^2RL2^2 ... \\
&\quad + C1^2L1^2 + RL2C1^2Ri2C2^2) + 39(C2^2RL2 + C2^2RL2^2 + C1^2Ri2 + RL1^2C1^2)
\end{aligned}
\tag{17}
$$

In conditions of unknown L_1, L_2, M, and frequency f, the expression of the load transferred active power is (18)

$$
\begin{aligned}
&PLss\_L1L2M\_PRL := 0.23 \cdot 10^{-33} f^6 M^2 \Big/ \Big[1 + (T_1 + T_4) \cdot f^8 + T_2 f^6 + T_3 f^4\Big] \\
&T_1 = 0.24 \cdot 10^7 \Big(-0.75 \cdot 10^{-43} L1\, L2 M^2 + 0.37 \cdot 10^{-43} M^4 + 0.37 \cdot 10^{-43} L1^2 L2^2\Big) \\
&T_2 = \Big(0.54 \cdot 10^{-32} L2\, M^2 + 0.15 \cdot 10^{-41} L2^2 - 0.54 \cdot 10^{-32} L1\, L2^2 + 0.54 \cdot 10^{-32} L1\, M^2\Big) \\
&T_3 = 1555.39 \begin{pmatrix} 0.77 \cdot 10^{-21} L1\, L2 + 0.19 \cdot 10^{-39} - 0.39 \cdot 10^{-21} M^2 - 0.19 \cdot 10^{-21} L2 ... \\ -0.22 \cdot 10^{-30} L2 - 0.69 \cdot 10^{-30} L1 + 0.19 \cdot 10^{-21} L1^2 \end{pmatrix} \\
&T_4 = 39.44 \cdot \Big(0.33 \cdot 10^{-19} - 0.28 \cdot 10^{-10} L1 - 0.28 \cdot 10^{-10} L2\Big)
\end{aligned}
\tag{18}
$$

The generation of functional $f =$ *myfunL1L2M_PRL (x, fj)* serves the purpose of optimizing the APL $P_2 = P_{RL}$, which is a function of parameters L_1, L_2, and M. The implementation in MATLAB of the program called "main_gradient_L1L2M_PRL" must determine the values for parameters L_1, L_2, and M, corresponding to the optimal power transfer.

The program operates via routine *fminunc*, and the extracted values of L_1, L_2, and M are the optimum values. The parameters under scrutiny, subjected to the optimizing procedure, start from the following intervals of variations:

$$
\begin{aligned}
&L_1 \in [1.6e-05,\ 2.0e-05],\ L_2 \in [1.6e-05,\ 2.0e-05], \\
&M \in [1.4e-06,\ 1.8e-06]
\end{aligned}
$$

Table 5 contains the results following the running of the program mentioned above for the objective function from (15a), with a *frequency between 107 and 1.4*

Table 5 Parameters for active power optimization

Parameters for optimizing active power PRL it_max = 41; L1_max = 2.0 e-05 H L 1n = 1.6748 e-05 H L2_max = 2.0 e-05 H L 2n = 1.6735 e-05 H M_max = 1.4 e-06 H Mn = 1.4899 e-06 H k_max = 0.07 kn = 0.089 Parameters for optimizing active power PRL it_min = 9 L1_min = 1.8 e-05 H L 1n = 1.6748 e-05 H L2_min = 2.0 e-05 H L 2n = 1.6735 e-05 H M_min = 2.0 e-06 H Mn = 1.4899 e-06 H k_min = 0.1 kn = 0.089	Parameters for optimizing active power PRL it_max = 42; L1_max = 2.0 e-05 H L 1n = 1.6748 e-05 H L2_max = 2.0 e-05 H L 2n = 1.6735 e-05 H M_max = 1.4 e-06 H Mn = 1.4899 e-06 H k_max = 0.07 kn = 0.089 Parameters for optimizing active power PRL it_min = 16 L1_min = 2.0 e-05 H L 1n = 1.6748 e-05 H L2_min = 2.0 e-05 H L 2n = 1.6735 e-05 H M_min = 2.0 e-06 H Mn = 1.4899 e-06 H k_min = 0.1 kn = 0.089	Parameters for optimizing active power PRL it_max = 13; L1_max = 2.0 e-05 H L 1n = 1.6748 e-05 H L2_max = 1.8 e-05 H L 2n = 1.6735 e-05 H M_max = 1.4 e-06 H Mn = 1.4899 e-06 H k_max = 0.07 kn = 0.089 Parameters for optimizing active power PRL it_min = 41 L1_min = 2.0 e-05 H L 1n = 1.6748 e-05 H L2_min = 2.0 e-05 H L 2n = 1.6735 e-05 H M_min = 2.0 e-06 H Mn = 1.48990 e-06 H k_min = 0.1 kn = 0.089

$\times$ *107* Hz*, and a frequency step of 0.2* $\times$ *107* Hz. To obtain the results, one applied three procedures used to calculate the Hessian matrix, included in the *fminunc* routine.

Figure 8 contains the variation of the with the frequency of APL P_{Lss}, respectively, of the efficiency *eta21ss* for the application of the third procedure of using the Hessian matrix 'HessUpdate','steepdesc' to the objective function (15a) for the frequency domain previously mentioned.

Consulting Table 5, one can remark the maximum values obtained for the active power P_{Lss} when the parameters L_1, L_2 and M have optimal values, being higher than the maximum of P_{Lss} corresponding to the nominal values of these parameters. The application of the same procedure for the objective functions (15b), (c), respectively (15d) delivered equivalent results summarized in Figs. 9, 10 and 11. When is analyzing Figs. 8, 9, 10, 11, one can find similar results for all four cases involving different objective functions.

Moreover, the results corresponding to the optimal values of parameters L_1, L_2, and M are practically identical for the three procedures for using the Hessian matrix (Table 5).

Whereas performing a thorough analysis of the results from Figs. 8, 9, 10, 11, one can extract interesting remarks:

1. The phenomenon of "splitting frequency" becomes evident for both sets of values of the parameters L_1, L_2, and M. The first set of values represents the nominal ones, whereas the second set of those corresponds to the maximum value of the power. Among the frequencies corresponding to the two maximum efficiency values recorded in this case is a significant difference.

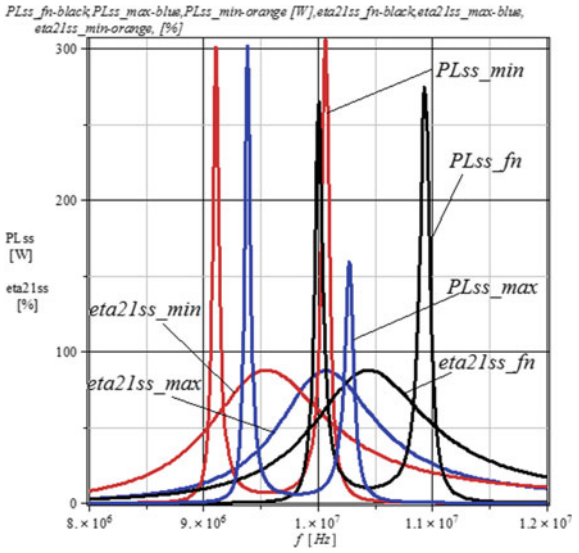


Fig. 8 Variations of APL and efficiency with the frequency for the objective function (19, a)-*optimum power*

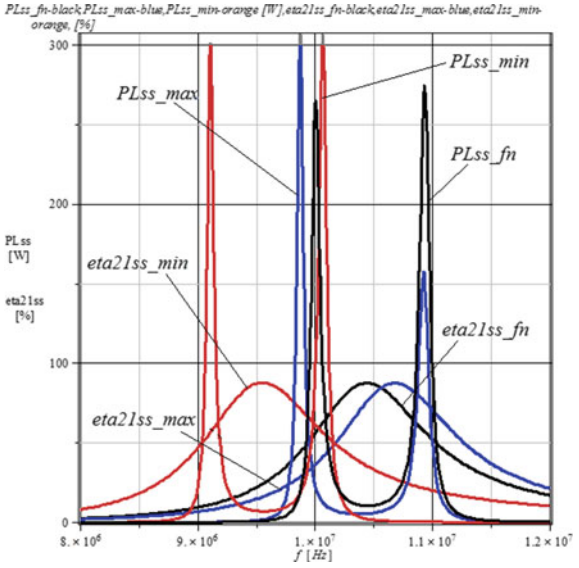


Fig. 9 Variations of APL and efficiency with frequency for the objective function (19, b)-*optimum power*

2. In the case when the parameters L_1, L_2, and M have the optimum values, the WPT efficiency has three maxima.
3. The second maximum value recorded for the APL is higher than the one corresponding to its first maximum.

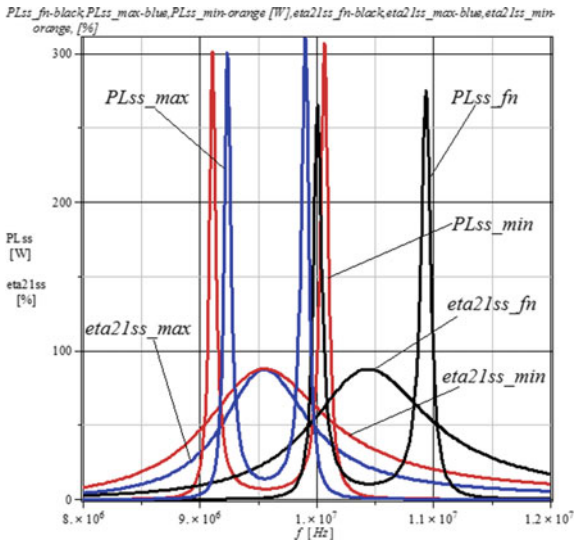


Fig. 10 Variations of APL and efficiency with the frequency for the objective function (19, c)-*optimum power*

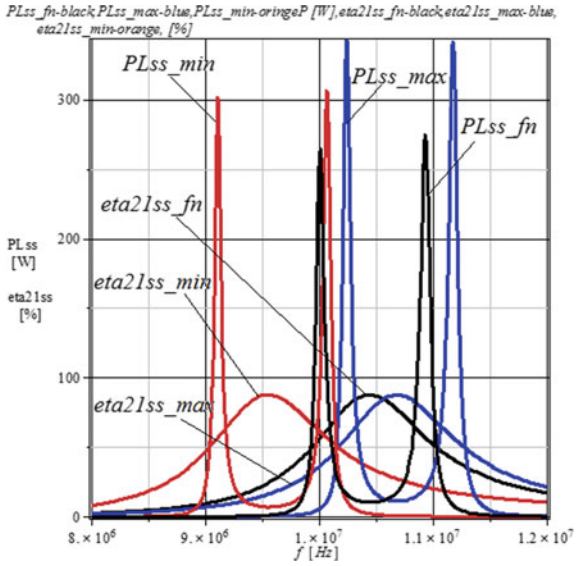


Fig. 11 Variations of APL and efficiency, with the frequency for the objective function (19, d)-*optimum power*

4. Looking at the values of the APL at the frequencies where the efficiency is maximum, one can observe a lower active power in for the optimal set of parameters L_1, L_2, and M, when compared to the case when the same parameters have nominal values.

5.2 *Efficiency Optimization Function of L1, L2, and M Parameters*

The WPT efficiency as a function of frequency, with unknown parameters L_1, L_2, and M is:

$$\begin{aligned} eta21ss\_L1L2M &:= 0.15 \cdot 10^{-15} f^4 M^2 / \left(T_5 f^4 + T_6 f^2 + 6.35\right) \\ T_5 &= 1555.39\left(0.12 \cdot 10^{-20} M^2 + 0.12 \cdot 10^{-20} L2^2\right) \\ T_6 &= 39.4384\left(-0.18 \cdot 10^{-9} L2 + 0.49 \cdot 10^{-19}\right) \end{aligned} \tag{19}$$

The generation of functional $f =$ *myfunL1L2M_eta21 (x, fj)* serves the purpose of optimizing the WPT efficiency, function of L_1, L_2, and M parameters. The implementation of the program called "main_gradient_L1L2M_eta21", must determine the values for parameters L_1, L_2, and M, corresponding to the optimal efficiency. The program, developed in MATLAB as well, operates via routine *fminunc* and the extracted values of L_1, L_2, and M are the optimum values.

The parameters under scrutiny, subjected to the optimizing procedure, start from the following intervals of variations:

$$\begin{aligned} &L_1 \in [1.6e-05,\ 1.8e-05],\ L_2 \in [1.6e-05,\ 1.8e-05], \\ &M \in [1.0e-06,\ 1.6e-06] \end{aligned}$$

Table 6 contains the results following the running of the program mentioned above for the objective function from (15a), with a *frequency between 107* Hz *and 1.4* $\times$ *107* Hz*, and a frequency step of 0.2* $\times$ *107* Hz. To obtain the results, one applied three procedures used to calculate the Hessian matrix, included in the *fminunc* routine.

Figure 12 contains the frequency variation of the active power *PLss*, respectively of the efficiency *eta21ss*, for the application of the third procedure of using the Hessian matrix 'HessUpdate','steepdesc' to the objective function (15a) for the frequency domain previously mentioned.

Consulting Table 6, one can remark the maximum values obtained for the efficiency *eta21ss* when the parameters L_1, L_2, and M have optimal values, being higher than the maximum of *eta21ss* corresponding to the nominal values of these parameters.

The application of the same procedure for the objective functions (15b), (15c), respectively (15d) delivered equivalent results summarized in Figs. 13, 14, and 15, respectively.

Analyzing Figs. 12, 13, 14, 15, one can find similar results for all four objective functions.

Moreover, the results corresponding to the optimal values of parameters L_1, L_2 and M are practically identical for the three procedures for using the Hessian matrix (Table 6).

Table 6 Parameters for efficiency optimizing

Parameters for optimizing the efficiency of eta21ss	Parameters for optimizing the efficiency of eta21ss	Parameters for optimizing the efficiency of eta21ss
it_max = 41;	it_max = 46;	it_max = 47;
L1_max = 1.6 e-05 H	L1_max = 1.8 e-05 H	L1_max = 2.0 e-05 H
L 1n = 1.6748 e-05 H	L 1n = 1.6748 e-05 H	L 1n = 1.6748 e-05 H
L2_max = 1.729 e-05 H	L2_max = 1.729 e-05 H	L2_max = 1.729 e-05 H
L 2n = 1.6735 e-05 H	L 2n = 1.6735 e-05 H	L 2n = 1.6735 e-05 H
M_max = 1.423 e-06 H	M_max = 1.423e-06 H	M_max = 1.423 e-06 H
Mn = 1.4899 e-06 H	Mn = 1.4899 e-06 H	Mn = 1.4899 e-06 H
k_max = 0.0855	k_max = 0.0806	k_max = 0.0806
kn = 0.089	kn = 0.089	kn = 0.089
Parameters for optimizing active power PRL	Parameters for optimizing active power PRL	Parameters for optimizing active power PRL
it_min = 49	it_min = 49	it_min = 46
L1_min = 1.6 e-05 H	L1_min = 1.6 e-05 H	L1_min = 1.8 e-05 H
L 1n = 1.6748 e-05 H	L 1n = 1.6748 e-05 H	L 1n = 1.6748 e-05 H
L2_min = 1.729 e-05 H	L2_min = 1.7294 e-05 H	L2_min = 1.729 e-05 H
L 2n = 1.6735 e-05 H	L 2n = 1.6735 e-05 H	L 2n = 1.6735 e-05 H
M_min = 1.423 e-06 H	M_min = 1.423 e-06 H	M_min = 1.423 e-06 H
Mn = 1.4899 e-06 H	Mn = 1.4899 e-06 H	Mn = 1.4899 e-06 H
k_min = 0.0855	k_min = 0.0855	k_min = 0.0807
kn = 0.089	kn = 0.089	kn = 0.089

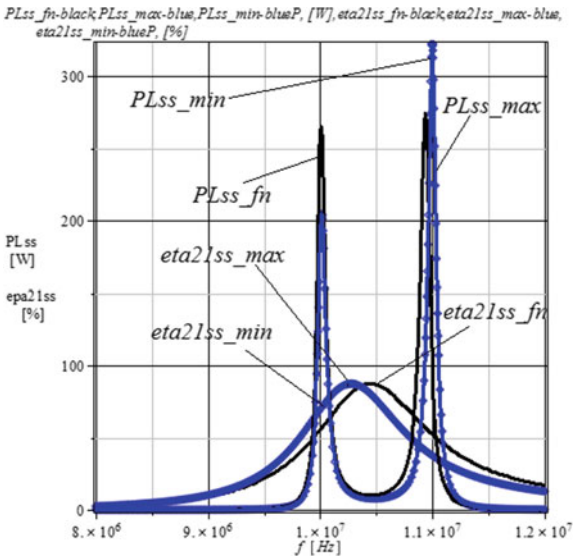


Fig. 12 The variations of the APL and efficiency with the frequency for the objective function (19, a)-*optimum efficiency*

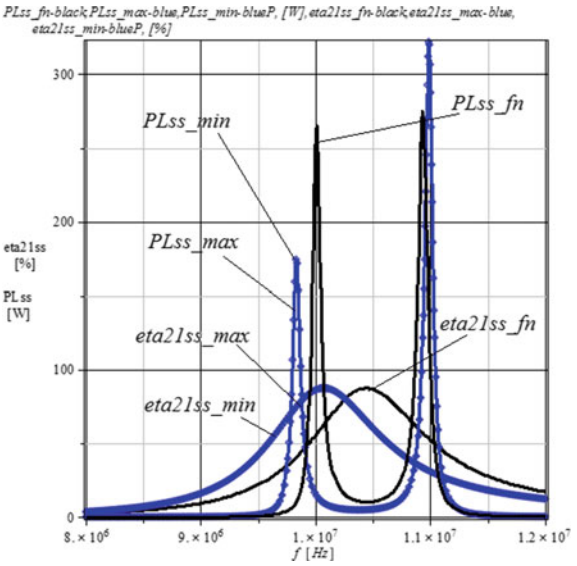


Fig. 13 The variations of the APL and efficiency with the frequency for the objective function (19, b)-*optimum efficiency*

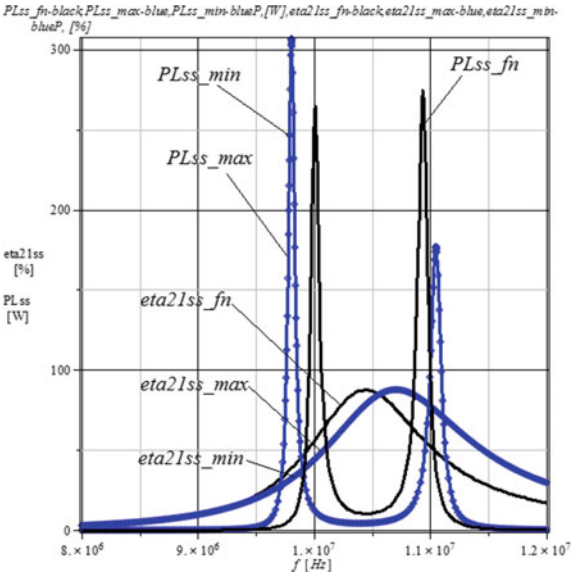


Fig. 14 The variations of the transmitted active power and efficiency with the frequency for the objective function (19, c)-*optimum efficiency*

Performing a thorough analysis of the results from Figs. 12, 13, 14, 15, one can extract interesting remarks:

1. The phenomenon of "splitting frequency" becomes evident for both sets of values of the parameters L_1, L_2, and M. The first set of values represents the nominal

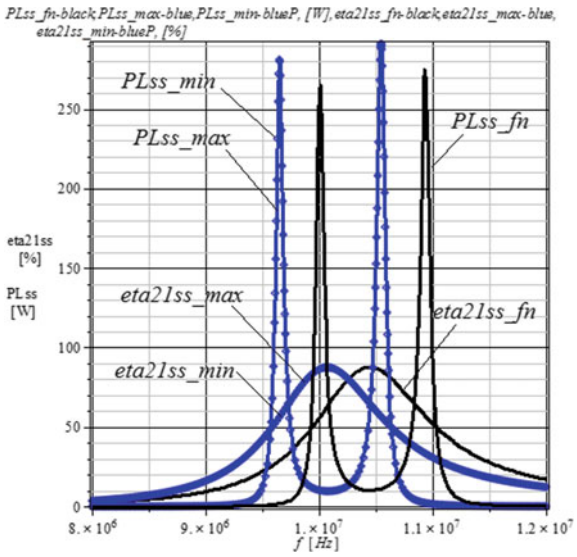


Fig. 15 The variations of the transmitted active power and efficiency with the frequency for the objective function (19, d)-*optimum efficiency*

ones, whereas the second set of those corresponds to the maximum value of the WPT efficiency.

2. In the case when the parameters L_1, L_2 and M have the optimum values, the WPT efficiency has three maxima.
3. The second maximum value recorded for the WPT efficiency is identical to the maximum WPT efficiency corresponding to the nominal values of the parameters.
4. The maximum values of the active power are lower for the optimal values of the parameters providing maximum efficiency than those corresponding to the nominal parameters.
5. The frequencies corresponding to the two maxima of the active power, for optimal parameters case, are different from those corresponding to the nominal parameters one.

5.3 Structure's Optimization Using ANSYS Q3D Extractor

When using ANSYS Q3D Extractor software, the mesh is essential for the accuracy of the solution. A major drawback of the program is given by the fact that the initial network can face convergence problems of an increased computational effort. The advantage of the mesh in ANSYS Q3D Extractor program is that the discretization of the coil takes place in the area of interest. So, it is no need to discretize the air in the vicinity of the coil or use conditions indefinitely. The resulting RLC matrices let us generate new matrices for any named parameters with field solutions, without having to compute new field solutions. Besides, any model successfully solved can have its matrix results exported to an equivalent circuit for further signal integrity

analysis. After the field solutions and matrix calculations are complete, the simulator performs an error analysis in each element in the mesh. Upon the next adaptive pass, there is a refinement process targeting the elements with the highest error. Therefore, in those areas, the obtained solution is more accurate.

The study from this section denotes the utilization of the program ANSYS Q3D EXTRACTOR [30–33] to simulate the WPT between differently configured two coils at three frequencies: *50* kHz, *5000* kHz, and *10,000* kHz. The best four configurations (cases) selected from a preliminary analysis are:

1. *Two spiral parallel coils*
2. *Two truncated cone-like coils with parallel bases, configuration A*
3. *Two truncated cone-like coils with non-parallel bases*
4. *Two truncated cone-like coils with parallel bases, configuration B*

1. *Two spiral parallel coils.* The constructive characteristics of the configuration: turn's initial radius $r = 10$ mm, pitch $p = 0.21$ mm, wire Sect. 0.2×0.2 mm^2, distance between coils $z = 20$ mm, and number of turns $N = 15$ (see Fig. 16). The determination of parameters for the configuration displayed in Fig. 16, using ANSYS Q3D EXTRACTOR [30–33] resulted in the values filling Table 7.
2. *Two truncated cone coils-like coils with the parallel bases, configuration A.* The constructive characteristics: turn's initial radius of a turn $r = 10$ *mm*, the distance between the centres of the two consecutive turns *0.21* mm, the distance between the centres of the two consecutive turns on the 0Y axis *0.21 mm*, the wire Sect. 0.2×0.2 mm^2, the distance between the coils $z = 20$ mm, and the number of turns $N = 15$. (see Fig. 17).

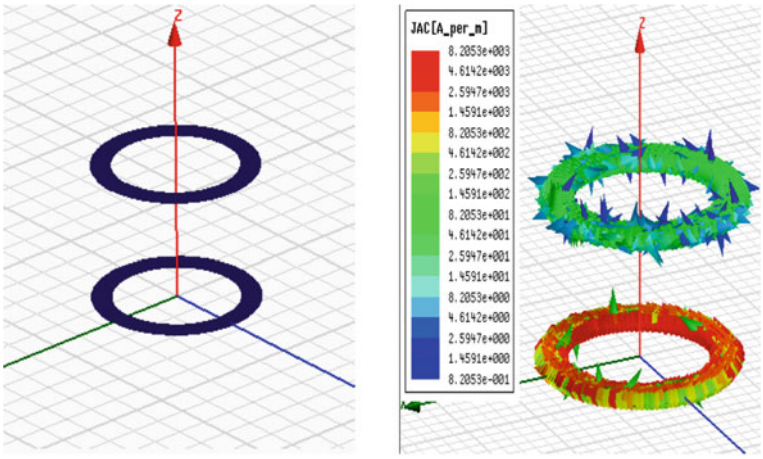


Fig. 16 **a** Configuration of the 2 spiral parallel coils; **b** Current density surface

Table 7 Results for spiral parallel coils

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.1635 & 0.0004 \\ 0.0009 & 0.1790 \end{pmatrix}$	$\begin{pmatrix} 2.1451 & 0.0394 \\ 0.0394 & 2.3414 \end{pmatrix}$	$\begin{pmatrix} 7.7438 & 0.1892 \\ 0.1892 & 7.6332 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 0.9601 & -0.2977 \\ -0.2977 & 0.9602 \end{pmatrix}$	$\begin{pmatrix} 0.9601 & -0.2977 \\ -0.2977 & 0.9602 \end{pmatrix}$	$\begin{pmatrix} 0.9601 & -0.2977 \\ -0.2977 & 0.9602 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 8.8699 & 0.4971 \\ 0.4971 & 8.8821 \end{pmatrix}$	$\begin{pmatrix} 8.5028 & 0.4950 \\ -0.75045 & 8.5032 \end{pmatrix}$	$\begin{pmatrix} 8.7034 & 0.4927 \\ 0.4927 & 8.7154 \end{pmatrix}$

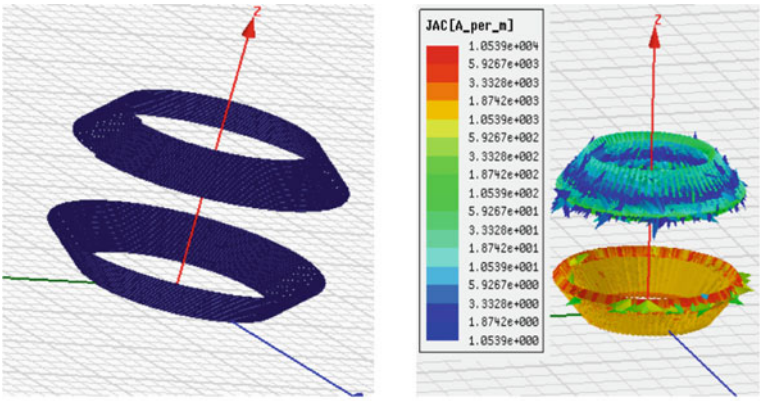


Fig. 17 **a** Truncated cone-like coils with the parallel bases, configuration 1; Current density surface

The determination of the parameters for the configuration displayed in Fig. 17, using ANSYS Q3D EXTRACTOR [30–33] resulted in the values filling Table 8.

3. *Two truncated cone-like coils with the nonparallel bases*, one of them is rotated with an angle of 45^0. The constructive characteristic of the configuration are: the initial radius of a turn $r = 10$ mm, the distance between the centres of the two consecutive turns 0.21 mm, the distance between the centres of the two consecutive turns on the 0Y axis 0.21 mm, the wire Sect. $0.2 \times 0\ 0.2\ mm^2$, the distance between the coils $z = 20\ mm$, and the turn number N = 15 (see Fig. 18). The determination of parameter for the configuration displayed in Fig. 18, using ANSYS Q3D EXTRACTOR [30–33] resulted in the values filling the Table 9.
4. *Two truncated cone-like coils with the parallel bases, configuration 2*. The constructive characteristics of the configuration are: the distance on the 0Z: $z = 20$ mm and the distance on the 0Y is $y = 15$ mm, the coil's initial radius of

Table 8 Results obtained for configuration A

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.1826 & 0.0012 \\ 0.0012 & 0.1807 \end{pmatrix}$	$\begin{pmatrix} 2.4503 & 0.0772 \\ 0.0772 & 2.5167 \end{pmatrix}$	$\begin{pmatrix} 4.1709 & 0.1374 \\ 0.1374 & 4.0615 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 1.3023 & -0.5986 \\ -0.5986 & 1.3031 \end{pmatrix}$	$\begin{pmatrix} 1.3023 & -0.5986 \\ -0.5986 & 1.3031 \end{pmatrix}$	$\begin{pmatrix} 0.9601 & -0.2977 \\ -0.2977 & 0.9602 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 13.348 & -1.9729 \\ -1.9729 & 13.339 \end{pmatrix}$	$\begin{pmatrix} 13.0880 & -1.9689 \\ -0.5986 & 13.173 \end{pmatrix}$	$\begin{pmatrix} 13.4930 & -1.9727 \\ -1.9727 & 13.5310 \end{pmatrix}$

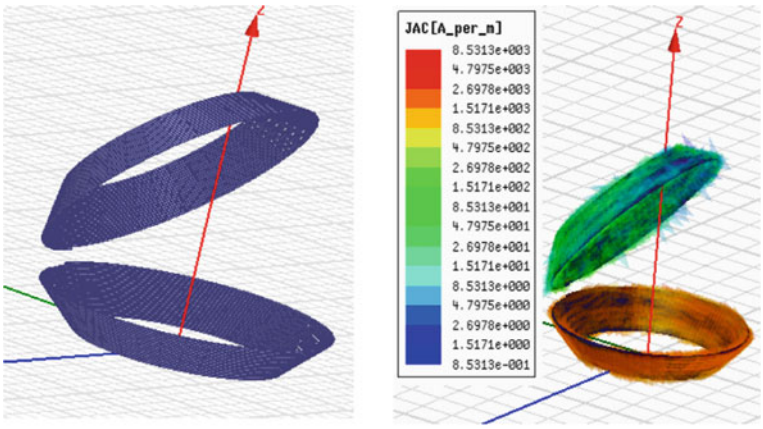


Fig. 18 **a** Truncated cone-like coils with nonparallel bases, one of them are rotated with an angle of 45^0; **b** Current density surface

Table 9 The results obtained for the truncated cone-like coils, with the nonparallel bases, one of them is rotated with an angle of 45°

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 0.1846 & 0.0015 \\ 0.0015 & 0.1787 \end{pmatrix}$	$\begin{pmatrix} 2.3284 & 0.0196 \\ 0.0196 & 2.3955 \end{pmatrix}$	$\begin{pmatrix} 4.5941 & 0.2569 \\ 0.2569 & 4.9967 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 1.3743 & -0.6593 \\ -0.6593 & 1.3667 \end{pmatrix}$	$\begin{pmatrix} 1.3743 & -0.6593 \\ -0.6593 & 1.3667 \end{pmatrix}$	$\begin{pmatrix} 1.3743 & -0.6593 \\ -0.6593 & 1.3667 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 13.3680 & 1.8555 \\ 1.8555 & 13.3730 \end{pmatrix}$	$\begin{pmatrix} 13.2090 & -1.9689 \\ 1.8560 & 13.0710 \end{pmatrix}$	$\begin{pmatrix} 13.5420 & 1.8485 \\ 1.8485 & 13.4200 \end{pmatrix}$

the turn is $r = 10$ mm, the distance between the centres of the two consecutive turns 0.21 mm, the section of the wire 0.2×0.2 mm^2 and the number of turns $N = 15$ (see Fig. 19).

The determination of the parameters for the configuration displayed in Fig. 19, using ANSYS Q3D EXTRACTOR [30–33] resulted in the values filling Table 10.

Due to the highest value of the mutual inductance, $M = 1.9729$ μH (Table 9) between the two magnetic coupled coils (resonators), the second configuration (see Fig. 17) appears to be the best possible out of the four studied configurations of coils. Whereas assessing the parameter values determined through ANSYS Q3D EXTRACTOR simulations performed at the three mentioned frequencies, one can conclude that the values of the parameters (L, M, C) remain unchanged, yet the

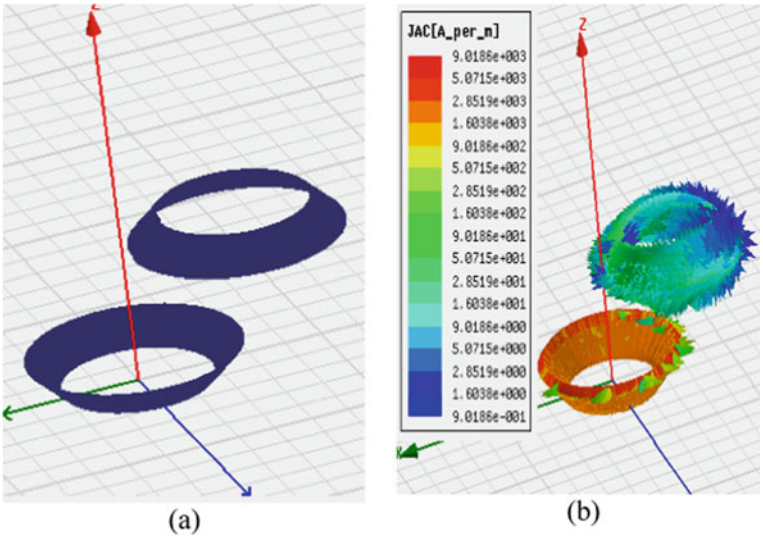


Fig. 19 **a** Truncated cone-like coils with the parallel bases, configuration B, **b** Surface current density

Table 10 Results obtained for configuration B

Frequency f [kHz]	50	5000	10,000
Resistance matrix [ohm]	$\begin{pmatrix} 183.3 & 0.9199 \\ 0.9199 & 181 \end{pmatrix}$	$\begin{pmatrix} 2720 & 32.731 \\ 32.731 & 2144.7 \end{pmatrix}$	$\begin{pmatrix} 4655.2 & 132.27 \\ 132.27 & 4715.4 \end{pmatrix}$
Capacity matrix [pF]	$\begin{pmatrix} 1.1973 & -0.4423 \\ -0.5523 & 1.975 \end{pmatrix}$	$\begin{pmatrix} 1.1973 & -0.4423 \\ -0.5523 & 1.975 \end{pmatrix}$	$\begin{pmatrix} 1.1973 & -0.4423 \\ -0.5523 & 1.975 \end{pmatrix}$
Inductivity matrix [μH]	$\begin{pmatrix} 13.366 & 0.6389 \\ 0.6389 & 13.347 \end{pmatrix}$	$\begin{pmatrix} 13.183 & 0.6387 \\ 0.6387 & 13.185 \end{pmatrix}$	$\begin{pmatrix} 13.4930 & 0.6357 \\ 0.6357 & 13.489 \end{pmatrix}$

Table 11 The mutual inductance: Values obtained through simulation versus measured

Method	1st Case	2nd Case	3rd Case	4th Case
M_m [μH]	0.5004	1.9875	1.8844	0.6482
M_p [μH]	0.4985	1.9862	1.876	0.6445
MQ3D [μH]	0.4971	1.9729	1.8555	0.6389

resistance sharply increases with the frequency due to the skin, respectively the proximity effect.

Table 11 contains the values of the mutual inductance (by measurements (M_m)), computed by integrating (M_p) in MATLAB, respectively and the ones obtained through simulations performed with ANSYS Q3D EXTRACTOR (*MQ3D*).

The values of mutual inductances determined with all three methods are very close for all four cases of coil configurations, suggesting compatibility of the three methods of determination.

6 Conclusions

The WPT's efficiency, function of frequency, is greatly influenced by the resonator parameters (L – self-inductance, M – mutual inductance, C – parasitic capacitance, and R – Ohmic resistances) of the coils, placed at different distances and angles in assemblies with several configurations. Consequently, parameter identification represents an essential objective during the configuration design stage.

The determination of the best possible configuration from a pool of four relied on the utilization of the ANSYS Extractor Q3D. The computation took place for the corresponding matrices for different configurations, structures, frequencies and distances between the resonators. Each configuration maintained the same relative position between the coils, the same number of turns, the same geometrical dimensions, and materials for conductors. The selected configuration, declared optimal, holds the highest mutual inductance between the coupled magnetic coils.

The optimization study had two directions. In the first direction of the study, one assessed the optimal parameter determination for the given configuration of two magnetically coupled coils, following two different outcomes: the maximum power transfer, respectively the maximum efficiency. The simulated tests involved the presence of four different objective functions for each of the outcomes and revealed interesting conclusions included together with the results. The second direction from the final part of the optimization study represented a simulation performed utilizing the ANSYS Q3D Extractor software. The output of the simulation consisted of sets of parameters as self-inductance, mutual inductance, capacitance, and resistance as a function of frequency. The coil resistance presents a strong dependency on the frequency due to the skin and proximity effects, whereas the other parameters maintain their constant values against frequency. Some factors are exercising a significant

influence on the mutual inductance parameter. They are the relative distance between the coils and the angle at which the field lines go through the receiver coil. The mutual inductance is a crucial parameter concerning the performances of WPT systems (i.e., the load power, the efficiency).

The mutual inductances computed with MATLAB utilizing the integration, the numerically calculated ones using ANSYS Q3D Extractor, and respectively those obtained through measurements show close values, indicating a consistency regarding all three methods of parameter determination.

References

1. Agbinya JI (2012) Wireless power transfer. River Publishers Series in Communications, 9000, Aalborg, Denmark
2. Kim KY (2011) Wireless power transfer—principles and engineering explorations. Published by InTech Janeza Trdine 9, 51000 Rijeka, Croatia
3. Zhang Z, Pang H, Georgiadis A, Cecati C (2019) Wireless power transfer—an overview. IEEE Trans Ind Electron 66(2):1044–1058
4. Zhang F, Liu L, Hackworth SA, Sclabassi RJ, Sun M (2009) In vitro and in vivo studies on wireless powering of medical sensors and implantable devices. In: Proceedings of life science systems and applications workshop. IEEE Xplore, 978-1-4244-4293-5/09/2009, pp 84−87
5. Karalis A, Joannopoulos JD, Soljačić M (2008) Efficient wireless non-radiative mid-range energy transfer. Ann Phys 323:34–48
6. Kurs A, Karalis A, Moffatt RA, Joannopoulos JD, Fisher P, Soljačić M (2007) Wireless power transfer via strongly coupled magnetic resonances. Science 317(5834):83–86
7. Moffatt RA (2009), Wireless transfer of electric power. Thesis for bachelor of science in physics (under the supervision of Soljačić M), MIT, Cambridge, MA, US
8. Tesla N (1914) Apparatus for transmitting electrical energy. US patent number1119732
9. https://www.witricity.com
10. https://electronicsca.blogspot.com/2015/07/wireless-power-transmission-market.html
11. Ochiana G, Ochiana L, Stănculescu M (2012) Electromagnetic field theory, Editura Printech, ISBN 978-606-521-879-6, 202
12. Jiang HC, Wang YC (2008) Capacity performance of an inductively coupled near field communication system. Proc IEEE Int Symp Antenna Propag Soc 1:1–4
13. Imura T, Hori Y (2011) Maximizing air gap and efficiency of magnetic resonant coupling for wireless power transfer using equivalent circuit and Neumann formula. IIEEE Trans Ind Electron 58(10):4746–4752
14. Cheon S, Kim YH, Kang SY, Lee ML, Lee JM, Zyung T (2011) Circuit model based analysis of a wireless energy transfer system via coupled magnetic resonances. IEEE Trans Ind Electron 58(7):2906–2914
15. Imura T, Okabe H, Hori Y (2009) Basic experimental study on helical antennas of wireless power transfer for electric vehicles by using magnetic resonant couplings. In: Proceedings of vehicle power and propulsion conference, pp 936–940. IEEE Xplore 978-1-4244-2601-4/010/2010
16. Mercier PP, Chandrakasan AP (2013) Rapid wireless capacitor charging using a multi-tapped inductively secondary coil. IEEE Trans Circuits Syst-I 60(9):2263–2272
17. Sample AP, Meyer DA, Smith JR (2011) Analysis, experimental results, and range adaptation of magnetically coupled resonators for wireless power transfer. IEEE Trans Ind Electron 58(2):544–554

18. Ricano JA, Rodriguez Torres H, Vazquez Leal H, Gallardo del Angel A (2010) Experiment about wireless energy transfer. In: 1st international congress on instrumentation and applied sciences, Cancun, Mexico, pp 1–10. Oct 26–29 2010
19. Bobaru L (2017) Magnetically coupled resonator coils optimization used in wireless power transfer systems, PhD Thesis, University Politehnica of Bucharest, Romania
20. Iordache M, Bucată V, Niculae D, Stănculescu M, Bobaru LM (2108) Design of wireless electromagnetic energy transfer systems, annals of the university of Craiova. Electr Eng Ser 42(1):15–20
21. Niculae D, Stănculescu M, Mihai Iordache M, Bobaru ML (2018) An analysis on efficiency of wireless transfer energy due to a misalignment of two coils annals of the university of Craiova. Electr Eng Ser 42(1):21–25
22. Stănculescu M, Bobaru L, Iordache M, Niculae D, Bucată V (2018) Wireless Power Transfer Analysis Using Scattering Parameters annals of the university of Craiova. Electr Eng Ser 42(1):26–31
23. Iordache M, Andronescu G, Bucata V, Iordache (Bobaru) ML, Stanculescu M, Niculae D, (2016) Design and simulation of wireless power transfer systems, annals of the university of Craiova. Electr Eng Ser 40(1):109–114
24. Bibirică C, Sandu C, Ene L, Iordache M (2017) Improving the performance of PCB inductors for WPT systems using magnetic shields. In: Proceedings of 5th international symposium on electrical and electronics engineering (ISEEE). Galați, Romania, pp 1–7. Oct 20–22 2017
25. Iordache (Bobaru) L, Iordache M (2016) Optimization of magnetic coupled resonator structure used in wireless electromagnetic energy transfer, U.P.B. Sci Bull Ser C – Electr Eng 78(4):137–148
26. Stanculescu M, Iordache M, Niculae D, Bobaru L, Bucata V (2018) Algorithm for computing s parameters and their use for studying efficiency of electromagnetic energy wireless transfer systems. Rev Roumaine Sci Tech. Serie Electrotechnique et Energetique, 63(2):138–144
27. Stanculescu M, Iordache M, Niculae D, Iordache Bobaru L, Bucata V (2016) Parameter computation and their use for electromagnetic energy wireless transmission. Int J of Comput Technol 15(9)
28. Iordache M (2015) Symbolic, numeric—symbolic and numeric simulation of analog circuits—user guides. MATRIX ROM, Bucharest, Romania
29. Niculae D, Iordache M, Dumitriu L (2011) Magnetic coupling analysis in wireless transfer energy, the 7th international symposium on advanced topics in electrical engineering (ATEE), May 12–14 2011. Bucharest, Romania, pp 1–4
30. Ansoft Q3D Extractor, *User Guide*, https://www.ansoft.com
31. ******Geting Started with Q3D Extractor A 3D PCB Via Model, ANSYS,*** Inc. 275 Technology Drive, Canonsburg, PA 15317 USA, 2011
32. https://www.ansoft.com/ansoftdesignersv
33. https://www.ansys.com/Products/Simulation+Technology/Systems+&+Embedded+Software/ANSYS+Simplorer.
34. Iordache M, Dumitriu L (2014) Computer-aided simulation of analog circuits—algorithms and computational techniques, Editura POLITEHNICA Press, vol. I and II, Bucharest, Romania
35. Iordache M (2016) Special works of electrotehnics. MATRIX ROM, Bucharest, Romania
36. Iordache M, Dumitriu L, Delion D (2000) SESYMGP-state equation symbolic generation program. Politehnica University of Bucharest, Romania, User Guide. Library of Electrical Department
37. Iordache M, Dumitriu L, Matei I (2002) SYMNAP– symbolic modified nodal analysis program. Politehnica University of Bucharest, User Guide. Library of Electrical Department
38. ******Optimization Toolbox™ User's Guide***, Matlab R2011b, The Math Works, Inc., 2011
39. Orasanu A, Dragomir A, Bobaru L, Iordache M, Deleanu S (2018) On optimization of wireless power transfer systems, 2018 international symposium on fundamentals of electrical engineering (ISFEE), Nov 1–3 2018. Bucharest, Romania, pp 1–6
40. Iordache M, Dumitriu L, Niculae D (2016) Chapter 1 "Power transfer by magnetic induction using coupled-mode theory", Book chapter in the book, *Wireless Power Transfer*, 2nd edn, Editor J. I. Agbinya, River Publishers Series in Communications, Denmark, (1), pp 1–69

41. Iordache M, Niculae D, Ene L, Sandu C, Bobaru L, Bibirică C (2017) On the procedures for optimal wireless energy transfer systems. In: Proceedings of electric vehicle international conference and show. Bucharest. Romania, pp 1–6. Oct 5–6 2107

Numerical Assessment of Electromagnetic Energy and Forces in Non-destructive Measurement Devices

Marilena Stanculescu, Paul Cristian Andrei, Horia Andrei, Sorin Deleanu, and Lavinia Bobaru

Abstract Non-destructive testing in the electromagnetic field is one of the fastest and least expensive testing techniques for pieces subject to degradation. This domain has evolved a lot in recent years, because of the increasing demands received by the scientific community from the industry. Non-destructive testing aims to detect defects (different types of cracks, structural inhomogeneity) in materials (conducting, ferromagnetic) without destroying the tested object. Therefore, the application of such techniques addresses many relevant domains that require high security of installations, domains such as aeronautical, nuclear, medical, or chemical industry. This chapter provides insight into the most commonly used non-destructive measurement devices, but also into the magnetic field analysis in nonlinear media. It presents the magnetization characteristic evaluation for ferromagnetic bodies.

Keywords Non-destructive testing · Numerical methods · Magnetic field · Measurement devices

M. Stanculescu (✉) · P. C. Andrei · L. Bobaru
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania
e-mail: marilena.stanculescu@upb.ro

P. C. Andrei
e-mail: paul.andrei@upb.ro

L. Bobaru
e-mail: lavinia.bobaru@upb.ro

H. Andrei
Department of Doctoral School, University Valahia, Targoviste, Romania
e-mail: hr_andrei@yahoo.com

S. Deleanu
Northern Alberta Institute of Technology, Edmonton, Canada
e-mail: sorind@nait.ca

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_21

Nomenclature

A. Acronyms

NDT	Non-destructive testing
FEM	Finite Element Method
FEM-BEM	Finite Element Method—Boundary Element Method
SST	Single Sheet Tester
VSM	Vibrating Sample Magnetometer
EMF	Electromotive force
PFPM	Polarization of the Fixed-Point Method

B. Symbols/Parameters

B	Magnetic flux density (induction)
H	Magnetic field strength (intensity)
D	Electric displacement
E	Electric field strength
ε	Absolute permittivity
μ	Absolute permeability
Ψ_{S_Γ}	Electric flux
Φ_{S_Γ}	Magnetic flux
P	Power
X_k	Generalized force
dx_k	Elementary variation
dL_k	Elementary mechanical work
W	Electromagnetic energy
w	Volumetric density of electromagnetic field energy
W_e	Energy of the electric field
W_m	Energy of magnetic field
$w_{h,cycle}$	Volumetric density of the energy transferred by the electromagnetic field to bodies during a full hysteresis cycle, electric and magnetic
w^*	Volume density of electromagnetic field coenergy
w_e^*	Volumetric density of electric coenergy
w_m^*	Volumetric density of magnetic coenergy
p_h	Volumetric density of power transferred by the field
$\overline{S}$	Poynting vector
$\mathbf{J}$	Current density
i	Current
(Σ)	Closed surface
(V_Σ)	Domain bounded by the closed surface (Σ)
(Γ_{el})	Closed contour corresponding to electric hysteresis cycle
$\left(\Gamma_{mg}\right)$	Closed contour corresponding to magnetic hysteresis cycle
$\mathrm{A_{el}}$	Area of electric cycle
$\mathrm{A_{mg}}$	Area of magnetic cycle
$\overline{dA}$	Area element

1 Introduction to Non-destructive Testing

Non-destructive testing(NDT) plays a fundamental part by ensuring the fact that pieces of equipment belonging to a particular structure perform their specific functions for a predetermined amount of time. The specialists in the non-destructive testing field have created and implemented tests to characterize the materials or to detect, localize and measure the flaws (defects) which can case plane crushes, nuclear power plant explosions, train derailment, fires and q whole range of events, less visible, but with very dangerous and with high impact [1].

Because NDT does not, affect in any way the integrity of the product under test, its utilization on a large scale in controlling the quality of the product [2] becomes obvious. Because of its non-destructive nature, the NDT is similar to the medical tests performed on humans or animals. Although, some of the terms used in NDT are synonyms, from the technical point of view by non-destructive evaluation one understands, first of all, the measurement and the description of a defect, the establishment of the its shape and position, the determination, for a material, of the its characteristics, physical, magnetic properties [3–5].

NDT is useful to explore the material integrity of the object under test. Whereas the measurements performed in several areas of science and technology, such as astronomy, radio, electricity (e.g., voltage and current measurement) qualify as non-destructive tests, yet they do not evaluate properties of materials, is specified manners. NDT practically deals with the performance of the equipment under test and answers the question: For how long one can use that piece be safely, and when it needs to be re-tested [6, 7]?

Non-destructive testingin the electromagnetic fieldis one of the fastest and least expensive testing techniques for pieces subject to degradation. This domain has evolved a lot in recent years, because of the increasing demands received by the scientific community from the industry.

Non-destructive testing aims to detect defects (different types of cracks, structural inhomogeneity) in materials (conducting, ferromagnetic) without destroying the tested object. Therefore, the application of such techniques addresses many relevant domains that require high security of installations, domains such as aeronautical, nuclear, medical, or chemical industry.

The researches in non-destructive testing in the electromagnetic field consider important the following two principal directions:

- Flaw (defect) detection
- Flaw shape reconstruction.

Flaw shape reconstruction is an actual domain, drawing huge financial efforts. There are numerous and prestigious international conferences and publications dedicated to this field [1, 8–37].

This chapter starts by giving an introduction in non-destructive testing by highlighting its importance regarding the safety that parts of a structure perform their working functions for a pre-determined period. Non-destructive testing does not

affect the integrity of the body under test, becoming therefore very important for quality control of the product, which is or will be in use. The specialists in this domain have created and implemented methods and devices to characterize the materials or to detect, localize and measure the flaws (defects) with the purpose of preventing events with a high-impact such as plane crushes, nuclear power plant explosions, derailment of trains and other events which are less visible but also dangerous.

The researches in the field of non-destructive testing in the electromagnetic field are oriented into two directions: flaw detection and flaw reconstruction. If the flaw is a crack, then the use of eddy currents testing is the most suitable procedure. For the ferromagnetic pieces, eddy currents testing has significant disadvantage: due to the big permeability, the penetration depth is very small, so only surface defects will be detected. For ferromagnetic body flaws, the best procedure is the use of difference static magnetic field. We have a magnetic field problem in steady-state, but we must deal with the non-linearity of the medium. Because flaw reconstruction implies a huge computation time for the direct problem, one of the main objectives is to obtain procedures for rapid computation of the electromagnetic field in linear/nonlinear media. Many papers in literature recommend the finite element method (FEM) or the hybrid method (Finite Element Method—Boundary Element Method (FEM–BEM)).

The modification of the B-H characteristic, in aged areas, suggests the possibility of elaborating some electromagnetic detection and reconstruction procedures. The computation of the difference static magnetic field seems to be the best procedure. Because the change of the B-H characteristic is very small, it is necessary to obtain a high accuracy computation of the magnetic field in the measurement points.

This chapter falls within the field of non-destructive assessment methods by determining the B-H characteristic of ferromagnetic bodies, by using, for example, a device and a procedure for measuring the first magnetization characteristic.

The main objective of this chapter is the electromagnetic field analysis, together with techniques for determining the corresponding relations. There will be used numerical methods and methods for nonlinearity treatment of the magnetization characteristics, such that the time and the resources allocated for computation to be minimized.

To accomplish the above-proposed objectives, in this chapter, there will be presented the main non-destructive measurement devices, in correlation with related applications. Special attention will be given to those devices used to characterize the magnetic materials widely used in energy generation, distribution and conversion, in telecommunication, aeronautics, etc.

It follows a presentation of the magnetic field analysis methods in nonlinear media. As an example, the study of the magnetization characteristic evaluation for ferromagnetic bodies will be given.

The chapter ends with conclusions and a large number of references in the non-destructive testing and electromagnetic field computation.

2 Non-destructive Measurement Devices

2.1 Description of the Main Devices Utilized for Characterization of the Magnetic Materials

Presently, the utilization of the magnetic materials covers essential domains such as energy distribution and conversion, information transmission and storage, in aeronautics. This fact gives an important reason to demand the measurement of magnetic properties with very high accuracy. Regarding the development of magnetic materials for research and industry use, one of the significant worldwide concern is maintaining the high standard of measurements [38].

At present, many laboratories are using a wide range of characterization devices to determine the magnetic materials properties. The most common devices are Epstein frame, Single Sheet Tester (SST), magneto-optic devices, and Vibrating Sample Magnetometer (VSM).

In [38, 39], the authors analyzed the most extensively used pieces of equipment and methods to determine the characteristics of magnetic materials. There are industrial and laboratory devices presented, accompanied by measurements performed to test their functioning and offer a solution for the complex characterization of soft magnetic material. There are essential aspects regarding the primary devices, and measuring procedures presented in [38–47].

2.2 Epstein Frame

Epstein frame is utilized to determine the electrical steel sheets and strips magnetic properties. The Epstein frame (Fig. 1) consists of four strips placed such that to form a square, with a standard width of 30 mm and length between 280 and 305 mm.

The measurement of electrical steel sheets magnetic properties using the Epstein frame demands the determination of the magnetic flux density (induction) B and the magnetic field strength (intensity) H.

The secondary coil of the frame measures the magnetic induction B.

Whereas the variation is sinusoidal, the induced electromotive force in the secondary coil is proportional to the magnetic induction.

Also, the magnetic computation of H requires measuring the value of the magnetization current (primary), followed by the application of Ampere's Theorem along the mean fiber of the Epstein frame's magnetic circuit [39, 40].

Because of its high reproducibility degree, the method of measuring the magnetic materials properties using the Epstein frame represents a widely accepted standard in the industry [11, 13].

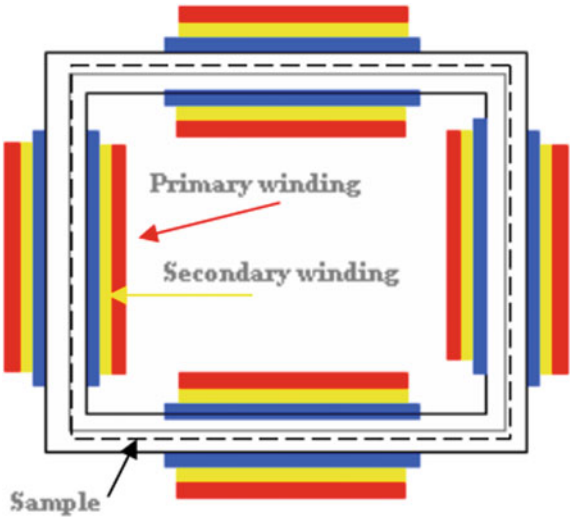


Fig. 1 Epstein frame

2.3 Single Sheet Tester (SST)

Since the Epstein method does not always give the best results, the Single Sheet Tester (SST) which allows the determination of the magnetic field. Several types of SST are used:

– horizontal tester with one frame;

Since the Epstein method does not always give the best results, the Single Sheet Tester (SST) which allows the determination of the magnetic field. There are several SST in use:

– horizontal tester with one frame;
– horizontal tester with two symmetrical frames;
– vertical tester with one frame;
– vertical tester with two symmetrical frames.

A characteristic of the SST method is the fact that there is only one magnetic sheet sample introduced in the space between the two windings: outer primary utilized for magnetization and secondary, which is inner and used to measure the magnetic induction.

For double frames horizontal SST, the frame assembly required the overlapping of metal sheets made from FeSi, with oriented crystals (GO), having a height of 0.1 m [41, 42].

The frame symmetry provokes a systematic error in the measurements because of the eddy currents, which appear in the lamination plane. The eddy currents are the results of the magnetic flux generated by the primary winding, which leaves the measured sample and enters the frame.

For single vertical frame SST, there is a constant error [11, 17–19, 38, 39], called end error when the sample is longer than the frame.

The end error can be eliminated by using symmetrical frames horizontal or vertical and by placing the sample between the two identical frames. If the target is the correct measurement of the tangential component of the magnetic field, which varies with distance, x at the sheet's surface, then two field coils are used. The measurement of the laminated sheets magnetic properties and the determination of the effective value of H, eliminating the approximation made regarding the length of the magnetic circuit, is done using the field coil method.

2.4 Magneto-optic Investigation Methods of the Magnetic Domain Structures

Due to this fact, the domain whose magnetizations are oriented differently, determine the different rotation of the reflected or transmitted light polarization plane [39, 43–45].

The measuring method based on the Kerr effect consists of directing a rectilinear polarized light fascicle using a Nicol prism. The reflected fascicle is going through an analyzer, which is, in fact, another Nicol prism. The transmission plane is practically in a right angle with the polarizer, ensuring, therefore, the extinction for a given magnetic state. The apparition of alternating bands, dark or bright when the prism rotates until it reaches an adequate position, underlines the existence of the domains. The contrast is optimum when the rectilinear incident vibration is almost parallel (t) or normal (n) to the plane of incidence, and the experience fully confirmed the theory [46].

The Kerr method requires special treatment of the sample surface but it applies to a vast range of temperatures and excitation field values.

2.5 Vibrating Sample Magnetometer (VSM)

Vibrating Sample Magnetometer (VSM) consists of a magnetized sample. VSM performs an oscillation movement inside a system of coils, such that inside, it induces an electromotive force (EMF) proportional to the sample's magnetic moment. If there is a magnetic dipole Q_{1z} which vibrates along the z-axis with amplitude δz, and the angular frequency of the vibration is ω, the dipole's position function of time is $z = \delta z(\cos\omega t)$. Such a scenario may be equivalent to the situation of a stationary quadrupole.

In his case, a system of coils system, sensitive to the quadrupole's vibrations can produce, a field with homogenous gradient $\partial H_z/\partial z$ with the assumption that an electric current is flowing through its windings. Inside the domain of homogeneity

for the gradient $\partial H_z/\partial z$, the sensitivity of coils with respect to the dipole's movement does not depend on the position of the volume of the dipole [41, 47].

The pair of coils, placed inside of the air-cored solenoid, may generate a magnetic field. The dipolar moment is equal to zero, being independent on the variations of the applied homogenous field. When the second-order derivative of the outer magnetic field has a significant value, then the detection coils can measure an induced EMF, leading to measurement errors. This fact recommends the introduction of an extra pair of coils able to cancel the induced EMF.

The connection of the outer coils is "series-opposition" having a big quadripolar moment to cancel the quadripolar moment from the central coils. A voltage divider was able to balance the system.

3 Numerical Assessment of Electromagnetic Energy and Forces

3.1 Electromagnetic Field Energy Theorem

The presentation of this theorem starts with two remarks [48]:

1. The first one refers to the importance of this theorem, importance that consists in the fact that it underlines *the electromagnetic fieldcapacity to accumulate and transfer energy.*
2. The second one refers to the hypothesis used to prove this theorem, more precisely: the bodies are at rest, and the media are linear, isotropic, and without permanent electric polarization and permanent magnetization. At the end of the proof, we have made observations regarding the influence of the adoption of these hypotheses upon the generality degree of the theorem.

Let us consider a closed surface (Σ) that bounds a domain (V_Σ) in which coexist the energetic interaction between the bodies (i.e., forms of matter, material substance) and the electromagnetic field. For the assumptions made above, the local forms of electromagnetic induction law, respectively magnetic circuit law for domains with continuity, show that in any point from domain (V_Σ) and for any instance of time, t there are the following relations in place:

$$\begin{aligned} curl\,\overline{E} &= -\frac{\partial \overline{B}}{\partial t} \\ curl\,\overline{H} &= \overline{J} + \frac{\partial \overline{D}}{\partial t}\,. \end{aligned} \tag{1}$$

Combining the two relations (1) results:

$$-\left(\overline{E}\cdot\frac{\partial\overline{D}}{\partial t}+\overline{H}\cdot\frac{\partial\overline{B}}{\partial t}\right)=\overline{E}\cdot\overline{J}+div\left(\overline{E}\times\overline{H}\right). \tag{2}$$

For linear media, isotropic and without permanent electric polarization or permanent magnetization, of absolute permittivity ε and absolute permeability μ, the constitutive laws have the forms $\overline{D}=\varepsilon\cdot\overline{E}$ and, respectively, $\overline{B}=\mu\cdot\overline{H}$, whereas the processing of the terms $\overline{E}\cdot\frac{\partial\overline{D}}{\partial t}$ and $\overline{H}\cdot\frac{\partial\overline{B}}{\partial t}$ gives:

$$-\frac{\partial}{\partial t}\left(\frac{\overline{D}\cdot\overline{E}}{2}+\frac{\overline{B}\cdot\overline{H}}{2}\right)=\overline{E}\cdot\overline{J}+div\left(\overline{E}\times\overline{H}\right). \tag{3}$$

Furthermore, to obtain global information concerning the electromagnetic phenomena from the domain (V_Σ), one can proceed as follows:

Computation of the space integral results into:

$$-\frac{\partial}{\partial t}\int\limits_{(V_\Sigma)}\left(\frac{\overline{D}\cdot\overline{E}}{2}+\frac{\overline{B}\cdot\overline{H}}{2}\right)\cdot dv=P+\int\limits_{(\Sigma)}\left(\overline{E}\times\overline{H}\right)\cdot\overline{dA}_{ext}. \tag{4}$$

Analyzing the above relation from a dimensional perspective and remarking that only one integral term contains the quantities defining the electromagnetic field, one can conclude that the surface integral

$$P_\Sigma=\int\limits_{(\Sigma)}\left(\overline{E}\times\overline{H}\right)\cdot\overline{dA}_{ext} \tag{5}$$

has the significance of a power, i.e., *the power transferred by that the electromagnetic fieldthrough the boundary* (Σ), while the volume integral

$$W=\int\limits_{(V_\Sigma)}\left(\frac{\overline{D}\cdot\overline{E}}{2}+\frac{\overline{B}\cdot\overline{H}}{2}\right)\cdot dv \tag{6}$$

has the significance of an energy, i.e., *the electromagnetic fieldenergy from the domain* (V_Σ).

The identity

$$-\frac{\partial W}{\partial t}=P+P_\Sigma \tag{7}$$

defines an equation referring to the power balance. It represents the mathematical expression of the electromagnetic energy theorem, which states: *The time variation*

of the electromagnetic field energy W from domain (V_Σ) *is equal to the sum between the power P transferred from the field to the bodies within the domain* (V_Σ) *during the electric conduction process, and the power* P_Σ *transferred by the field to the domain's exterior through the boundary* (Σ).

In the power balance equation, the quantity W is positively defined, whereas the quantities P and P_Σ can be either positive or negative depending on the problem under scrutiny. For the computation of term P_Σ, the vectorized area element $\overline{dA}$ points towards the exterior of the closed surface (Σ).

Meanwhile, the signs in front of the scalars P, P_Σ and $\frac{\partial W}{\partial t}$ show the real directions of the energy transfer, as follows:

(1) $P > 0$: the electromagnetic field transfers power from the domain (V_Σ) to the bodies from the domain (V_Σ) during the electric conduction process
(2) $P < 0$: the bodies from domain (V_Σ) transfer power to the electromagnetic field during the electric conduction process
(3) $P_\Sigma > 0$: the electromagnetic field transfers power through the boundary (Σ)
(4) $P_\Sigma < 0$: the electromagnetic field receives power through the boundary (Σ)
(5) $\left(-\frac{\partial W}{\partial t}\right) > 0$: the energy of the electromagnetic field from the domain (V_Σ) decreases in time
(6) $\left(-\frac{\partial W}{\partial t}\right) < 0$: then the energy of the electromagnetic field from the domain (V_Σ) increases in time.

Figure 2 illustrates the case: $P > 0$, $P_\Sigma > 0$.

According to the energy localization principle, the quantity:

$$w = \frac{\overline{D} \cdot \overline{E}}{2} + \frac{\overline{B} \cdot \overline{H}}{2} \tag{8}$$

represents the volumetric density of the electromagnetic field energy from the domain (V_Σ) and it has two components:

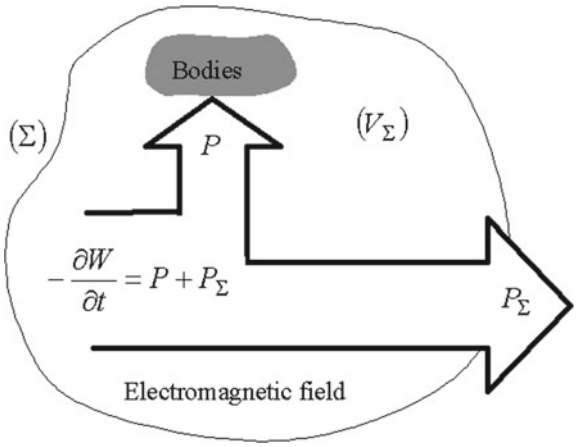


Fig. 2 Qualitative illustration of the electromagnetic field

(a) The volumetric density of electric energy:

$$w_e = \frac{\overline{D} \cdot \overline{E}}{2} \tag{9}$$

(b) The volumetric density of the magnetic energy

$$w_m = \frac{\overline{B} \cdot \overline{H}}{2}\,. \tag{10}$$

Consequently, the next identity results from a volume integration covering the domain (V_Σ):

$$W = \int\limits_{(V_\Sigma)} \left(\frac{\overline{D} \cdot \overline{E}}{2} + \frac{\overline{B} \cdot \overline{H}}{2} \right) \cdot dv = \int\limits_{(V_\Sigma)} w_e \cdot dv + \int\limits_{(V_\Sigma)} w_m \cdot dv = W_e + W_m\,, \tag{11}$$

Relation (11) demonstrates that the energy W of the electromagnetic field is the sum between the energy W_e of the electric field (12) and the and energy W_m of the magnetic field (13).

$$W_e = \int\limits_{(V_\Sigma)} w_e \cdot dv \tag{12}$$

$$W_m = \int\limits_{(V_\Sigma)} w_m \cdot dv. \tag{13}$$

The quantity

$$\overline{S} = \left(\overline{E} \times \overline{H}\right) \tag{14}$$

is called Poynting vector and has as the significance of the electromagnetic power flux density through the boundary (Σ), respectively the "flow" speed of electromagnetic energy through this surface:

$$P_\Sigma = \int\limits_{(\Sigma)} \overline{S} \cdot \overline{dA}_{ext} = \int\limits_{(\Sigma)} \left(\overline{S} \cdot \overline{n}\right) \cdot dA\,. \tag{15}$$

The vector $\overline{S}$ has the form

$$\overline{S}^{'} = \left(\overline{E} \times \overline{H}\right) + \overline{S}^{''} \tag{16}$$

with $\overline{S}''$ a zero-divergence vector.

3.2 *Electromagnetic Field Energy and Coenergy for Nonlinear Media*

Let us consider again a closed surface (Σ) that bounds a domain (V_Σ) in which coexist bodies and electromagnetic field in energetic interaction. From the set of adopted hypotheses required to prove the electromagnetic energy theorem, one must eliminate the assumption referring to the linear character of media.

However, the hysteresis of media does not play any role this time (i.e., media without hysteresis).

As shown previously, the immobility of the bodies and the inexistence of the hysteresis indicates that $P_m = P_h = 0$ and subsequently, the first principle of thermodynamics has the form:

$$-\frac{\partial W}{\partial t} = P + P_\Sigma \,. \tag{17}$$

The relationship (18)

$$-\left(\overline{E}\cdot\frac{\partial \overline{D}}{\partial t} + \overline{H}\cdot\frac{\partial \overline{B}}{\partial t}\right) = \overline{E}\cdot\overline{J} + div\left(\overline{E}\times\overline{H}\right) \tag{18}$$

was proved in the above paragraph using the local forms of the electromagnetic induction law and magnetic circuit law and, consequently, stays valid for any type of media, including the nonlinear ones.

From (19)

$$-dW = (P + P_\Sigma)\cdot dt \tag{19}$$

and (20)

$$-\int\limits_{(V_\Sigma)} \left(\overline{E}\cdot d\overline{D} + \overline{H}\cdot d\overline{B}\right)\cdot dv = (P + P_\Sigma)\cdot dt \tag{20}$$

results the expression of variation dW of electromagnetic field energy during that transformation (21):

$$dW = \int\limits_{(V_\Sigma)} \left(\overline{E}\cdot d\overline{D} + \overline{H}\cdot d\overline{B}\right)\cdot dv \,. \tag{21}$$

Using again the principle of energy localization (22)

$$dW = \int\limits_{(V_\Sigma)} dw \cdot dv\,, \tag{22}$$

results the variation dw of the volumetric density of electromagnetic field energy as:

$$dw = \overline{E} \cdot d\overline{D} + \overline{H} \cdot d\overline{B} \tag{23}$$

The two components of dw are:

– variation dw_e of the volumetric density of electric energy

$$dw_e = \overline{E} \cdot d\overline{D} \tag{24}$$

– variation dw_m of the volume density of magnetic energy

$$dw_m = \overline{H} \cdot d\overline{B} \tag{25}$$

If the media does not have permanent electric polarization, nor permanent magnetization, then the volumetric densities of electric energyw_e, respectively of the magnetic energy w_m corresponding to an arbitrary state characterized by the pair values $\left(\overline{E}, \overline{D}\right)$ respectively, $\left(\overline{H}, \overline{B}\right)$ result from integrations (26) and (27) while the reference state assumes all of the electromagnetic field state quantities equal to zero:

$$w_e = \int\limits_0^{\overline{D}} \overline{E} \cdot \overline{dD} = \int\limits_0^{D} E \cdot dD \tag{26}$$

respectively:

$$w_m = \int\limits_0^{\overline{B}} \overline{H} \cdot \overline{dB} = \int\limits_0^{B} H \cdot dB \tag{27}$$

The volumetric density w of electromagnetic field energy is:

$$w = w_e + w_m\,. \tag{28}$$

The volumetric density of electric coenergy is:

$$w_e^* = \int_0^{\overline{D}} \overline{E} \cdot \overline{dD} = \int_0^D E \cdot dD \tag{29}$$

with the volume density of magnetic coenergy:

$$w_m^* = \int_0^{\overline{B}} \overline{H} \cdot \overline{dB} = \int_0^B H \cdot dB \tag{30}$$

and the volume density w^* of electromagnetic field coenergy is

$$w^* = w_e^* + w_m^* . \tag{31}$$

Figure 3 contains two nonlinear dependencies, one in terms of the electric field (see Fig. 3a) and one in terms of the magnetic field (see Fig. 3b) are qualitatively depicted. The geometric interpretation comes immediately:

- The volumetric density of electric energy w_e, corresponding to electric field state, and characterized by the pair of values (E, D), is represented by the horizontally hatched area bordered by the nonlinear curve $D = D(E)$, OD axis and the horizontal line $D =$ constant,
- The volumetric density of electric coenergyw_e^*, corresponding to electric field state characterized by the pair of values (E, D), is represented by the vertically hatched area included between the nonlinear curve that gives the dependence $D = D(E)$, OE axis and the vertical line $E =$ constant.
- The volumetric density w_m of the magnetic energy, corresponding to magnetic field state characterized by the pair of values (H, B), is represented by the horizontally hatched area bounded between the nonlinear curve that gives the dependence $B = B(H)$, OB axis and the horizontal line $B =$ constant,

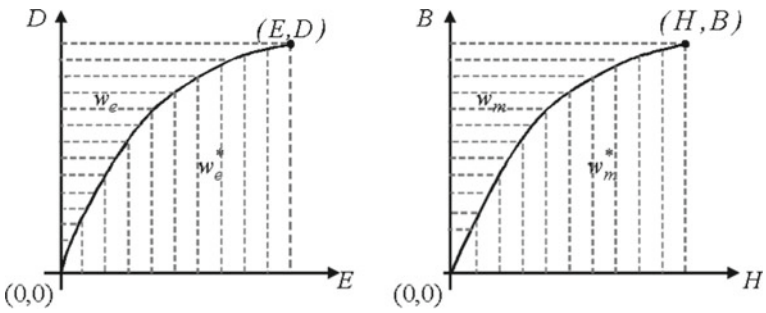


Fig. 3 Generic (qualitative) nonlinear dependencies $D = D(E)$ and $B = B(H)$

– The volumetric density w_m^* of magnetic coenergy, corresponding to magnetic field state characterized by the pair of values (H, B), is represented by the vertically hatched area bounded between the nonlinear curve that gives the dependence $B = B(H)$, OH axis and the vertical line $H = \text{constant}$.

These interpretations prove that:

$$w_e + w_e^* = D \cdot E = \overline{D} \cdot \overline{E} \tag{32}$$

$$w_m + w_m^* = B \cdot H = \overline{B} \cdot \overline{H}, \tag{33}$$

Relations (32) and (33) in the particular case of linear media become:

$$w_e = w_e^* = \frac{\overline{D} \cdot \overline{E}}{2} \tag{34}$$

$$w_m = w_m^* = \frac{\overline{B} \cdot \overline{H}}{2}\,. \tag{35}$$

In this, way we found again the expressions obtained for volume densities of electric energy and, respectively, magnetic energy, this time for linear media.

3.3 Warburg Theorem

The purpose of this theorem is to evaluate a specific experimental finding in terms of quantity. According to the specific experiments, one found that in nonlinear media with hysteresis, the electromagnetic field transfers power to the bodies during the hysteresis cycle. Let us consider again a closed surface (Σ) that bounds a domain (V_Σ) in which coexist bodies (i.e., material) and electromagnetic field in energetic interaction. The end of Sect. 1 shown that the electromagnetic field energy variation (V_Σ) obeys the first principle of thermodynamics, mathematically reflected by the relation:

$$-\frac{\partial W}{\partial t} = P + P_m + P_h + P_\Sigma\,, \tag{36}$$

For immobile bodies $(P_m = 0)$ has the particular form:

$$-\frac{\partial W}{\partial t} = P + P_h + P_\Sigma\,. \tag{37}$$

In this power balance equation, P_h represents the power transferred by the electromagnetic field to the bodies from the domain (V_Σ) through hysteresis. However, the expression of the electromagnetic field energy W is unknown.

The relation

$$-\int\limits_{(V_\Sigma)} \left(\overline{E} \cdot \frac{\partial \overline{D}}{\partial t} + \overline{H} \cdot \frac{\partial \overline{B}}{\partial t} \right) \cdot dv = P + P_\Sigma \tag{38}$$

proven in Sect. 1 using the local forms of electromagnetic induction law and magnetic circuit law. Consequently, it is also valid for media with hysteresis.

Carrying out an elementary transformation results:

$$-dW = (P + P_\Sigma) \cdot dt + P_h \cdot dt \tag{39}$$

and, respectively,

$$-\int\limits_{(V_\Sigma)} \left(\overline{E} \cdot d\overline{D} + \overline{H} \cdot d\overline{B} \right) \cdot dv = (P + P_\Sigma) \cdot dt \tag{40}$$

whereas subtracting member by member:

$$-dW + \int\limits_{(V_\Sigma)} \left(\overline{E} \cdot d\overline{D} + \overline{H} \cdot d\overline{B} \right) \cdot dv = P_h \cdot dt \tag{41}$$

From the principle of energy localization and, respectively, of the power, the variation of electromagnetic field energy volumetric density dw in the domain (V_Σ):

$$dW = \int\limits_{(V_\Sigma)} dw \cdot dv \tag{42}$$

The volumetric density p_h of power transferred by the field to the bodies from the domain (V_Σ) by hysteresis:

$$P_h = \int\limits_{(V_\Sigma)} p_h \cdot dv\,. \tag{43}$$

Then Eq. (41) becomes, assuming the immobile bodies:

$$
\begin{aligned}
&-\int\limits_{(V_\Sigma)} dw \cdot dv + \int\limits_{(V_\Sigma)} \left(\overline{E} \cdot d\overline{D} + \overline{H} \cdot d\overline{B}\right) \cdot dv = \\
&= \left(\int\limits_{(V_\Sigma)} p_h \cdot dv\right) \cdot dt = \int\limits_{V_\Sigma} (p_h \cdot dt) \cdot dv
\end{aligned}
\tag{44}
$$

The equality (41) is true whichever the domain (V_Σ) is, such that the integral equality implies the equality of the integrands:

$$
-dw + \left(\overline{E} \cdot d\overline{D} + \overline{H} \cdot d\overline{B}\right) = p_h \cdot dt \tag{45}
$$

To avoid the complication of finding the instantaneous transfer of power in the hysteresis process, we consider a very frequently case met in practice, namely the case of the electromagnetic fields with periodic time variations (called alternating fields). Let T be the period necessary to traverse an electric hysteresis cycle $\overline{D} = \overline{D}\left(\overline{E}\right)$ and, respectively, a magnetic one, $\overline{B} = \overline{B}\left(\overline{H}\right)$. Integrating with respect to the time variable, one can obtain:

$$
-\int\limits_0^T dw + \int\limits_0^T \left(\overline{E} \cdot \overline{dD} + \overline{H} \cdot \overline{dB}\right) = \int\limits_0^T p_h \cdot dt, \tag{46}
$$

in which the integral

$$
\int\limits_0^T p_h \cdot dt = w_{h,cycle} \tag{47}
$$

represents the volumetric density $w_{h,cycle}$ of the energy transferred by the electromagnetic field to bodies during a full hysteresis cycle, electric and magnetic. Because both, the electromagnetic energy W and the volumetric energy density w (whose expression is unknown) are state functions, it results that:

$$
\int\limits_0^T dw = w(T) - w(0) = 0\,. \tag{48}
$$

Moreover,

$$
\int\limits_0^T \left(\overline{E} \cdot \overline{dD} + \overline{H} \cdot \overline{dB}\right) = \int\limits_0^T (E \cdot dD + H \cdot dB) \tag{49}
$$

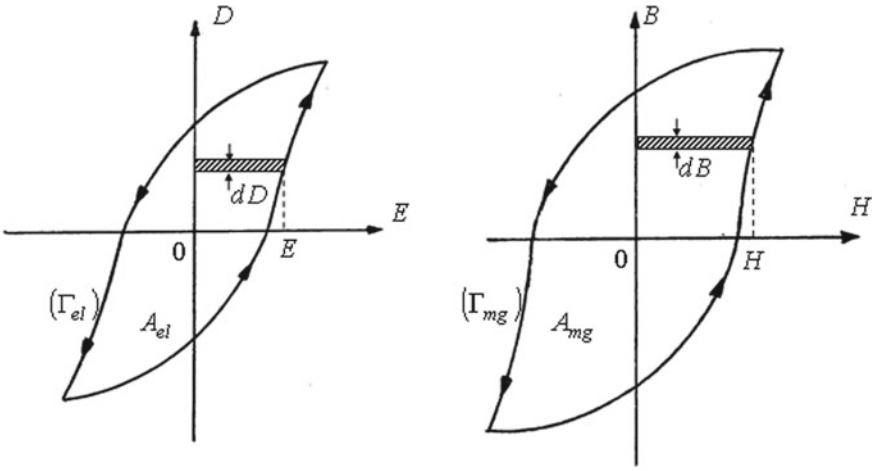


Fig. 4 Hysteresis cycles

the time integration can be interpreted as a space integration for the closed contours (Γ_{el}) and (Γ_{mg}) corresponding to the two hysteresis cycles electric and, respectively, magnetic one (see Fig. 4).

With these assumptions, the volumetric density of the electromagnetic energy for a hysteresis cycle is:

$$w_{h,cycle} = \int_0^T E \cdot dD + \int_0^T H \cdot dB = \int_{(\Gamma_{el})} E \cdot dD + \int_{(\Gamma_{mg})} H \cdot dB = w_{h,el} + w_{h,mg}\,. \tag{50}$$

Relation (50) emphasizes the Warburg theorem: *The two components of the volumetric energy density irreversibly transferred by the electromagnetic fieldto bodies during the two-hysteresiscycles are*:

$$w_{h,el} = \int_{(\Gamma_e)} E \cdot dD = \mathrm{A}_{el} \tag{51}$$

$$w_{h,mg} = \int_{(\Gamma_m)} H \cdot dB = \mathrm{A}_{mg} \tag{52}$$

The terms of relations (51), (52) and have the significations of the areas $\mathrm{A_{el}}$ and $\mathrm{A_{mg}}$ of the cycles. This energy transfer leads to the heating of heating ferroelectric and ferromagnetic bodies by transforming the electromagnetic energy in caloric energy.

3.4 The Generalized Forces Theorem in the Electromagnetic Field

The bodies placed inside the electromagnetic field are subjects to ponderomotive actions: forces, torques, superficial tensions, pressures. This fact inspired the scientists to introduce and utilize the concepts of *generalized coordinate* and *generalized force.* The body exposure to ponderomotive actions can results in the modification of its state due to translations, rotations, or deformations caused by the mechanical work involved. Let us consider again a closed surface (Σ) that bounds a domain (V_Σ) in which coexist bodies (i.e., material) and electromagnetic field in energetic interaction. Moreover, the term: system of bodies with *n degrees of freedom* designates a system completely described by a number of *n scalar independent variables.* The variables are defined as *generalized coordinates* and denoted by $x_1,\ x_2, \ldots, x_n$. If one of the bodies from the system modifies its position, for example, one of its coordinates, x_k may encounter one elementary variation dx_k. This fact signifies an elementary mechanical work dL_k.

The elementary variations dL_k and dx_k are linearly linked (53)

$$dL_k = X_k \cdot dx_k \tag{53}$$

by the proportionality coefficient X_k, called the *generalized force*. It becomes a force (in the current acceptance of the notion) whereas x_k is a length, a torque when x_k is an angle, a superficial tension when x_k is an area, a pressure when x_k is a volume. Furthermore (54) expressed the energy variation in terms of power:

$$-\frac{\partial W}{\partial t} = P + P_m + P_\Sigma \tag{54}$$

and

$$-\int\limits_{(V_\Sigma)} \left(\overline{E} \cdot \frac{\partial \overline{D}}{\partial t} + \overline{H} \cdot \frac{\partial \overline{B}}{\partial t} \right) \cdot dv = P + P_\Sigma \,. \tag{55}$$

Subtracting member by member the relation (54) and (55) one obtains:

$$-\frac{\partial W}{\partial t} + \int\limits_{(V_\Sigma)} \left(\overline{E} \cdot \frac{\partial \overline{D}}{\partial t} + \overline{H} \cdot \frac{\partial \overline{B}}{\partial t} \right) \cdot dv = P_m \,. \tag{56}$$

Theory proven in the past that the elementary mechanical work

$$dL_k = P_m \cdot dt \tag{57}$$

depends on the elementary performed transformation performed, while the generalized forces do not depend on it. Therefore, one can choose, the elementary transformation meant to simplify most of the computations, without diminishing the generality of the conclusions. This elementary transformations refer to the electric fluxes Ψ_{S_Γ} and magnetic fluxes Φ_{S_Γ}, which are time-invariant, through any surface S_Γ. The cancellation of the time derivatives of these fluxes leads to:

$$\frac{\partial \Psi_{S_\Gamma}}{\partial t} = \frac{\partial}{\partial t} \int\limits_{(S_\Gamma)} \overline{D} \cdot \overline{dA} = \int\limits_{(S_\Gamma)} \frac{\partial \overline{D}}{\partial t} \cdot \overline{dA} = 0 \tag{58}$$

and, respectively,

$$\frac{\partial \Phi_{S_\Gamma}}{\partial t} = \frac{\partial}{\partial t} \int\limits_{(S_\Gamma)} \overline{B} \cdot \overline{dA} = \int\limits_{(S_\Gamma)} \frac{\partial \overline{B}}{\partial t} \cdot \overline{dA} = 0\,. \tag{59}$$

Consequently, using the previous equations, results:

$$-\frac{\partial W}{\partial t} = P_m\,, \tag{60}$$

conditions in which, for an elementary transformation, the power balance equation:

$$-dW = P_m \cdot dt = dL_k = X_k \cdot dx_k \tag{61}$$

demonstrates that the elementary mechanical work dL_k is realized exclusively on the decrease of the electromagnetic field energy.

The relation:

$$X_k = -\left(\frac{\partial W}{\partial x_k}\right)\Bigg|_{\Psi,\Phi=ct} \tag{62}$$

represents the mathematical form of the generalized forces theorem in electromagnetic field. This theorem has the following statement: "*The generalized force X_k exerted along the generalized coordinate x_k is equal to the negative of the partial derivative of the electromagnetic field energy W, function of the generalized coordinate x_k, in conditions of time-invariant electric and magnetic fluxes through any surface.*"

As

$$W = W_e + W_m \tag{63}$$

it results that

$$X_k = -\left(\frac{\partial W}{\partial x_k}\right)\Bigg|_{\Psi,\Phi=ct} = -\left(\frac{\partial W_e}{\partial x_k}\right)\Bigg|_{\Psi=ct} - \left(\frac{\partial W_m}{\partial x_k}\right)\Bigg|_{\Phi=ct}, \tag{64}$$

followed by the particular relations defining the generalized forces theorem in the electric field:

$$X_k = -\left(\frac{\partial W_e}{\partial x_k}\right)\Bigg|_{\Psi=ct} \tag{65}$$

respectively the generalized forces theorem in the magnetic field:

$$X_k = -\left(\frac{\partial W_m}{\partial x_k}\right)\Bigg|_{\Phi=ct} \tag{66}$$

There are few observations valid:

- X_k is a scalar component (affected by sign) of the vector component $\overline{X}_k$ representing the generalized force $\overline{X}$ with respect to the direction of the generalized coordinate x_k
- The sign of the scalar X_k gives the information regarding the direction of the vector component $\overline{X}_k$ from generalized force $\overline{X}$: if $X_k > 0$, then the component $\overline{X}_k$ is oriented in the increasing direction of the generalized coordinate x_k, and if $X_k < 0$, then the component $\overline{X}_k$ is oriented in the decreasing direction of the generalized coordinate x_k
- To compute the generalized force X_k, there is necessary to determine the system degrees of freedom (i.e., their number), to express the electromagnetic field energy W as a function of fluxes and the generalized coordinate x_k and then perform a derivation of the energy expression with respect to x_k, assuming as the electric and the magnetic fluxes are time-invariant.

4 Case of Study

Nowadays, the existing devices and procedures, for determining the ferromagnetic materials' **B-H** relationship, use samples from the studied material. The geometry is imposed by the installation [38, 39]. The excitation current is a time variable. For instance, the Epstein frame measure **B-H** relationships for sheets of precise dimensions, which constitute the 4 branches of the magnetic circuit. The metal sheets interleaving at the frame's corners reduce the reluctance of the air gap between the branches.

Along the magnetic circuit, it is assumed that the magnetic field strength and the magnetic flux density are constant and as a consequence, the value of the current

which power the excitation coils is proportional **H**, and **B** is proportional to the voltage's time integral across the coil which measures the time-varying magnetic flux.

Single Sheet Tester (SST) is a device where the sample consists of a single sheet, which makes the yokes reluctance almost negligible. The magnetic field strength is achieved by measuring the excitation current. In this case, a Rogowski coil is utilized along with the sheet. In this way, the magnetic voltage drops from the yokes are canceled. The magnetic field must be a time-variable.

The devices presented above have some drawbacks, among which we can list:

- they need individual geometry samples (with dimensions which must be well-defined);
- the magnetic field is non-uniform, and the flux leakage is not null;
- the excitation current is time variable to produce a voltage at the terminals of the coil used for measurement;
- a higher voltage value at the terminals of the measuring coil involves a faster time change of the magnetic field, which can lead to the occurrence of eddy currents, thus inducing measurement errors.

One proposes an innovative measuring device [38] for assessing the **B-H** relationship of ferromagnetic materials, using a procedure accurate enough to obtain efficient results. The proposed device has two constructive shapes: one shape for samples, where materials of precise dimensions and geometric shapes are used, respectively, an in-situ shape, which can be used to assess a material of any dimensions or geometric shape.

The proposed device consists of a magnetic circuit (1, 2, 3, 4) like in Figs. 5 or 6, where one of the exterior branches contains the piece/sample (4) whose **B-H** relationship one wants to assess. The materials of the other branches of the magnetic circuit (1, 2, 3) have known **B-H** relationships, with an excellent static relative magnetic permeability. In the center, there is a median yoke (2, 3) interrupted by a small air gap. In the magnetic circuit windows, are placed the coils (5, 6), and excited by the currents i_1 and i_2. One energizes coil 2 with the current i_2, whereas its value is increasing in steps. For each value of i_2, one searches the value of current i_1 from coil 1 for which the magnetic flux through the median yoke (2, 3) is null (=0).

In the median yoke air gap, along the air gap, there are several Hall probes (7) placed in point equally distributed, such that the sum of the transversal components of the magnetic flux densities measured by these probes to allow us to decide whether the magnetic flux in the median yoke is zero (=0).

In Table 1 are described the components from which the measuring device is built.

Following the measurements presented above, one obtains a string of pairs ($i_1^{(k)}$, $i_2^{(k)}$), representing the currents from coils 1 and 2, for which the magnetic flux in the median yoke is zero. The procedure involved in the calculation of the **B-H** relationship for the piece (sample) is the following:

(1) For the first pair ($i_1^{(1)}, i_2^{(1)}$) one computes the magnetic field (**B, H**). The following equations are verified:

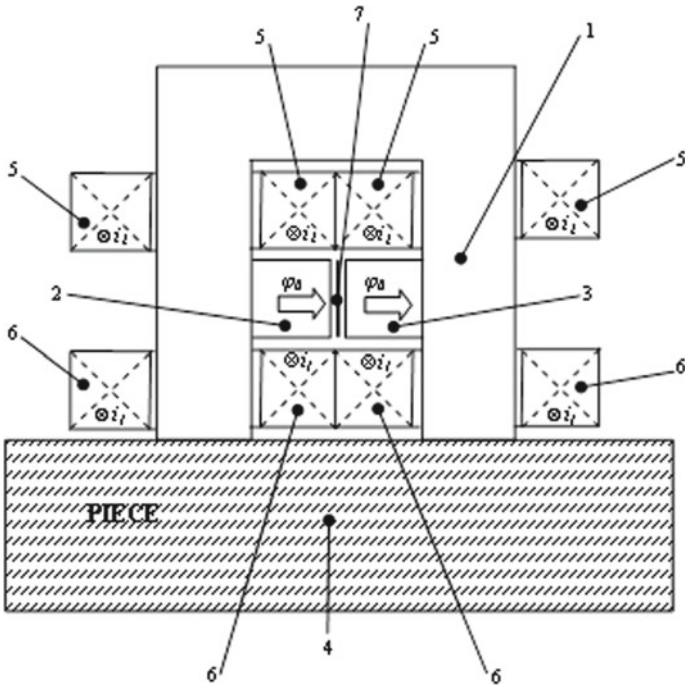


Fig. 5 In-situ measuring device

$$\nabla \times \mathbf{H} = \mathbf{J}, \tag{67}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{68}$$

$$H = \frac{B}{\mu_0}, \text{ in air}, \tag{69}$$

$$\mathbf{H} = F(\mathbf{B}), \text{ in the magnetic material of the device}, \tag{70}$$

where $\mathbf{J}$ is the current density imposed by the currents $i_1^{(1)}$, $i_2^{(1)}$ circulating through the two coils. For the studied sample, one assumes that the *B*-*H* relationship starts linearly from the origin, with a slope μ(i.e., considering the sample material isotropic and homogenous):

$$H = \frac{B}{\mu} \tag{71}$$

(2) Then one determines the flux ϕ_0 and searches the value $\mu^{(1)}$ for which $\phi_0 = 0$, then calculates the maximum value $B^{(1)}$ from the sample. For $B \in [0, B^{(1)}]$, the

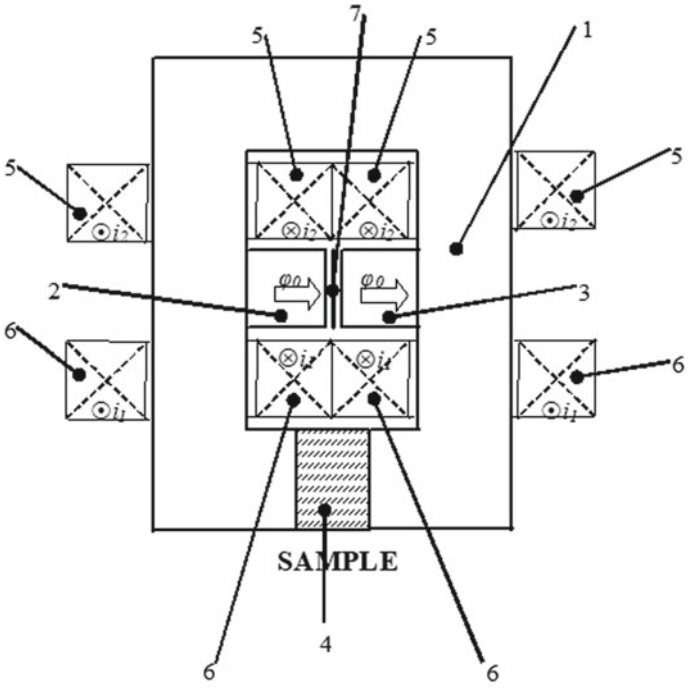


Fig. 6 Sample measuring device

Table 1 Components of the measuring device

Crt. no	Constitutive element
1	Device magnetic circuit—lateral and transversal yokes
2	Left median yoke
3	Right median yoke
4	Piece/sample
5	Coil_2
6	Coil_1
7	Hall probes

B-H relationship is described by the line of slope $\mu^{(1)}$:

$$H = \frac{B}{\mu^{(1)}}. \tag{72}$$

To determine the value of $\mu^{(1)}$, the secant method can be used.

(3) Next, for the pair $(i_1^{(2)}, i_2^{(2)})$ one determines the magnetic field which verifies the Eqs. (67)–(70). For the studied sample, one searches the linear extension of

the *B-H* relationship over the value $B^{(1)}$ and then searches for the slope $\mu^{(2)}$ for which $\phi_0 = 0$. This slope is valid for $B \in [B^{(1)}, B^{(2)}]$, where $B^{(2)}$ represents the newly sample's maximum value of the magnetic flux density. The extension of the B-H relationship becomes:

$$H = \begin{cases} \frac{B}{\mu^{(1)}}, & pt.\ B \in [0, B^{(1)}] \\ H - H^{(1)} = \dfrac{B - B^{(1)}}{\mu^{(2)}}, & pt.\ B \in [B^{(1)}, B^{(2)}] \\ \ldots & \\ H - H^{(n)} = \dfrac{B - B^{(n)}}{\mu^{(n+1)}}, & pt.\ B \in [B^{(n)}, B^{(n+1)}] \end{cases} \tag{73}$$

(4) The application of the method/algorithm continues in the same way, eventually obtaining the piecewise linear *B-H* relationship of the sample. The piecewise linear approximation is as smoother as we choose a smaller step Δi_2.

This device has the following advantages:

- One has to measure only two currents of the coils under the conditions of zero value of the magnetic flux through the median yoke ($\phi_0 = 0$).
- The measurement of the magnetic flux through the median yoke is not necessary; one just needs to ensure that it has a zero value (like Wheatstone bridge).
- When computing the magnetic field, the material and the form of the median yoke can be randomly adopted.
- The measuring procedure is not conditioned by the time variation of the currents through the coils.

The algorithm used to reconstruct the B-H relationship is based on a FEM program coupled with the *polarization of the fixed-point method* (PFPM) to solve a stationary magnetic field problem.

The program has two components: the magnetic field direct problem (used to determine the pairs of currents (i_1, i_2) for which the magnetic flux through the median yoke is null), and the field inverse problem (the reconstruction of the **B-H** relationship using the currents obtained in the direct problem).

4.1 The Magnetic Field Direct Problem—The Determination of Currents (i_1, i_2)

The magnetic circuit of the device is made of a zero-remanence ferromagnetic material. For the studied sample, one had taken into consideration two arbitrary materials with a known dependence: 1018 Steel, respectively 455 Steel. The three **B-H** dependencies are shown in Fig. 7.

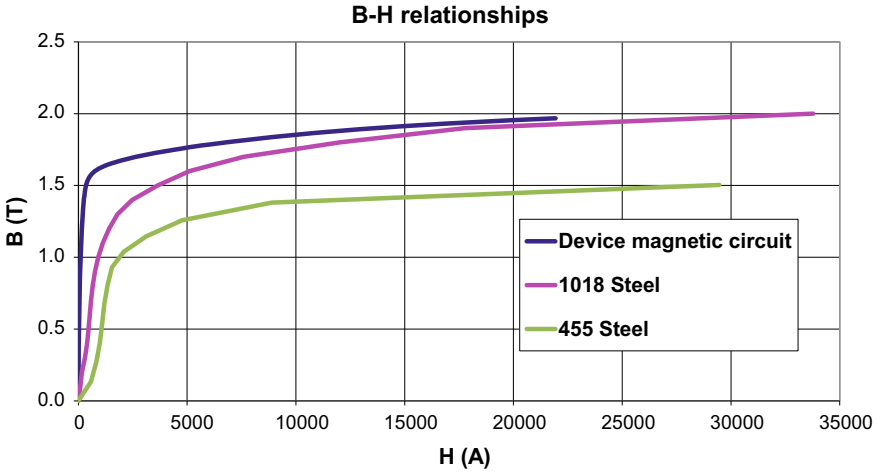


Fig. 7 B-H relationships of the device magnetic material and of the sample

Table 2 Currents (i_1, i_2) and the values of the maximum magnetic flux density for 1018 Steel sample, obtained from the direct problem

i_1 (A)	i_2 (A)	Bmax (T)
0.003	0.0989	0.1396
0.0057	0.1973	0.2751
0.009	0.296	0.4119
0.012	0.394	0.5496
0.0148	0.493	0.6872
0.018	0.593	0.82
0.0217	0.692	0.9461
0.0266	0.7915	1.0641
0.0336	0.8911	1.1754
0.0421	0.9914	1.2813
0.055	1.0916	1.3819
0.078	1.1925	1.4968
0.1144	1.294	1.6061
0.1764	1.3966	1.6645
0.2568	1.5032	1.7061
0.3712	1.6351	1.7494
0.5982	1.8948	1.8161
0.9892	2.3339	1.8934
1.752	3.0192	1.9922

Table 3 Currents (i_1, i_2) and the values of the maximum magnetic flux density of the 455 Steel sample, obtained from the direct problem

i_1 (A)	i_2 (A)	Bmax (T)
0.00860945	0.0989	0.1328
0.0128367	0.1973	0.2678
0.0163371	0.296	0.4025
0.0194525	0.394	0.5382
0.0226214	0.493	0.6736
0.0261479	0.593	0.8035
0.0302432	0.692	0.9311
0.0375819	0.7915	1.0384
0.0535341	0.8911	1.1457
0.0774609	0.9914	1.2579
0.124701	1.0916	1.38
0.284154	1.1925	1.5039
0.481376	1.294	1.62277

Tables 2 and 3 contain the pairs of currents (i_1, i_2) obtained by computing the magnetic field direct problem, respectively the samples' values of the maximum magnetic flux densities.

Figures 8 and 9 show the $i_2(i_1)$ dependencies.

Figures 10 and 11 show the magnetic flux density lines and spectre corresponding to the first pair of currents (i_1, i_2) from Tables 2 and 3. One can observe that the magnetic field is null in the median yoke.

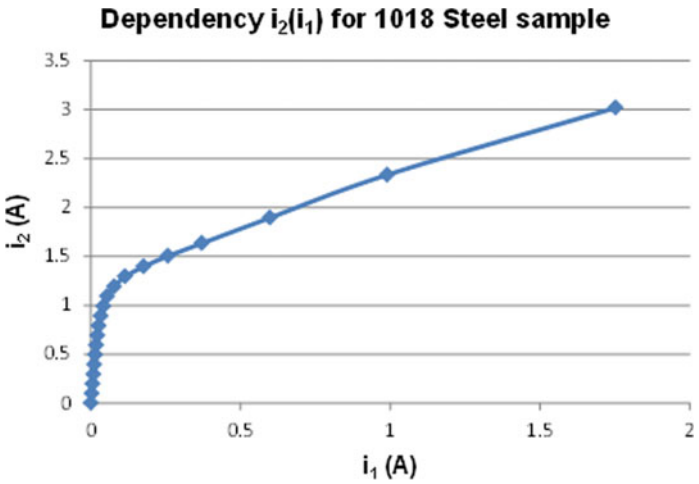


Fig. 8 Current dependency for 1018 Steel sample

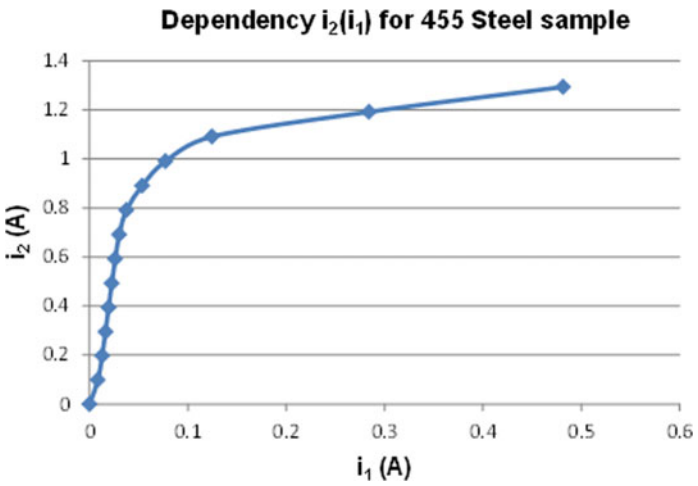


Fig. 9 Current dependency for 455 Steel sample

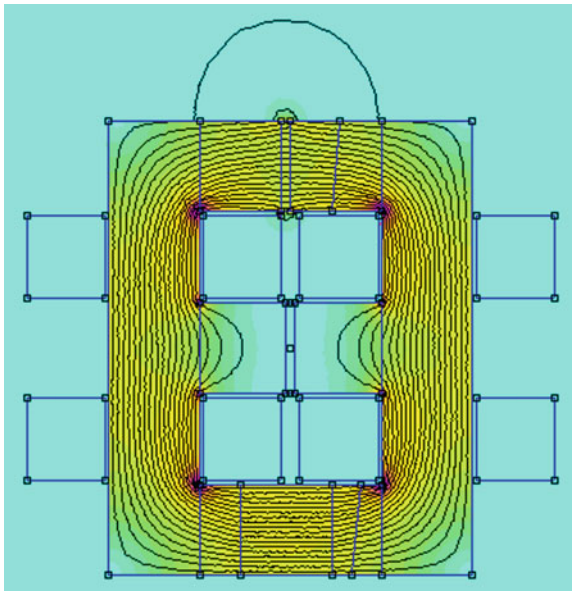


Fig. 10 Magnetic flux density lines for $(i_1, i_2) = (0.003, 0.0989)$ A—1018 Steel

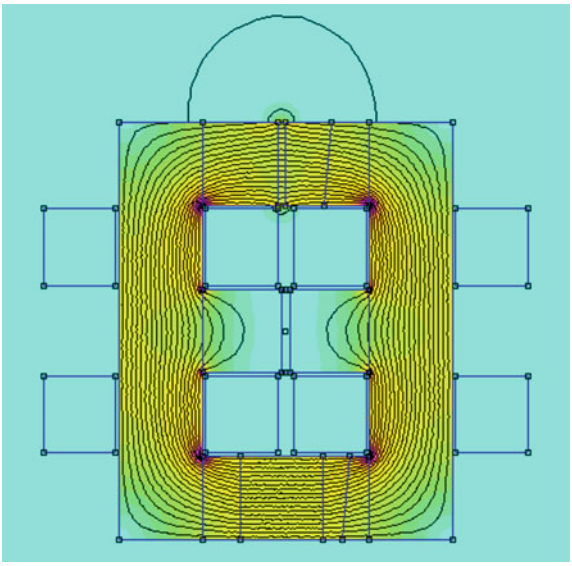


Fig. 11 Magnetic flux density lines for $(i_1, i_2) =$ (0.00860945, 0.0989) A—455 Steel

4.2 The Magnetic Field Inverse Problem—The Reconstruction of the B-H Relationship Using Currents (i_1, i_2)

Using the values of currents (i_1, i_2) from Tables 2 and 3, one reconstructs the B-H relationship of the two samples, by solving the magnetic field inverse problem. The results are shown in Figs. 12 and 13, noticing that the computed relationship is perfectly overlapped on the original one, given by the manufacturer. This proves the correctness of the proposed algorithm and device.

The device described above allows the determination of the B-H relationship by measuring the currents from two coils placed on the device columns. In the proposed procedure, one chooses a value for one of the currents. Then it searches the second one such that the value of the magnetic flux through the median yoke is zero. As a consequence, one must control the null value of the flux, a method similar to measuring resistances with the Wheatstone bridge. The values of the imposed current are ascending, also resulting in growing values for the second current. Thus, one can obtain the first magnetization curve, avoiding secondary hysteresis cycles. Then, one searches for the piecewise linear B-H relationship, by successively determining each segment of it. In the inverse problem, a segment obtained at an individual iteration describes the B-H relationship between the starting point (i.e., the maximum value of the magnetic flux density obtained in the sample for the previous pair of currents) and the ending point (i.e., the maximum magnetic flux density value obtained in the sample at the current iteration). Every magnetic flux computation presumes the

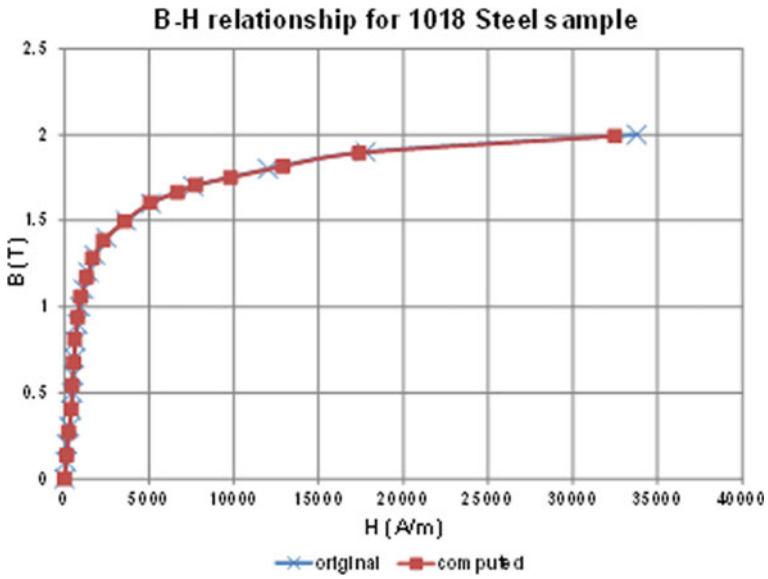


Fig. 12 Original and computed B-H relationship for 1018 Steel sample

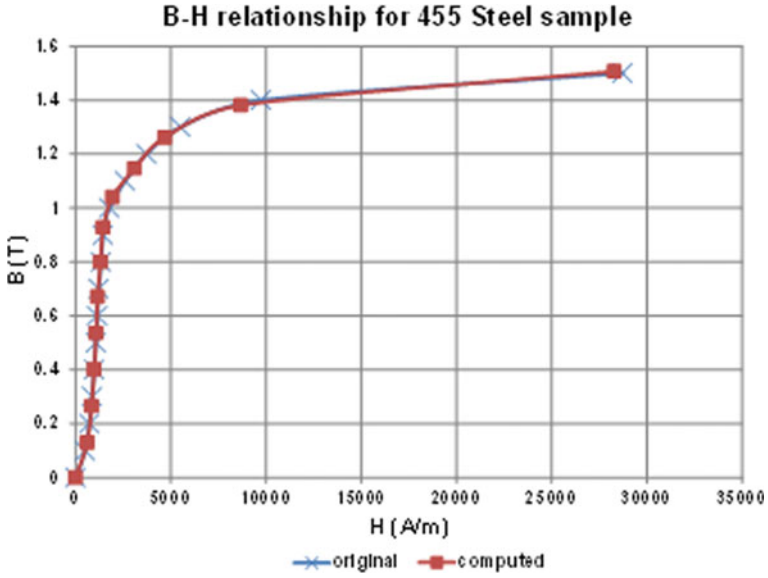


Fig. 13 Original and computed B-H relationship for 455 Steel sample

solving of a stationary magnetic field problem. The non-linearity of the B-H relationship is treated by using the iterative polarization fixed point method, and the magnetic field computation at each iteration had been made through a FEM program.

5 Conclusions

This chapter introduces the non-destructive testing field by highlighting its importance regarding safety. Non-destructive testing does not affect the integrity of the body under test, becoming therefore very important for the quality control process, subjecting the product in use or scheduled to be. The specialists in this domain have created and implemented methods and devices to characterize the materials or to detect, localize and measure the flaws (defects) with the purpose of preventing events with a high-impact such as plane crashes, nuclear power plant explosions, trains' derailment and other events which are dangerous but not very visible. The researches in NDT in the electromagnetic field look in two directions: flaw detection and flaw reconstruction. If the flaw is a crack, then the use of eddy currents testing is the most suitable procedure. For the ferromagnetic pieces, eddy currents testing has a significant disadvantage: due to the high permeability, the penetration depth is minimal, the fact which allows only surface defects detection. The best procedure of flaws detection in ferromagnetic bodies is to use the difference static magnetic field. We have a magnetic field problem in steady-state, but we must deal with the nonlinearity of the medium. Because flaw reconstruction implies a considerable computation time for the direct problem, one of the main objectives is to obtain procedures for the rapid computation of the electromagnetic field in linear/nonlinear media. Many papers in literature recommend the finite element method (FEM) or hybrid method (Finite—Boundary Element Method (FEM–BEM)).

The modification of the B-H characteristic, in aged areas, suggests the possibility of elaborating some electromagnetic detection and reconstruction procedures. The computation of the difference static magnetic field seems to be the best procedure. Because the change of the B-H characteristic is minimal, it is necessary to obtain a high accuracy computation of the magnetic field in the measurement points. This chapter falls within the field of non-destructive assessment methods by determining the B-H characteristic of ferromagnetic bodies, by using, for example, a device and a procedure for measuring the first magnetization characteristic. The main objective of this chapter is the electromagnetic field analysis, together with techniques for determining the corresponding relations. There several numerical methods available, as well as methods used to treat for the nonlinearity of the magnetization characteristics. However, only those methods able to deliver results in minimal time, with minimally allocated resources present interest. Therefore, there are presented the primary non-destructive measurement devices, in correlation with related applications. Special attention was granted to those devices used to characterize the magnetic materials widely used in energy generation, distribution and conversion, in telecommunication, aeronautics. The followed a presentation of the magnetic field analysis methods

in nonlinear media. As given an example, there is the study of the magnetization characteristic evaluation for ferromagnetic bodies. The device described in Sect. 4, allows the determination of the B-H relationship by measuring the currents from two coils placed on the device columns. The chapter ends with conclusions and many references in the non-destructive testing and electromagnetic field computation.

References

1. Stănculescu M (2009) Contribuţii la recunoaşterea formelor defectelor în defectoscopia magnetică nedistructivă (Contributions to flaw shape reconstruction in nondestructive testing). PhD thesis (in Romanian language), University "Politehnica" of Bucharest, Romania, 2008
2. SR EN ISO 9000-2001, Sisteme de management al calităţii. Principii fundamentale şi vocabular (Quality management systems. Vocabulary and fundamental principles)
3. www.asnt.org/ndt
4. IzfP Fraunhofer-Institut for NDT. https://mm.fhg.de/depts/izfp-e.html
5. Wenk SA, McMaster RC (1987) Choosing NDT: applications, costs and benefits of nondestructive testing in your quality assurance program. American Society for Nondestructive Testing, Columbus, OH
6. McMaster RC, Wenk SA (1951) A basic guide for management's choice of nondestructive tests. Special technical publication no. 112. American Society for Testing and Materials, Philadelphia, PA
7. Juran JM (1988) JURAN's quality control handbook, 4th edn. McGraw-Hill, NewYork
8. Diederichs R (2003) The advantages of the internet in the field of NDT, Internet article
9. Miller RK (1996) Nondestructive testing handbook, 2nd edn, vol 1, 2, 3, 4, 5, 6, 7, 8, 9 şi 10, American Society for Nondestructive Testing, USA
10. Cecco VS, Van Drunnen G, Sharp FL (1981) Eddy current manual: test method, vol 1, AECL-7523. Chalk River, Ontario
11. McMaster RC, McIntire P, Mester ML (eds) (1986) Non-destructive testing handbook: electromagnetic testing, vol 4, ASNT
12. de la Pintiere L (1985) Multifrequency eddy current examination of heat exchanger tubing. In: Lord W (ed) Electromagnetic methods for non-destructive testing. Gordon and Breach Science Publishers, New York, pp 195–303
13. Rao BPC, Baldev R, Jayakumar T, Kalyanasundaram P, Arnold W (2001) A new approach for restoration of Eddy current images. J Non-Destruct Eval 20:61–72
14. Ramuhalli P (2002) Electromagnetic NDE signal inversion by function-approximation neural network. IEEE Trans Magn 38(6):3633–3642
15. Takagi T, Miya K (2000) ECT Round-Robin test for steam generator tube. J Japanese Soc Appl Electromagn Mech 8:121–128
16. Chen Z, Miya K (1998) A new approach for optimal design of eddy current testing *probes*. J Nondestr Eval 17(3):105–116
17. Koch T, Browsing and searching internet resources. https://www.ub2.lu.se/nav_menu.html
18. Nondestructive evaluation *at SwRI*. https://www.nde.swri.edu/index.html
19. NTIAC. https://www.ntiac.com
20. Center for Nondestructive Evaluation at Iowa State University. https://www.cnde.iastate.edu/
21. European networks for structural integrity. https://science.jrc.nl/www/jrc/iam/sci-unit/networks/networks.html
22. DGZfP: German NDT Society. https://www.dgzfp.de
23. British Institute for NDT. https://www.powertech.co.uk/bindt/
24. The Canadian Society for Nondestructive Testing. https://www.csndt.org/
25. The NASA Technical Report Server. https://techreports.larc.nasa.gov/cgi-bin/NTRS

26. Panametrics. https://www.panametrics.com
27. Krautkramer. https://www.krautkramer.com
28. IRT—Inspection Research Technology. https://www.irt.co.il
29. Qnet. https://www.qnetworld.com
30. ASTM E 1316–1992 terminology for nondestructive examination
31. Mihai C (2003) Procedee electromagnetice nedistructive pentru evaluarea defectelor în materiale (in Romanian language), PhD thesis
32. Raj B, Jayakumar T, Rao BPC (1995) Review of NDT techniques for structural integrity. Sadhana, Acad Proc Eng Sci 20:5–38
33. Nichols RW (1982) Advances in NDE for structural integrity. Applied Science Publishers, London
34. Bllitz J (2000) Electrical and magnetic methods for non-destructive testing. Adam Hilger, Bristol
35. Lovejoy DJ (1993) Magnetic particle inspection: a practical guide. Chapman & Hall
36. Borucki JS (1991) Development of automated magnetic particle testing systems. Mater Eval 49:324–329
37. Lord W (1985) Electromagnetic methods of NDT. Gordon and Breach, New York
38. Andrei P (2015) Dispozitiv de determinare a caracteristicii de primă magnetizare pentru materiale feromagnetice (Device for determining the first magnetization curve for ferromagnetic materials). PhD thesis, PhD supervisor, Prof. PhD. Eng. I.F. Hantila, University Politehnica of Bucharest
39. Păltânea V (2008) Sisteme Avansate de Caracterizare a Materialelor Magnetice Moi (Advanced systemd for soft magnetic materials characterization). PhD thesis, PhD supervisor: Prof. PhD. Eng. H. Gavrilă, Octombrie
40. Kappel W, Caracterizarea avansată şi industrială a oţelurilor magnetice, contract 57/B Ad 1411/97/VIII, I.C.P.E
41. IEC Standard Publication 60404—2, Methods of measurement of the magnetic properties of electrical steel sheet and strip by means of the Epstein frame. IEC Central Office, Geneva, 1996
42. Sievert J, Ahlers H, Fiorillo F, Rocchino L, Hall M, Henderson L (2001) Magnetic measurements on electrical steels using Epstein and SST methods. Summary report of the EUROMET comparison project no. 489, PTB–Bericht, E−74, pp 1–28
43. Fiorillo F (2004) Measurement and characterization of magnetic materials. Elsevier Academic Press
44. Noţingher P (1995) Etude du dispositif a bande unique dans les conditions extremes de caracterisation magnetique, Stage de fin d'etudes, L.E.G.-I.N.P.G
45. Farrell E, Allen CJ, Whilden MW, Tripp JH, Usoskin A, Sheth S, Brittenham GM (2007) Magnetic measurement of liver iron stores: engineering aspects of a new scanning susceptometer based on high-temperature superconductivity. IEEE Trans Magnet 43(11):4030–4036
46. Tsukada K, Kiwa T (2007) Magnetic measurement of moisture content of grain. IEEE Trans Magnet Volo 43(6):2683–2685
47. Zijlstra H (1967) Experimental methods in magnetism, vol IX, Measurement of magnetic quantities. North–Holland publishing company, Amsterdam
48. Ochiana G, Ochiana L, Stănculescu M (2011) Teoria câmpului electromagnetic (Electromagnetic field theory), Editura Printech, 2011, ISBN 978–606–521–756–0, (214 pagini) (cod CNCSIS 54)

Optimal Integration of Electric Vehicles in Smart Grid Energy Flow

Sorin Deleanu, Marilena Stanculescu, Dragos Niculae, Paul Cristian Andrei, Lavinia Bobaru, and Horia Andrei

Abstract This chapter provides an insight into the rapidly growing domain represented by the plug-in electric vehicles (PEV) in connection with the smart grid power system (SGPS) and bi-directional power flow. The chapter starts with an introductive Sect. 1, followed by Sect. 2, containing a review of the main aspects, which define and characterize the interaction between the EVs and the SGPS. The analysis regarding the integration of the EVs into the SGPS under the vehicle to grid (V2G) considered the three main directions: SGPS efficiency improvement, cost-effectiveness, and the reduction in greenhouse gases. After a presentation of the SGPS structure, Sect. 2 includes a brief description of the primary power produced by the wind turbine as well as the one produced by the photovoltaic panels. Both represent input powers to the SGPS, fact which classifies these two elements of SGPS as power sources, although with an intermittent character. Following Sect. 3 dedicated to the charging stations, the chapter continues with Sect. 4 allocated to modeling, simulation, and results. Firstly, this section contains detailed models dedicated to the analysis of battery charging and battery discharging, with applications to individual vehicles and considering the presence of renewable energy sources. The models expanded

S. Deleanu (✉)
SAST, Northern Alberta Institute of Technology, Edmonton, Canada
e-mail: sorind@nait.ca

M. Stanculescu · D. Niculae · P. C. Andrei · L. Bobaru
Department of Electrical Engineering, University Politehnica Bucharest, Bucharest, Romania
e-mail: marilena.stanculescu@upb.ro

D. Niculae
e-mail: dragos.niculae@upb.ro

P. C. Andrei
e-mail: paul.andrei@upb.ro

L. Bobaru
e-mail: lavinia.bobaru@upb.ro

H. Andrei
Doctoral School of Engineering Sciences, University Valahia, Targoviste, Romania
e-mail: hr_andrei@yahoo.com

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_22

to multiple vehicles, scenarios, and simulations, including the discussions accompanying the results. In the last Sect. 5, the authors present the conclusions. The chapter concludes with an up to date section of references.

Keywords Electric vehicles · Smart grid power system · Charging stations · Energy flow

Nomenclature

A. Acronyms

PEV	Plug in Electric Vehicle
EV	Electric Vehicle
SGPS	Smart Grid Power System
RES	Renewable Energy Sources
V2G	Vehicle to Grid
PSS	Power System Stabilizer
SOC	State of Charge
SC	Solar Cells
ASC	Area of the Solar Cells
PV	Photovoltaic Panels
EENC	Expected Energy Not Charged
CS	Charging Subsystem
ESS	Energy Storage System
EPDS	Electric Power Distribution Subsystem
AES	Auxiliary Energy Subsystem

B. Symbols/Parameters

P	Power
P_{rated}	Rated power
P_{wind}	Power delivered by the wind turbine
ε	Albert Betz constant
ρ	Air density
A	Area cleared by the wind turbine
V	Speed
$v_{rated,w}$	Rated speed of the wind turbine
$v_{min,w}$	Minimum wind speed to allow the wind energy production
$v_{max,w}$	Maximum wind speed allowed for wind turbine safe operation
$P_{SC}(t)$	Output power of the solar cells
A_{SC}	Area of the solar cells
η_{SC}	Efficiency of the solar cells
$P_{PV}(t)$	Output power of the photovoltaic panels
A_{PV}	Area of the photovoltaic panels

η_{PV}	Efficiency of the photovoltaic panels
$\mu(t)$	Solar insolation
$V_{\text{int},rec}$	Rectified (DC) voltage
V_R	Voltage across the equivalent resistor of the charger
V_{cap}	Voltage across the equivalent capacitor of the battery
V_{out}	Voltage at the output of the charger operating as rectifier
i_1	Current at the input to the charger
i_2	Current at the output of the charger
η_C	Efficiency of the charging station operating as rectifier (charging)
R	Resistor
C	Capacitor
V_{in}, V_{out}	Input, respectively output voltage
i_{in}, i_{out}	Input, respectively output current
P_{in}, P_{out}	Input, respectively output power
T_C	Charging time constant
T_{max}	Duration necessary for battery to fully charge from zero power
P_{max}	Maximum power at the battery
t_1	Initial moment of the charging process
$P(t_1)$	Battery power level at the beginning of the charging
$P(t'_1)$	Battery power level at the beginning of the discharging
P_{BPEV}	Instantaneous power level of plug in electric vehicle battery
$P_{BPEV,max}$	Maximum power level of plug in electric vehicle battery
$P_{BPEV,dem}$	Instantaneous power demand of plug in electric vehicle battery
α,α_i	Charging constant of the plug in electric vehicle battery
$P_{BPEV,max}$	Maximum power level of plug in electric vehicle battery
$P_{BPEV,dem}$	Instantaneous power demand of plug in electric vehicle battery
E_{BPEV}	Energy necessary for the plug in electric battery to fully charge
Δt_i, $i = 1\ldots4$	Time interval
ΔT_{Fault}	Duration of the fault interval faced by the PEV battery during charging
γ_{dis}, γ_{disk}	Discharging constant of the plug in electric vehicle battery
n_{PEV}	Number of electrical vehicles in charging/discharging process
N_{st}	Number of charging stations
$P_{BPEV,inj}$	Power injected into the supply by the batteries of the PEV fleet
η_i	Efficiency of the charging station converter operating as inverter (V2G)
$P_{BPEV,dis}$	Power level of the plug in electric vehicle battery following discharging
P_{Lim}	Lower level specified for the plug in electric vehicle battery power
$P_{BPEV,net}$	Net power exchanged between the supply by the batteries of the PEV fleet
$P_{Renewables}$	Power injected into the SGPS by RES at the point of common coupling with the charging station
G_{Power}	Power gain at the connection point between the parking lot and power supply

G_{En} Energy gain at the connection point between the parking lot and power supply

1 Introduction

All the EV utilizing rotational motors for producing the tractive effort, display similar configurations of their power systems responsible for torque production and its conversion into the traction force applied at the wheel. Such systems consist of a primary energy source, converter(s), and rotational electric motors, eventually a gearbox and traction wheels. In the latest applications, one can observe a trend to integrate the assembly motor–traction wheel in a subcomponent eliminating the need for the differential gearbox. The autonomous electric vehicles, presented in the first section of this chapter, operate based on a primary source of energy, including a system of rechargeable batteries. Their connection to the power grid generally occurs during their parking time. Purely electric vehicles rely on batteries exclusively, whereas plug-in hybrid vehicles (PHEV) contain a fossil fuel-based primary energy source as well.

The implementation in the last decade of newly restrictive norms regarding the pollution reduction and climate change mitigation, targeting agglomerated metropoles from several developed countries, determined a significant increase in the percentage of EV from the total number of vehicles on the road.

Consequently, the autonomous EV fleet became a direct and essential player regarding the processes of energy conversion, transfer and distribution, working within SGPS which includes renewable energy sources (RES) as well. Ignoring the interaction between the electric vehicles and the SGPS became practically impossible, given the bi-directional character of the energy transfer. The electric energy flows from the grid to the vehicles during the battery recharging process, and conversely from the vehicles to the grid, mostly during the hours of highly local power demand. In the last case the electric vehicles, having their batteries charged enough, act as generators whereas the grid absorbs the electrical energy.

Section 2 of this chapter provides details regarding the bi-directional transfer of electric energy. The inherent involvement of the power electronic converters in the process of the electric energy transmission, no matter their location, inside the grid or belonging to the electric vehicles, may result in the generation of current harmonics. These harmonics can circulate through the power system, influencing the operation of its components through the production of further losses, heating, and determining an overall efficiency decrease. Some components of the SGPS may become victims of this phenomenon, which overall affects the power quality of the system. This section includes details regarding the classification of components interacting within an SGPS. It delivers insight regarding the ancillary services provided by the EVs, and their implications in terms of battery lifespan. There is a presentation of standards and

norms concerning both the EVs and SGPS reflecting the complexity of the interaction level, including economical aspect as well as energy management.

Section 3 focuses on the charger's classification and performances, as well as on solutions regarding the charging stations and the correlation between the driving habits and the status of the battery.

The main objective of Sect. 4 is the analysis of the optimal integration of electric vehicles into the smart grid. Such a problem is very complex presenting significant considerations related to the distribution of the parking areas and their types of charging stations, the locations assigned to the power regulators and power system stabilizers (PSS) and the performances of such devices, the types of vehicles in the system, the demand in recharging power at the moment or conversely the power delivered by the vehicles and consumed by the grid, the minimization of the losses and costs required by the bi-directional energy transfer through the ongoing efficiency maximization. The case of study and the adjacent simulations performed utilizing the well-known MATLAB program answers to the requirements highlighted above.

The final section contains the conclusions of this work, and the chapter ends with numerous and consistent list of up-to-date bibliographical references.

2 Electric Vehicles, Smart Grid Power System and Bi-directional Energy Flow

The understanding of EV integration into SGPS, nowadays, requires familiarity with the vehicle to grid (V2G) concept, firstly introduced by Kempton [1], in 1997, which proposes the utilization of EVs as sources for the SGPS. An SGPS includes RES, storage units, and eventually, classical energy sources, based on fossil fuel (see Fig. 1) [2]. The structure of an SGPS, according to the National Institute of Standards and Technology (US) contains the following divisions [3]:

- *Generation*
- *Transmission*
- *Distribution*
- *Service Providers*
- *Consumers (customers)*
- *Energy Market*
- *Operations*

At this point, there is no direct reference to the terms such as V2G, aggregator, whereas the customers are not candidates for energy storage, suitable for involvement in V2G action.

The storage units can include or be exclusively EVs, whereas small scale applications may not contain classical sources. Continuously increasing EV fleet around the globe comes with the drawback of the massive charging demand, which strongly impacts the power grid components such as generators, transformers, distribution

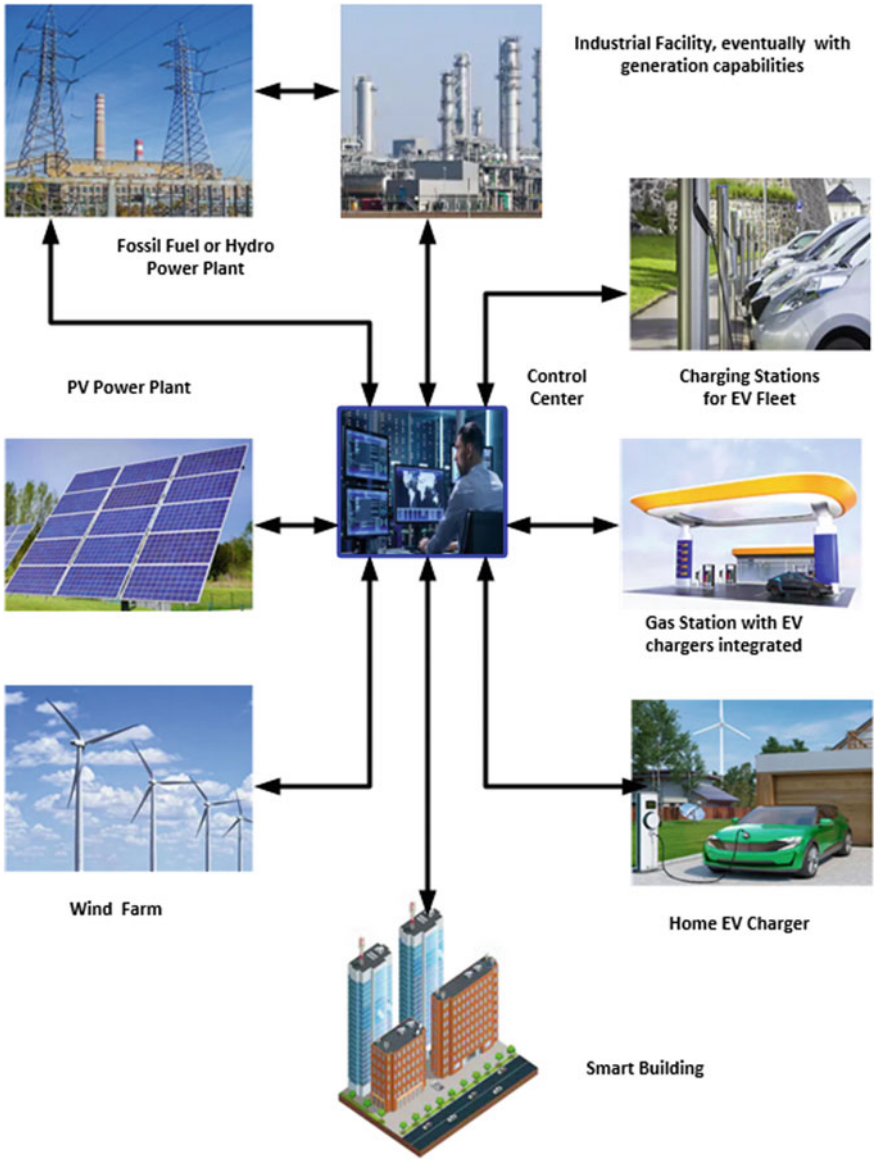


Fig. 1 One smart grid configuration

cables. Interconnection between the local power system, RES and charging stations may represent the critical solution for the mitigation of the drawbacks mentioned above. Moreover, it can lead to a decrease in the amount of greenhouse gases released into the atmosphere and in the EVs charging cost, as well [2–6].

Assuming substantial scale participation of RES, one must account for their significant dependence on weather conditions, generally resulting in either surplus or deficit of delivered energy and, consequently, hardship concerning dispatching it. Participation of the EVs considered working under the V2G technology, can significantly improve the overall SGPS stability by smoothing the power and voltage profiles versus the time, whereas playing the character of storage system for the energy delivered by RES or being an intelligent consumer through an intelligently planned and coordinated EVs battery charging process.

The interaction between the SGPS, RES, and EVs follows three distinct directions [2]:

- *SGPS efficiency improvement*, utilizing the EVs batteries as energy buffers; EVs charging takes place when the RES energy is in excess to the grid absorption availability, whereas EVs batteries deliver electric power to the supply when an SGPS experiences power deficit.
- *Cost-effectiveness*, results in preoccupation for minimization of electricity generation, charging and operation and maximization of the provider's profit.
- *Reduction in greenhouse gas emissions.*

The main difference between EVs and the vehicles utilizing engines with internal combustion resides in the nature of the engine itself. The autonomous EVs primary's source is either a power generator (e.g., diesel-electric or gas-turbine based locomotives, in which cases the prime mover of the generator is an internal combustion engine) or a storage system existing on-board, mostly batteries. Most of the current EV employs the last type of source. There are two main categories of EVs [2]:

Battery Electric Vehicles: The power plant of EVs compatible to V2G technology consists of a rechargeable battery storage system, an electric motor (permanent magnet synchronous, induction or switched reluctance), and a four-quadrant fully-controlled power converter (i.e. composed from two fully-controlled rectifiers, the one 's positive output rail being the negative for the other, and conversely).

Plug-in Hybrid Electric Vehicles: Relying on a battery storage system and a fuel tank as well, the power plant contains two motors for a combined propulsion system, one electric and another one with internal combustion, although they don't operate simultaneously. With the vehicle in the charge depleting mode, the electric motor drives the vehicle, getting its input electric power from the batteries, until the battery reaches the lowest charge threshold. At this point, the control system switches the traction responsibility to the internal combustion motor, which carries the extra responsibility of charging the battery system to the highest state of charge threshold limit, giving the possibility to the controller to switch back to the electric motor. Such an operation is called parallel-hybrid [2]. Lithium-ion batteries, the main elements fulfilling the energy storage function of the duality SGPS–EVs, possess a high density of energy, long life span, and minimal environmental impact. However, they require special attention from the control systems apropos the charging/discharging process, improper voltage/current variations resulting in battery damage. RES is the main component of an SGPS, and the most known types are the wind energy and solar energy, yet both are highly weather dependent.

2.1 Wind Energy Based RES

The appropriate placement of wind farms improves the efficiency of energy conversion from wind to electrical and increases the output of electrical energy. The wind speed determines the amount and the quality of wind energy, and there are several proposed methods for the evaluation of the wind speed and power summarized in [2], using the approach of Peterson and Hennessey:

$$\left(v \middle/ v_r\right) = \left(h \middle/ h_r\right)^a \tag{1}$$

In (1), v is the wind speed at height h, whereas the wind speed v_r is known for the reference height h_r, and the coefficient a belongs to the interval 0.1–0.4.

The amount of wind energy with the potential of conversion into electrical energy, for the wind speed (1), strongly depends on atmospheric conditions such as wind speed and air density [2]:

$$P = 0.5\varepsilon\rho A v^3 \tag{2}$$

The coefficient α designates the constant of Albert Betz, ρ (kg/m^3) is the air density, A (m^2) is the cleared area by the wind turbine, whereas v (m/s) is the wind speed. The employment of several available expressions may correlate the wind power to the wind generator rated value, as in (3a–3c) [2]:

$$P_{wind} = 0, \text{ if } v < v_{\min,w} \text{ or } v > v_{\max,w} \tag{3a}$$

$$P_{wind} = P_{rated}\left(\frac{v^3 - v_{\min,w}^3}{v_{rated,w}^3 - v_{\min,w}^3}\right), \text{ if } v_{\min,w} < v < v_{rated,w} \tag{3b}$$

$$P_{wind} = P_{rated}, \text{ if } v_{rated,w} < v < v_{\max,w} \tag{3c}$$

In (3a–3c), v is the wind speed, $v_{min,w}$ represents the minimum speed of the wind for which the wind turbine is capable to produce electric power, $v_{max,w}$ the maximum wind speed at the wind turbine can safely operate, P_{wind} is the wind power, P_{rated} is the rated power of the wind generator, $v_{rated,w}$ is the rated wind speed.

2.2 Solar Energy Based RES

Representing the primary energy for the other principal RES, solar energy conversion into electrical energy occurs directly through photovoltaic panels (PV) or indirectly from solar to thermal, then into electrical. In the second case, specialized installations, utilizing mirrors and lenses to focus sunlight, target water tanks placed on tall

reinforced mechanical structures. By concentrating the solar power on the radiation absorbent surface of the water tank, the water is transformed into steam, reaching the appropriate parameters for driving a steam turbine as a prime mover for a generator.

PVs are constructed with solar cells (SC), and display an efficiency which slightly varies with the temperature, practically decreasing with 0.2–0.5% for every 10 °C increase [2].

In addition to the SC efficiency η_{SC}, the electrical power output P_{SC} for an aggregated area of SC depends on the surface of its area A_{SC} (m^2), respectively by the solar irradiation s (W/m^2) (4), which is practically the only variable parameter [2]:

$$P_{SC}(s) = \eta_{SC} s A_{SC} \tag{4}$$

The output electric power of a PV solar panel depends on its area A_{PV} (m^2), its efficiency η_{PV}, respectively on its solar insolation $\mu(t)$ [2] (5):

$$P_{PV}(t) = \eta_{PV} \mu(t) A_{PV} \tag{5}$$

Modeling of wind and solar energy to improve the quality of feasibility studies regarding this RES placement, integration, and utilization, required linear, nonlinear and artificial intelligence to predict the wind and solar energy in specific locations [2–6]. The integration of the EVs into the SGPS occurs under the surveillance of a third party, called *aggregator*. Even though the main feature of the V2G technology is the bi-directional power flow between EVs and SGPS, meaning that the EVs are electrical energy buffers utilized to compensate for the RES intermittency and weather determined irregularities in energy delivery, shaving the peaks and leveling the load "seen" by SGPS, there are ancillary services requested from EVs [4]:

- *Power regulation*, which means the frequency regulation, matching the energy generation with the load demand.
- *Warm or Spinning reserve*, representing the short term (less than 10 min) fast response additional generation capacity to act in case of outage rapidly.
- *Reactive power compensation*, utilizing the reactive power delivered by the capacitor present in the DC link of the fully controlled four-quadrant drive.

The owners of the EVs integrated into the SGPS must have the acknowledgment for the battery lifespan reduction due to a continuously variable charging-discharging cycle, in the form of proper compensation. The integration of EVs into SGPS has a multitude of impacts on the SGPS. Many authors agree [3, 7–28] on the EVs presence impacting the operation of the *distribution systems*, *system's equipment*, *load capacity*, *power quality*, *economy, social life* and *nevertheless, the environment.*

The impact of the EVs penetration on distribution system reliability, presented in [11] assumes different levels of EVs penetration in the SGPS, in the presence of V2G technology.

The factor named Expected Energy Not Charged (EENC), measures the reliability of the system from the viewpoint of EVs integrated with the SGPS, and requires the addition of the energy loss during charging to the energy loss because of unscheduled

V2G. The authors signal better reliability of the SGPS with distributed generation in the presence of EVs.

For example, whereas applying a stochastic model of the EVs energy needs and starting from a load flow study, Anastasiadis et al. [12] concluded that the voltage levels at the nodes where EVs are present fluctuate in a prescribed range, with the application of a smart charging strategy within V2G.

In [13], there is an assessment of the transformers aging due to the penetration of the EVs, using a stochastic method modeling of the transformer's life consumption. Here, the load model applied to EVs, originated from an analytical solution used for predicting a cluster of EV chargers. The authors disclosed results that show a severe 44.1 h of the loss of life for every 24 h of operation, during the summer and with an EV penetration of 50%, whereas a penetration of 10% or less does not decrease transformer's life. However, the transformer's life loss depends on the load level and its temperature at each moment. High EVs penetration level unfavorably impacts the power transformer lifespan, following a study involving the calculation and analysis of the dielectric oil deterioration [14], showing consistency with [13].

Furthermore, there are economic and social aspects of the EVs integration debated in [15–18]. For example, in [15], a study targeting EV users in Germany concluded the presence of substantial concern, especially for high mileage drivers, regarding the range decreasing and loss of availability to drive the vehicle during the long hours of connection to the SGPS. The authors found that a promising remuneration for participating in V2G couldn't mitigate the concern. The costs regarding infrastructure, electricity sale by the vehicles and the battery degradation appeared the most important factors regarding the feasibility of the V2G implementation, to obtain income from the energy supplied to a commercial building [16].

The cost of EVs operation, together with the rules governing the energy market, have an influence on the participation of the EVs into a centralized V2G system [17]. This study indicates a profitable grid (here called virtual power plant), a system operator with fulfilled services and EVs being on the money-losing side. Cost elements face the revenue element in a study focused on determining the critical economic factors of the V2G technology implementation [18]. The factors considered in the *cost* side are the battery lifespan reduction and replacement, the V2G capable power electronic converters (i.e., many vehicles need retrofitting and installation of such converters), the infrastructure of charging stations, nonresidential and residential, the hardware equipment and its specialized software required for communication and the aggregator as network manager. An aggregator has proven to be costly (i.e., may necessitate up to 50% revenue for covering its expenses and obtaining some profit) as well, directly associated to its main tasks of EVs control, networking services, contract management with EV owners, activity on energy markets and availability for unexpected services. The elements offering potential *revenues* are the pool of aggregated EVs, daily number plug-in hours, the power availability (kW) for transfer from EV fleet to SGPS, the level of available energy (kWh) for specific conditions of demand and price.

One can perceive the EVs as receptive loads, whereas integrated into SGPS provide ancillary services and reserve power under proper scheduling from the aggregator, responding to the energy and reserve requirements. Such scheduling positively influences the load profile by flattening it through charging/discharging time allocation [19]. However, in [20] the authors extrapolate the aggregator's role from the business controller into the highly bonded interface between the EVs and SGPS, meant to implement and monitor the EVs function of providing ancillary services.

The EVs integration into SGPS displays not only potential benefits but very many challenges, as well. The overall interaction between the EVS and SGPS under the V2G technology requires proper scheduling mechanisms, to address the challenges, fact which turns the need for *optimal scheduling* into a priority [21–28]. There are many objective functions possible to define and undertaken by scientists, very much trying to answer to individual objectives, associated with real applications. In [21, 22], the authors raise the problem of dimensioning and placement for a distributed generation system, built with small power units, issues followed by the bi-directional power flow investigation. In [21], the debated problem has the first stage, while the clear objectives are the power loss minimization and voltage regulation. In the second stage, the objective function represents the amount of RES, including the EVs, followed by the optimization of the rate of charge. Consequently, to the application of a genetic algorithm to a system with 30 buses, the authors found the optimum power ratings of the distributed generators as 450 kW and 550 kW, respectively. Moreover, the minimization of the power loss indicated the placement for EV chargers.

In [22], there is an inclusion of RES (i.e., 10 wind turbines and 117 photovoltaic panels) in a distributed generation system, which becomes capable of sophisticated control of the EV batteries charging and discharging, minimizing the power loss and regulating the voltage as well. Intelligent scheduling considering the EV batteries charging and discharging process in conditions of energy price constrains, initiated a study focused on maximum profit demanded by the EV owners through a proper charging/discharging time rate [23].

Although the applications from [21–23] addressed common parking lot facilities for EVs, in [24], the authors proposed a study aiming to solve the optimal control of smart buildings with EV charging capabilities, with the objectives of minimizing the power loss and maximizing the comfort for customers. The optimizing technique is a combination between the particle swarm algorithm and the multi-agent technology. Then, a carried-out simulation employed an extended system with 500 wind generators and 400 photovoltaic panels in a time frame of 24 h. Simulation has proven the possibility to acquire a high level of comfort, even during periods with energy deficit recorded for isolated buildings.

A solution to the energy concern in a small-scale system, including aggregated EVs and working within V2G technology, exists in the form of a robust optimization model based on linear programming [25]. In such a model, the authors suggest an assessment of the V2G impact on the system's energy management. In the case of EVs presence in large numbers, there is a demand for coordination between the EV

fleets given their role as renewable sources as well as storage units within the bi-directional power flow under V2G technology. Simulation results from [26], recommend an optimal scheduling power flow inside the SGPS, whereas the EV fleets are stationary. Frequency regulation is one of the primary ancillary services potentially provided by EVs, together with their role as load leveling and on-waiting (i.e., instant readiness) reserves. However, SGPS frequency regulation may result in Li-ion battery degradation [27]. The mitigation of such a drawback became possible with the application of a smart algorithm pointing to the maximization of the battery life span. More precisely, the battery of an EV connected to the grid can deliver energy only if it is fading, evaluated through a proper forecast model about reaching its lowest acceptable level. If did, the respective battery can only receive electrical energy from the supply.

Frequency regulation with the participation of EVs is the main topic in [28] as well. This time the deciding factor being the battery state-of-charge (SOC), assessed by calculating the owner's driving demand with a specially designated module. Financial aspects of the EVs integration into SGPS are essential for the EV owners, willing to allow their cars to provide auxiliary services to the grid, as well as a stand-by reserve to the grid. However, two studies from 2016, [29, 30] recognized that the revenues from aggregators are by far insufficient to make the V2G technology attractive (i.e., from 20 times less to 27,500 less) under the present-day market and regulations in place. The EV battery in V2G is facing a permanent cycle of charging-discharging if stays connected to the power system, and consequently, its life span shortens, and there is no mechanism in place to encourage the owners to participate in V2G other than potential government subsidies. Moreover, the relatively reduced number of EVs didn't assure the stand by (aka fast) firm reserve for more than a few hours. The most optimistic prediction regarding V2G cost-effectiveness mentions the year 2030. Consequently, one must address the amount of annual profit to satisfy all, the aggregator and the EVs owners into a joint venture towards V2G. The integration of EVs into SGPS, under several business models, resulting in contracts between the SGPS owners and the aggregators, must consider the EVs battery charging stations, containing data centers, as essential participants in SGPS regulations [3]. The National Institute for Standards and Technology (NIST, US) decides the standardization direction for security and reliability on the smart grid. For example, the guidelines for SGPS cybersecurity have its place into NISTIR 7628. To be feasible, an actual SGPS must fulfill many conditions, such as [3]:

(1) *Participation of RES with possibility of increasing it*
(2) *High power quality*
(3) *High resilience to disruptive events*
(4) *Increased operational efficiency of its components and overall*
(5) *Robustness to disturbances and self-healing*
(6) *Capability to accommodate drive storage units*
(7) *High level of cyber security and capability of ongoing improving it*
(8) *Flexibility in terms of consumer's selections*
(9) *Smart revenue metering*

(10) *Reduced amount of greenhouse gases by encouraging the progressively larger participation of EVs, actively working under V2G technology: bi-directional electric power flow.*

Subsequently, understanding the actual standards applicable to EV and SGPS becomes vital. Whereas Table 1 presents the standards for EVs elaborated by the Society of Automation Engineers (SAEJ), in the Table 2 (IEC), respectively Table 3 (IEEE) one can find the most critical standards in place, referring SGPS, from all over

Table 1 Most important SAEJ standards for EV, US [3]

SAEJ standard	Domain
1772	Defines the functions of the of the vehicle supply equipment as: coupler, rectification and voltage regulation
2293	Addresses the EV requirements and the off-board equipment used for EV charging
2836	Communication requirements for integrating the vehicle in a V2G technology, specifically to assure the existence of the functions of energy transfer and or energy storage
2847	Communication between the EV and the fast, high power DC off-board charger
2894	Off-board charging practices

Table 2 Most important IEC standards for SGPS [3]

IEC standard	Domain
61000-4-30	Power quality
618500	Communications
14908-3	Power line communication
618500-7-420	Communications including reference to microgrids
61968-9	Advanced/Automated metering
62351	Protection of smart grid
61970	Communication and metering data exchange
62056	Communication and metering data exchange
61968	Communication in distribution
11518-2	Communication between smart grid networks
TC57-WG13	Cybersecurity and reliability in smart grids
60870	Inter-control center protocol
62056	Metering, load control an tariff
14543	Home electronic system architecture
61400-25	Operation of the wind power plants

Table 3 Most important IEEE standards for SGPS [3]

IEEE standard	Domain
P2030	Interoperability requirements
C37.118.1-2011	Security
1588	Smart grid interoperability and control
2030-2011	Smart grid interoperability and control, communication infrastructure
1377	Advanced/Automated metering
1547.4	Micro grid
1547.6	Distributed networks
1451	Smart sensors
145.1	Wireless sensor network
1646	M2M communication
802.16	M2M communication
61499	Control in smart grid
ZigBee 802.15.4b (wireless, V2G, dedicated for installation at client)	Communication frequency: 2.4 GHz (anywhere in the world) Distance covered: 10–100 m
Wi-Fi 802.11 g (wireless, V2G)	Communication frequency: 5.85–5.95 GHz Distance covered: 0.5–1 km
Bluetooth 802.15.1a (wireless, V2G)	Communication frequency: 2.4 GHz Distance covered: 1–100 m
WiMAX 802.16 (wireless, V2G)	Metropolitan area network Frequency: 2–6 GHz Distance covered: 2–5 km
6loWPAN 802.15.4	Communication
WRAN 802.22	Communication
901	Power line communication
802.3ah	Optic fiber communication
802.3	Ethernet
1701	Metering devices
1702	Metering devices

the world [3], and their development responded in essence to six directions: operation, power quality, metering, communication, protection & control, and cybersecurity.

3 Requirements for PEV Charging Stations and Smart Grid Power System

A station used for charging EVs is a system consisting on four main subsystems (see Fig. 2), namely:

- Charging subsystem (CS)
- Energy Storage System (ESS)
- Electric Power Distribution subsystem (EPDS)
- Auxiliary Energy subsystem (AES)

The duty cycle of a charging station for EVs can have a duration of 24 h. Such a duty cycle strongly depends on the user's daily schedules, with the highest demand recorded during the daytime because of the heavy traffic. Consequently, the modeling of the EV charging process considers three periods:

– *Model 1*, called of high energy consumption, with a duration denoted with *Tr* between 7.00 am and 3.00 pm
– *Model 2*, called of medium energy consumption, with a duration *Tm*, between 3.00 pm and 11.00 pm
– Model 3, called of low energy consumption, with a duration *Ts*, between 11.00 pm and 7.00 am next day

Using the previous notations becomes obvious that $Tr + Tm + Ts = 24\ h$. The demand for efficiency in conditions of lowest possible electrical energy market price encouraged the EVs users to proceed for battery charging only during night time, between 11.00 pm and 7.00 am. There are several charging possibilities available for EV users, whereas the Standard IEC 61851 defines four charging modes [3]:

Charging mode 1. In this mode, the charging location represents a standard outlet, generally used for domestic appliances at 120 or 230 V AC.

Whereas the EV is in *charging mode 2*, the electrical connection of the EV battery to an AC outlet for general-purpose requires a standard cable that incorporates a control device, capable of interacting with the outlet. Such a control device keeps the battery charging current below its maximum threshold, monitoring the outlet quality as well, more precisely the voltage drop level as a good indicator for an over temperature condition. Automatically, the battery charger turns-off in case of a fault.

In the *charging mode*, *3* the outlet comes connected to a specially designated circuit; this fact assures the real-time communication between the EV and the electric installation. There is a need for the installation of a specialized electronic device at the location.

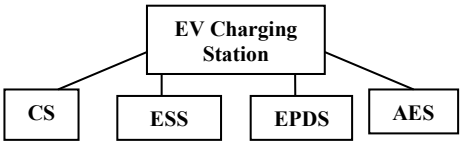


Fig. 2 The structure of a station used for charging EV

Charging Mode 4 refers to the fast charge using a DC outlet, utilizing an external charger which includes the charging cable as well. The DC charger delivers the charging voltage itself. Characteristics of such a charging type are the elevated values for the voltage and current. Whereas the DC charging is in place, the connection cable doesn't separate from the charger itself and, consequently doesn't exist any connection to the battery terminals. One can find this type of outlet only on the fast charging terminals. Such an assembly must comply with the world standard CHAdeMO.

The AC charging of type E matches the domestic outlets becoming suitable for charging modes 1 and 2. The so-called plugs of type 1 or type 2 have a clear designation for EVs and belong to the charging stations or the domestic chargers such as Walbox. The domestic or "at home" charging comprises 95% of the total EV battery charging. For single-family homes, the whole process becomes quite simple and relatively easy to monitor. However, not every EV user lives in such a home, a fact which complicates the charging issue.

Moreover, for certain EV brands, the Walbox Mode 3 is a must, as their connection to the standard outlet is not possible, excluding the charging Mode 1 as an option. The Walbox, operating in charging Mode 3, comes highly recommended by its security features, yet financial reasons make many users reluctant to it. Several hours of charging, absorbing 8–10 A of current can potentially overheat the electric cables existing on and involved in the process. Consequently, the whole charging process requires careful monitoring and control. Walbox Mode 3 presents the advantage of the highly safe operation, practically eliminating any electric hazard and providing elevated protection to people and livelihood. For such a solution, the internal microprocessor can permanently control the voltage, current, the charge duration, and its cost, data that becomes accessible to users due to the communication with the EV.

The regular charging outlets from public spaces assure 1–3 h of charging, depending on their power rating and the type of EV as well. Furthermore, the DC fast charging outlets impress the direct current into the batteries, through their rectifiers. The DC charger communicates the same essential information (current, voltage, power, and cost) to the EV through the charging cable.

One classifies the DC outlets as intelligent due to their ability to communicate directly with the EV, delivering data regarding the energy consumed for charging, optimally, and safely. First-generation EV batteries allow the recharging to 80% state of charge from a DC charger in about 20–30 min. Nowadays, the lithium-ion technology made possible manufacturing new batteries with a reduced charging time of 5–15 min to the same 80% state of charge. A typical DC charger can deliver a voltage of 400–500 V, a current of 100–250 A, for a total power of approx. 50 kW. Battery replacement gains terrain at the level of small EV, after proven effective for electric buses used for public transportation. The replacement process takes place in about 3 min (i.e., duration similar to the filling of a full tank) inside of special fully automated dedicated stations. According to SAEJ1772, each type of chargers has three levels, whereas fulfilling the recommended following specifications [3]:

- The AC type charger, Level 1 has single phase voltage, rated at 120 V, current ratings of 12 A or 16 A and consequently the power ratings of 1.4 kW, respectively 1.9 kW. The charging time assigned for the EV is 17 h and for PHEV 7 h.
- The AC type charger, Level 2 has either single phase or three-phase voltage rated at 240 V, a current rating of 80 A, and consequently the power rating of 19.2 kW. The charging time assigned for the EV is 7 h and for PHEV 3 h.
- The AC charger, Level 3 must assure more than 20 kW, without specifying any other characteristics for the present day.
- The DC type charger Level 1 comes with a voltage rating between 200 and 500 V, a current rating less than 80 A, a power rated at 40 kW, a three phase supply connected to the input of the rectifier. The charging time assigned for the EV is 72 min and for PHEV 22 min.
- The DC type charger Level 2 comes with a voltage rating between 200 and 500 V, a current rating less than 200 A, a power rated at 100 kW, a three-phase supply connected to the input of the rectifier. The charging time assigned for the EV is 20 min and for PHEV 10 min.
- The DC type charger Level 3 comes with a voltage rating between 200 and 600 V, a current rating less than 400 A, and a power rated at 240 kW. There are no other specifications for the present time.

Wireless charging requires the presence of two coils. The first one, called the emitter, installed in a box fixed at the ground level and connected to a power supply, produces a magnetic field. The magnetic coupling between the emitter and second coil called receiver placed below the EV allows the magnetic field to induce a voltage across the second coil. The output of the receiver which is directly connected to the input of a rectifier, converts the induced voltage into DC voltage. Furthermore, a filter flattens the voltage shape, which is then applied to the battery terminals, charging it. High losses and low efficiency of about 20%, maintain this charging mode at the level of work in progress, necessitating future research work.

Presently the EV user especially counts on public charging stations, whereas most of the owners depend on home chargers, fast-charging stations or battery replacement stations. Nowadays, there are not enough fast-charging stations nor battery replacement stations: the public charging infrastructure satisfies only 75% of the "100 km/day" EV users. This lack of enough fast charging and or replacement facilities, forces the EV users towards home charging solutions, triggering a significant impact on the energy market. Such a standpoint makes the charging stations to look like completion to home charging outlets. However, current limitation makes the fast charging impossible at home level.

For home charging, one can count for maximum 1/3 of battery charging capacity in three hours; if this is not enough, then highly demanded public fast-charging stations can bring the EV batteries to full charging state in less than 1 h. Although, the large-scale deployment of solar cells, integrated into SGPS under continuous development, may result in the opportunity to provide homes with solar panels enhancing performances of the home chargers. The power and energy densities are the two critical

factors concerning the battery fabrication technologies which face rapid development. One can expect an increase in these two factors predicting an upturn in the EV autonomy, as well.

The integrated power stations represent a new concept, and the construction of such a station must fulfill specific rules, for risk mitigation. Integrated charging stations may be candidates for "safe distance" like requirements, presently applied to infrastructure adjacent to gas stations.

In one suggestive example, one affirms that the safe distance between the charging equipment and the electric power station must be at least 18 m. Moreover, the electrodes from inside the lithium-ion batteries according to their fabrication material belong either to class C (i.e., solid combustible) or class D. In this way the battery represents a synthesis of different materials, possibly assessed as depletion of class C materials.

In another example, one aims the safe distance between the supplied equipment and the warehousing area for batteries: such a distance should be more than 15 m. Following numerous standards and experiences, one can choose the appropriate placement of the charging stations for EVs, mitigating the risks associated with technologies of operation when integrating the chargers into gas stations. A safe distance of 18 m between the gasoline tank and charger's equipment allows the addition of chargers to the gas station, as shown in Fig. 3.

There are two solutions regarding battery charging management, suitable for *integrated charging stations*:

(1) The first solution requires the battery charging to take place inside the charging stations, with the following advantages and drawbacks:

   (1a) one can replace the fading batteries at the right time; there can be an improvement in the EV battery utilization

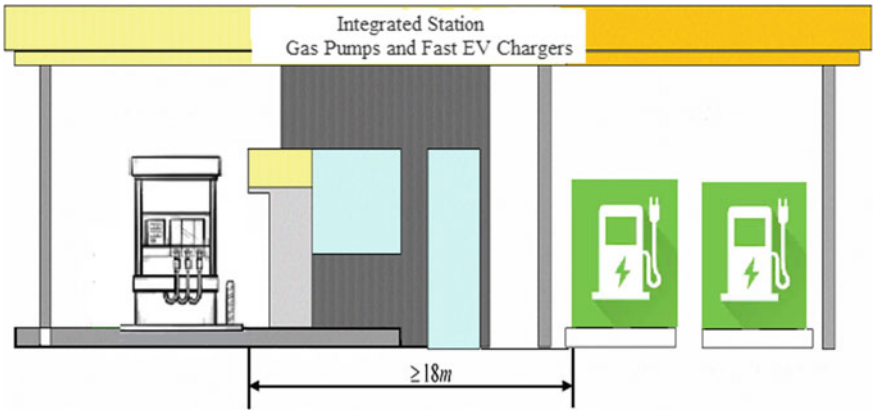


Fig. 3 Gasoline station with integrated chargers for EVs

(1b) integrated charging stations demand a more performant energy distribution capability, to assure enough energy for battery charging
(1c) more electric equipment in operation during the battery charging process requires a higher level of safety

(2) The second solution considers the battery charging done inside of special stations built with this purpose in mind, followed by the battery distribution to integrated stations in need. This solution has some advantages and drawbacks, such as:

(2a) requires a distribution system
(2b) both special and integrated stations require a larger battery depositing space, a fact which results in improperly low utilization of space
(2c) the delay of battery replacement requires more batteries available for exchange to satisfy EVs necessities.

4 Case Study: Modelling Individual and Compounded V2G

This section contains a study regarding the integration of the EVs into the SGPS, addressing the processes of PEV battery charging and discharging.

4.1 Battery Charging Modelling

In the model representing the charging process, the location of the battery is always at the output of the power electronic converter which operates as a rectifier [11, 31]. The battery described by an equivalent capacitor *C* (see Fig. 4) represents a fair alternative regarding the analysis of the charging and discharging processes.

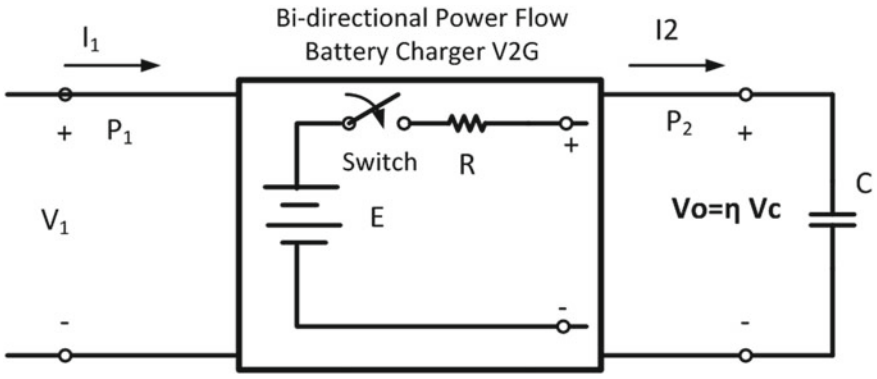


Fig. 4 Bi-directional power flow battery charger (V2G): equivalent circuit

However, the charging circuit is not a 100% efficient one, fact which explains the presence of a resistor R that characterizes the losses accounted for the charger. The rectifier, being the intermediate circuit between the charger's power supply and its output is stiff enough to provide a constant DC (rectified voltage) $V_{\text{int},rec}$ voltage. The switch meant to close the battery charging circuit turns on at the instant t_1, triggering the appearance of the charging current $i_1(t)$.

$$V_{\text{int},rec} = V_R(t) + V_{out}(t) \tag{6a}$$

$$V_{\text{int},rec} = Ri_1(t) + \eta_C V_{Cap}(t) = \eta_C \frac{1}{C} \int_{t_1}^{t} i_2(\tau) d\tau \tag{6b}$$

The definition of the charger's efficiency in per-unit involves the ratio between the charger's output power to its input one (7).

$$\eta_C = \frac{P_{out}}{P_{in}} = \frac{V_{out} i_{out}}{V_{in} i_{in}} \tag{7}$$

One can express the Eq. (6b) in terms of currents following a time derivation, with the charging time constant $T_C = RC$:

$$R\frac{di_2(t)}{dt} + \frac{\eta_C}{C} i_2(t) = \frac{dV_{\text{int},rec}}{dt} = 0 \Rightarrow \frac{T_C}{\eta_C}\frac{di_2(t)}{dt} + i_2(t) = 0 \tag{8}$$

The solution for differential equation (8) comes the easiest whereas performing an intermediate Laplace conversion:

$$\frac{T_C}{\eta_C}[sI_2(s) - I_2(0)] + I_2(s) = 0 \Rightarrow I_2(s) = \frac{\frac{T_C}{\eta_C} I_2(0)}{\left(s + \frac{T_C}{\eta_C}\right)} \tag{9}$$

The initial value of the charging current is equal to $I_2(0) = V_{\text{int},rec}/R$, fact which leads to the current [31]:

$$i_2(t) = \frac{V_{\text{int},rec}}{R} \exp(-\eta_C t / T_C) \tag{10}$$

Introducing (10) into (6a and 6b), one can express the instantaneous voltage across the battery, considering the battery voltage at the beginning of the charging process being equal to $V_{out}(t_1)$:

$$v_{out}(t) = V_{\text{int},rec}\left[1 - \exp(-\eta_C t / T_C)\right] + V_{out}(t_1) \tag{11}$$

The battery status in terms of power requires the multiplication of each term of the Eq. (11) with the current expressed according to (10). It results the expression (12) which links the instantaneous power to the maximum power value and to the initial power as well.

$$P_{out}(t) = \frac{V_{\text{int},rec}^2}{R}\left[1 - \exp(-\eta_C t/T_C)\right]\exp(-\eta_C t/T_C) + \cdots$$
$$V_{out}(t_1)\frac{V_{\text{int},rec}}{R}\exp(-\eta_C t/T_C) \tag{12}$$

Whereas defining the maximum power at the battery as P_{max} (13a), respectively the initial power as $P(t_1)$ (13b) [11, 31]:

$$P\text{max} = \frac{V_{\text{int},rec}^2}{R}\exp(-\eta_C t/T_C) \tag{13a}$$

$$P(t_1) = V_{out}(t_1)\frac{V_{\text{int},rec}}{R}\exp(-\eta_C t/T_C) \tag{13b}$$

One can express the instantaneous charging power as in (14).

$$P_{out}(t) = P\text{max}\left[1 - \exp(-\eta_C t/T_C)\right] + P(t_1) \tag{14}$$

Considering the case of a PEV batteries during the charging process (power demand), the instantaneous power of the battery is [11, 31]:

$$P_{BPEV}(t) = P_{BPEV,\text{max}}\left[1 - \exp(-\alpha\eta_C t/T\text{max})\right] + P(t_1) \tag{15}$$

The relation (16) presents the link between the efficiency of the charger η_C, charging constant of the battery α, charging time constant T_C, respectively the maximum time T_{max}. However, relation (15) lasts until the battery charging level reaches its fully charged required value $t_1 \leq t \leq t_{char}$. T_{max} represents the duration of battery to fully charge from zero power.

$$T_C = \left(\frac{\eta_C}{\alpha}\right)T\text{max} \tag{16}$$

The initial PEV battery power level determines the duration of the charging process, if the battery reaches its prescribed maximum power value. In such a way, one can find the duration necessary for full charging as (17):

$$t_{char} = \begin{cases} 0 & P(t_1) = P_{BPEV,\text{max}} \\ T & P(t_1) = 0 \\ t_{char} & 0 < P(t_1) < P_{BPEV,\text{max}} \end{cases} \tag{17}$$

Combining the relations (15)–(17) and imposing $P_{BEV}(t_{char}) = P_{BEV,max}$, one can express the charging time as following (18):

$$\begin{aligned} &P_{BPEV}(t_{char}) = P_{BPEV,\max} \\ &P_{BPEV,\max} = P_{BPEV,\max}\left[1 - \exp(-\alpha\eta_C t_{char}/T_{\max})\right] + P(t_1) \\ &P_{BPEV,\max}\exp(-\alpha\eta_C t_{char}/T_{\max}) = P(t_1) \\ &\alpha\eta_C t_{char}/T_{\max} = -\ln\left(\frac{P(t_1)}{P_{BPEV,\max}}\right) \Rightarrow t_{char} = -\frac{T_{\max}}{\alpha\eta_C}\ln\left(\frac{P(t_1)}{P_{BPEV,\max}}\right) \end{aligned} \tag{18}$$

The PEV power battery level, expressed for all the conditions is (19):

$$P_{BPEV}(t) = \begin{cases} P_{BPEV,\max}\left[1 - \exp(-\alpha\eta_C t/T_{\max})\right] + P(t_1) & t_1 < t < t_{char} \\ P_{BPEV,\max} & t \geq t_{char} \end{cases} \tag{19}$$

One can estimate the total power demanded by the PEV battery for a full charge, starting from its initial status $P(t_1)$ as (20):

$$P_{BPEV,dem}(t) = P_{BPEV,\max} - P_{BPEV}(t) \tag{20}$$

The expression (20) detailed for the full charging interval becomes (21a, 21b):

$$P_{BPEV,dem}(t) = P_{BPEV,\max} - P_{BPEV,\max}\exp(-\alpha\eta_C t/T_{\max}),\ t_1 < t < t_{char} \tag{21a}$$

$$P_{BPEV,dem}(t) = 0,\ t \geq t_{char} \tag{21b}$$

Several scenarios in place lead to the evaluation of the battery charging in normal conditions and/or in presence of faults. For instance, if the PEV battery charging started at t_1 is scheduled for a duration Δt_1 until the full charging completion, usually measured in hours [11]. A fault occurrence may stop the charging process prior to the full battery charge, let's say after an interval Δt_2. The battery power level in such a case is (22):

$$P_{BPEV,dem}(t_1 + \Delta t_2) = P_{BPEV,\max}\exp(-\alpha\eta_C \Delta t_2/T_{\max}) - P(t_1),\ \Delta t_2 < \Delta t_1 \tag{22}$$

If the fault occurs after the time scheduled for the PEV battery to reach its maximum charging level, then:

$$P_{BPEV,dem}(t_1 + \Delta t_2) = P_{BPEV,\max},\ \Delta t_2 > \Delta t_1 \tag{23}$$

Whereas the fault is restored at the time $t_1 + \Delta t_3$, assuming no battery discharge during the time of fault occurrence and the time of operation restoration, one can assume a new initial power status for the battery power level (24):

$$P_{BPEV,dem}(t_1 + \Delta t_3) = P_{BPEV,\max}(t_1 + \Delta t_2),\ \Delta t_3 > \Delta t_2 \tag{24}$$

Finally, the maximum battery power level appears after $t_1 + \Delta t_3 + \Delta t_1 - \Delta t_2$. One can calculate the overall energy required for the vehicle battery to fully charge, whereas the charger is unavailable because of a fault which lasts a certain time $\delta T_{Fault} = \Delta t_3 - \Delta t_2$ in terms of power and time (25).

$$E_{BPEV} = \begin{cases} \int_{t_1}^{t_1+\Delta t_2} P_{BPEV,dem}(\tau)d\tau & \Delta t_2 < \Delta t_1,\ t_1 < \tau < t_1 + \Delta t_2 \\ 0 & t_1 + \Delta t_2 < \tau < t_1 + \Delta t_3 \\ \int_{t_1+\Delta t_3}^{t_1+\Delta t_3+\Delta t_1-\Delta t_2} P_{BPEV,dem}(\tau)d\tau & t_1 + \Delta t_3 < \tau < t_1 + \Delta t_3 + \Delta t_1 - \Delta t_2 \end{cases} \tag{25}$$

In a parking lot with identical chargers (i.e., the same efficiency for each charger), whereas assuming the same level of maximum power, yet different charging rates α_i and different charging durations t_{chari}, one can express the total instantaneous power demand as [31] as a function of the total number of the PEVs.

$$P_{BPEV,dem}(t) = n_{PEV,C} P_{BPEV,\max} - P_{BPEV,\max} \sum_{i=1}^{n_{PEV,C}} \exp(-\alpha_i \eta_C t_{chari} / T\max) \tag{26}$$

4.2 Vehicle to Grid Modelling

The battery of a single PEV connected to the SGPS delivers power into the grid through the power electronic converter which operates in the inverter mode, with an efficiency η_{inv}. The battery discharge rate is γ_{dis}. If the process of power delivery into the grid starts at t'_1, with the initial battery power level $P(t'_1)$, then the power discharging level at an instant t' is equal to:

$$P_{BPEV,dis}\left(t'\right) = P\left(t'_1\right) \exp\left(-\gamma_{dis}\eta_i t' / T\max\right) - P_{BPEV,Lim} \tag{27}$$

The power limit level $P_{BPEV,Lim}$ represents the PEV battery power threshold necessary for starting the PEV engine, or driving it in electrical mode. In the latest situation, the battery cannot start the engine, yet the charger must do it, whereas the battery

must sustain the EV driving [31]. The power injected into the grid in equal to:

$$P_{BPEV,inj}\left(t^{'}\right) = P\left(t_1^{'}\right) - P_{BPEV,dis}\left(t^{'}\right) \tag{28}$$

Consequently, introducing (28) into (27), one can express the power injected by the battery of a single PEV into the grid as (29).

$$P_{BPEV,inj}\left(t^{'}\right) = P\left(t_1^{'}\right)\left[1 - \exp\left(-\gamma_{dis}\eta_i t^{'}/T\max\right)\right] \tag{29}$$

The duration of the power injection into the grid strongly depends on the battery discharging constant γ_{dis} as well as well as the limit power P_{Lim}. At the end of the allowed discharging time t'_{dis}, the PEV battery discharging level is equal to zero.

$$P_{BPEV,dis}\left(t_{dis'}^{'}\right) = 0 = P\left(t_1^{'}\right)\exp\left(-\gamma_{dis}\eta_i t_{dis}^{'}/T\max\right) - P_{BPEV,Lim} \Rightarrow$$
$$t_{dis}^{'} = -\frac{T\max}{\gamma_{dis}}\ln\left[\frac{P_{BPEV,Lim}}{P\left(t_1^{'}\right)}\right],\ P_{BPEV,Lim} \le P\left(t_1^{'}\right) \tag{30}$$

If the power is injected into the grid for a duration equal to $\Delta t'$, in complete fulfillment of condition (27), then the energy impressed into the grid is:

$$E_{BPEV,inj} = \int_{t^{'}}^{t^{'}+\Delta t^{'}} P\left(t^{'}\right)\left[1 - \exp\left(-\gamma_{dis}\eta_i t^{'}/T\max\right)\right]dt;\ t^{'} \le t_{dis}^{'} \tag{31}$$

If instead of one PEV there are n_{PEV} PEVs, injecting power to the grid for the durations $t_{sup,k}$, having the discharging constants of the batteries $\gamma_{dis,k}$ and the initial powers $P_k(t'_1)$, one can evaluate the total power impressed into the grid at the instant t' as (32):

$$P_{BPEV,inj}\left(t^{'}\right) = \sum_{k=1}^{n_{PEV,D}} P_k\left(t_1^{'}\right)\left[1 - \exp\left(-\gamma_{dis,k}\eta_i t_{\sup,k}/T\max\right)\right] \tag{32}$$

However, the presence of the overall V2G interaction has a dynamic character, there are PEV with batteries under charging, vehicles which retrieve energy into the SGPS, and their number is in continuous change. The authors of [31] assembled a model of a sized parking lot, suitable for estimating the net value of the power submitted or received from the SGPS, as well as the net energy impressed to the SGPS, whereas the vehicles are parked and connected. Using probabilistic calculations, the modeling study continues with the inclusion of the influx, respectively outgoing vehicles, the service duration allocated to vehicles, and the approximation of the battery initial charging level. The approach to the PEV diversity relies on the

differences in charging and discharging coefficients. Finally, the addition of some renewables to the PEV parking lot offer a more complete picture.

4.3 Modelling of the EV Parking Lot Operation

The resultant (net) power exchange between the PEV fleet and the grid (33) represents the difference between the injected power and the power demanded by the vehicles for charging (26), (32):

$$P_{BPEV,net} = \sum_{k=1}^{n_{PEV,D}} P_k\left(t_1^{'}\right)\left[1 - \exp\left(-\gamma_{dis,k}\eta_i t_{\sup,k}/T\max\right)\right] - P_{BPEV,\max}\left[n_{PEV,C} - \sum_{i=1}^{n_{PEV,C}} \exp(-\alpha_i \eta_C t_{chari}/T\max)\right] \quad (33)$$

The net energy injection into the grid results from the integration of (33) over a certain amount of time of interest. Moreover, imposing the non-existence of charging over a 24 h time span, one can estimate the energy retrieved into a storage facility and obtain important data to design such a facility [31]. Such an energy expression displays the form:

$$E_{BPEV,net} = \sum_{k=1}^{n_{PEV,D}} P_k\left(t_1^{'}\right)\left\{t_{\sup,k} - \frac{T\max}{\eta_i t_{\sup,k}}\left[1 - \exp\left(-\gamma_{dis,k}\eta_i t_{\sup,k}/T\max\right)\right]\right\} - P_{BPEV,\max}\frac{T\max}{\eta_C}\sum_{i=1}^{n_{PEV,C}}\left\{\frac{\eta_C t_{chari}}{T\max} - \frac{\left[1 - \exp(-\alpha_i \eta_C t_{chari}/T\max)\right]}{\alpha_i}\right\} \quad (34)$$

The power, respectively the energy gain factors are defined by dividing the power, respectively the energy inserted into the supply, respectively by the product of the number of charging stations to the maximum power of a station.

$$G_{Power} = \frac{\sum_{k=1}^{n_{PEV,D}} P_k\left(t_1'\right)\left[1 - \exp\left(-\gamma_{dis,k}\eta_i t_{\sup,k}/T\max\right)\right]}{N_{st}P\max} \quad (35)$$

$$G_{En} = \frac{\sum_{k=1}^{n_{PEV,D}} P_k\left(t_1'\right)\left\{t_{\sup,k} - \frac{T\max}{\eta_i t_{\sup,k}}\left[1 - \exp\left(-\gamma_{dis,k}\eta_i t_{\sup,k}/T\max\right)\right]\right\}}{N_{st}P\max} \quad (36)$$

The renewable energy presence generally comes through arrays of photovoltaic panels (PV), wind farms. However, due to the irregularities in energy delivery recorded by renewables, the aggregators strongly recommend the access to a power

system containing stable power sources from thermal and nuclear power plants, hydroelectric plants and so on. In presence of renewables the net power is (37).

$$P_{BPEV,net} = \sum_{k=1}^{n_{PEV,D}} P_k\left(t_1^{'}\right)\left[1 - \exp\left(-\gamma_{dis,k}\eta_i t_{\sup,k}/T\max\right)\right] - P_{BPEV,\max}\left[n_{PEV,C} - \sum_{i=1}^{n_{PEV,C}} \exp(-\alpha_i \eta_C t_{chari}/T\max)\right] + P_{renewables} \quad (37)$$

5 Case Study: Simulations, Results and Discussion

In the case of individual vehicles, simulations targeted the battery power status of two different plug in electric vehicles during the power demand (charging), respectively power supply to the grid (discharging). The hypothetical PEVs under scrutiny have the characteristics of interest displayed in Table 4.

This assessment procedure applied to the PEVs battery charging and discharging processes relies on a methodology derived from the one proposed in [31] and continued in [11]. At the first step of simulations, one subjected the batteries of both PEVs to a full charge from 0 to 100% power. The charge efficiency values (i.e., operating in this case as rectifier), are: 50, 70, 90% and the ideal of 100%. The advantage of a faster charging process at high efficiency comes with the drawback of increased power demand from the supply, according to Fig. 5. For the same efficiency value, the battery charging performance, according to (15) and (19) is an intrinsic characteristic of an individual PEV.

In a similar manner, simulations regarding the PEV battery discharging during the time allocated for energy supply to the grid, regards the higher converter efficiency (i.e., inverter operation of the charger) due a faster discharge process (see Figs. 6 and 7), according to (29), which means a faster response to the grid demand. Although, faster charging/discharging processes, even benefiting the PEV operation, respectively the grid may result in severe shortening of the battery lifespan.

Table 4 Characteristics of plug-in electrical vehicles under study

Characteristic	PEV1	PEV2
Battery power at full charge (kW)	150	75
Maximum charge duration T_{max} (h)	4	8
Charging coefficient α_{C1}	9.75	8.25
Charging coefficient α_{C2}	6.75	4.75
Discharging coefficient γ_{dis1}	10.15	8.75
Discharging coefficient γ_{dis2}	5.00	4.0

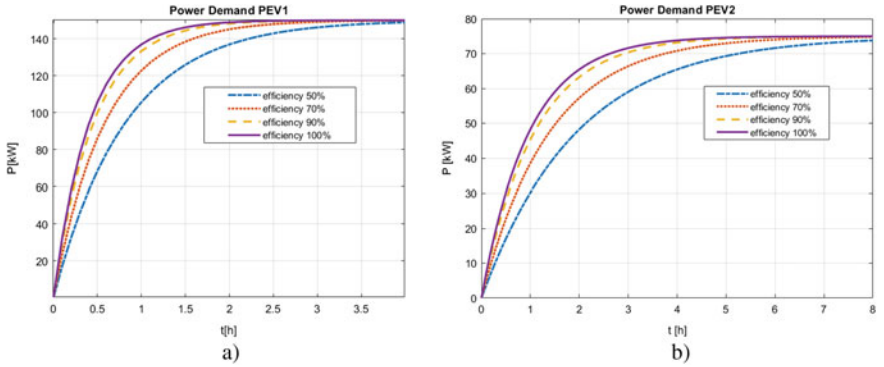


Fig. 5 Time dependency of the power demand for PEV1 (**a**) and PEV2 (**b**) influenced by the battery charger efficiency

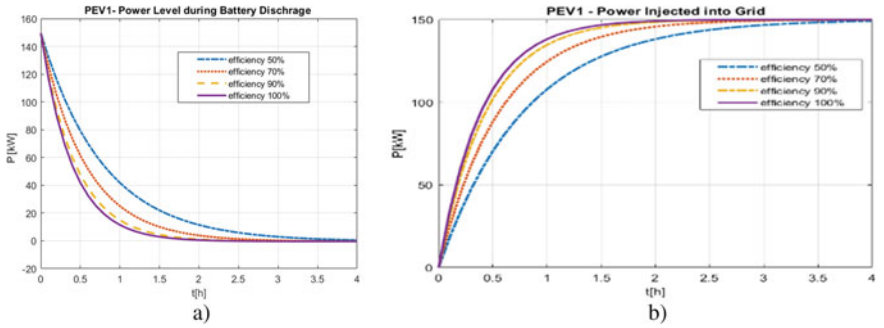


Fig. 6 Time dependency of the power discharge (**a**) and power injected into the grid (**b**) for PEV1 influenced by the inverter efficiency

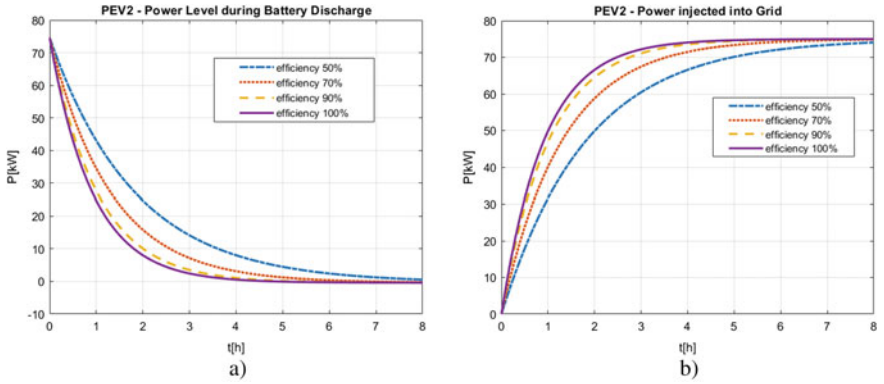


Fig. 7 Time dependency of the power discharge (**a**) and power injected into the grid (**b**) for PEV2 influenced by the inverter efficiency percent value

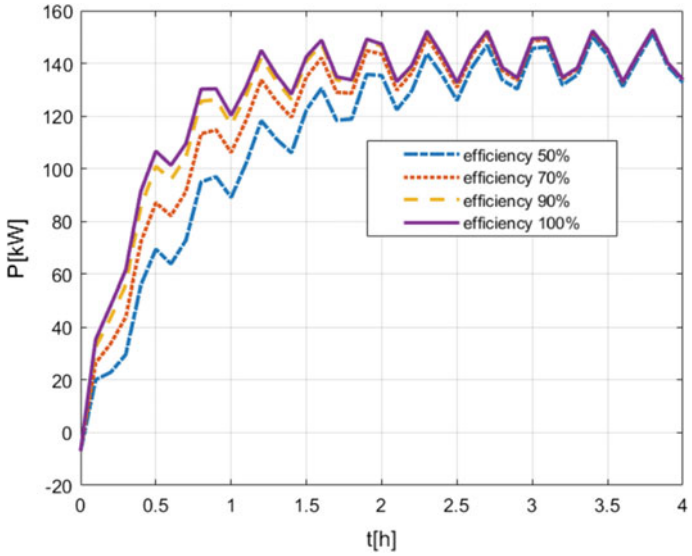


Fig. 8 Time dependency of the power demand for PEV1 in presence of renewable energy sources, for a battery charging coefficient $\alpha_C = 9.75$

Conisdering the presence of the renewable energy sources, the power demand, reflected at the level of individual PEVs appears in Figs. 8, 9, 10 and 11. A time alternating function of approximatively 2 cycles per hour mimics the presence of wind generators. In such a scenario, the SGPS doesn't have a connection to storage units like stationary battery arrays or main power system which includes classical (i.e. fossil fuel based) energy sources. The presence of renewables leads to approximate power ripples of 15% for PEV1, respectively 30% for PEV2. One can observe the fact that charging coefficients play a role as well: the higher value of the charging coefficient, the faster the process is (Figs. 9 and 10).

Taming the intermittency and randomness from the existing power-time profile of renewables (see Figs. 12 and 13), requires the presence of fast compensation function at the point of common coupling between the PEV, SGPs, stationary power storage batteries and eventually main power supply.

In a similar assumption of present renewable energy sources, the PEVs deliver power to the grid, having a profile, reflected at the level of individual PEVs appearing in Figs. 14, 15, 16 and 17. Renewable sources are identical to the ones present in the simulations regarding individual battery charging. The presence of renewables leads to approximate power ripples of 15% for PEV1, respectively 30% for PEV2, as well. One can observe the fact that discharging coefficients play a role as well: the higher value of the discharging coefficient, the faster the process is (Figs. 14 and 16).

Charging process regarding a parking lot requires scenarios in which the involvement of multiple vehicles requires certain assumptions made regarding the initial level of battery charging, respectively the durations necessary to reach certain levels

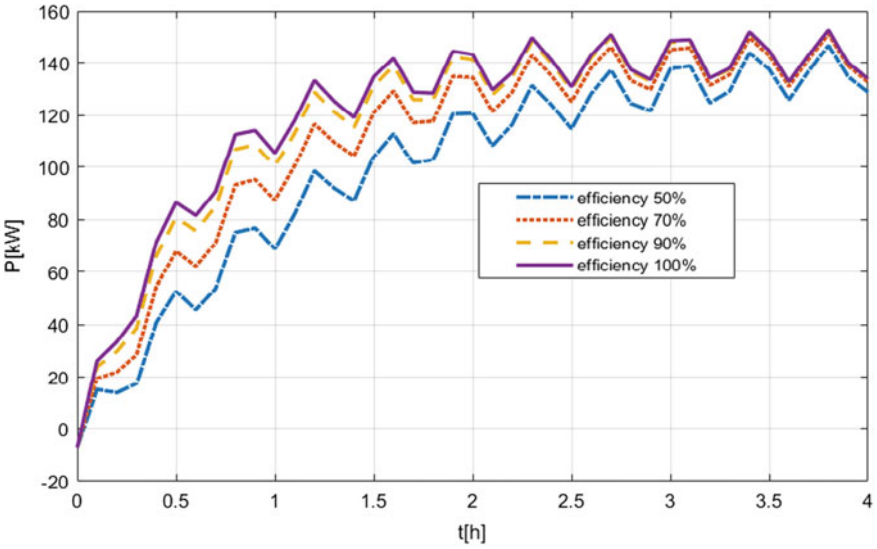


Fig. 9 Time dependency of the power demand for PEV1 in presence of renewable energy sources, for a battery charging coefficient $\alpha_C = 6.75$

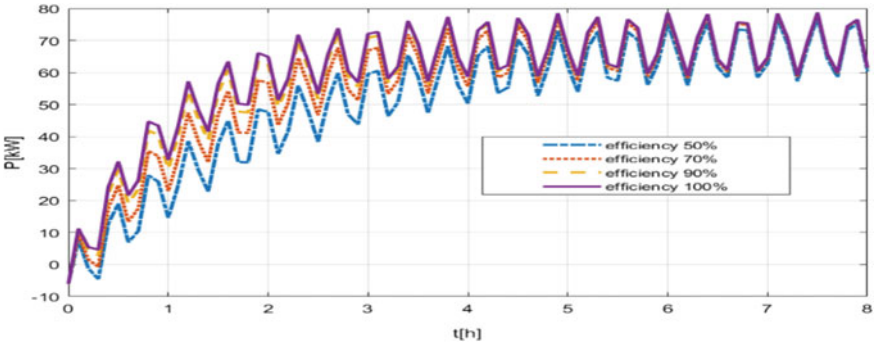


Fig. 10 Time dependency of the power demand for PEV2 in presence of renewable energy sources, for a battery charging coefficient $\alpha_C = 8.25$

of battery charging. For the multiple vehicles charging scenario, we've assumed the presence of 30 PEVs: 15 PEV1s and 15PEV2s. The efficiency of the charger, operating as rectifier is 95%, whereas the upper limit of the battery power charging is 90% of the maximum power for both types of PEVs using Level 2 chargers:

- 5 PEV1s at 0% initial power level before starting charging, 6 PEV1s at 20% initial power level, respectively 4 at 40% power level.
- 8 PEV2s at 0% initial power level before starting charging, 5 PEV1s at 30% initial power level, respectively 2 at 50% power level.

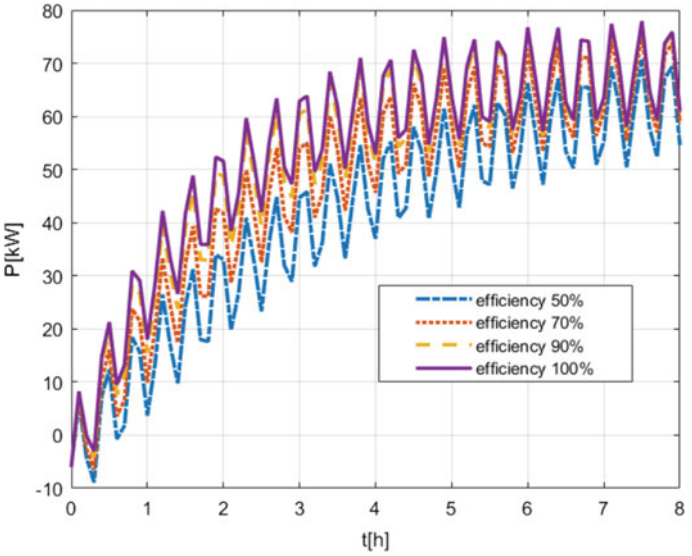


Fig. 11 Time dependency of the power demand for PEV2 in presence of renewable energy sources, for a battery charging coefficient $\alpha_C = 4.75$

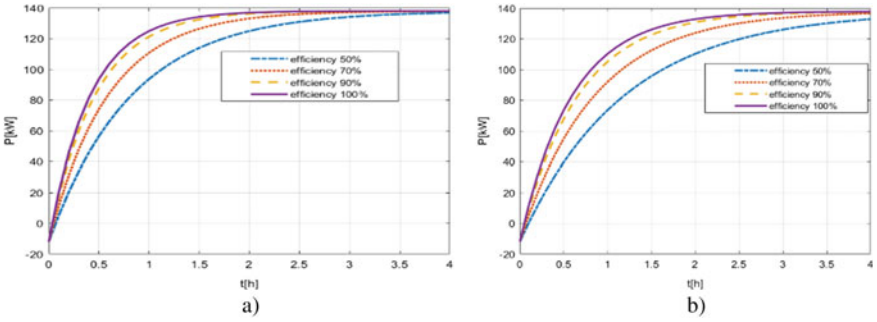


Fig. 12 Time dependency of the power demand for PEV1 in presence of compensated renewable energy sources, for a battery charging efficiency $\alpha_C = 9.75$ (**a**), respectively $\alpha_C = 6.75$ (**b**)

- For PEV1 is necessary 3/4 of the maximum charging time to reach 90% level from 0% initial power level, 1/2 of the maximum charging time to reach 90% level from 20% initial power level, respectively 1/4 of the maximum charging time to reach 90% level from 40% initial power level.
- For PEV2 is necessary 4/5 of the maximum charging time to reach 90% level from 0% initial power level, 1/2 of the maximum charging time to reach 90% level from 30% initial power level, respectively 1/4 of the maximum charging time to reach 90% level from 50% initial power level.
- The maximum charging durations of PEV1 and PEV2 appear in Table 4.

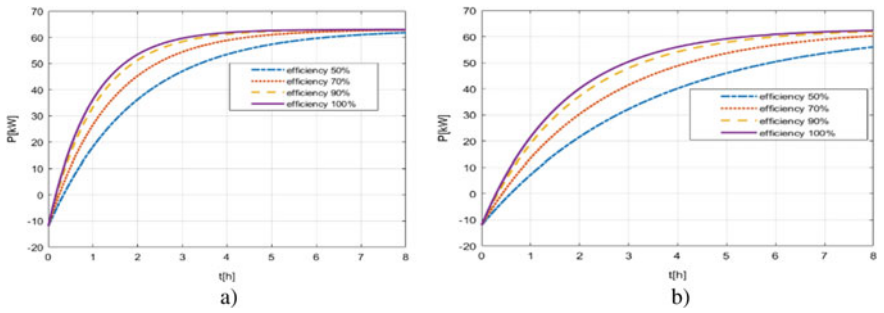


Fig. 13 Time dependency of the power demand for PEV2 in presence of renewable energy sources, for a battery charging efficiency $\alpha_C = 8.25$ (**a**), respectively $\alpha_C = 4.75$ (**b**)

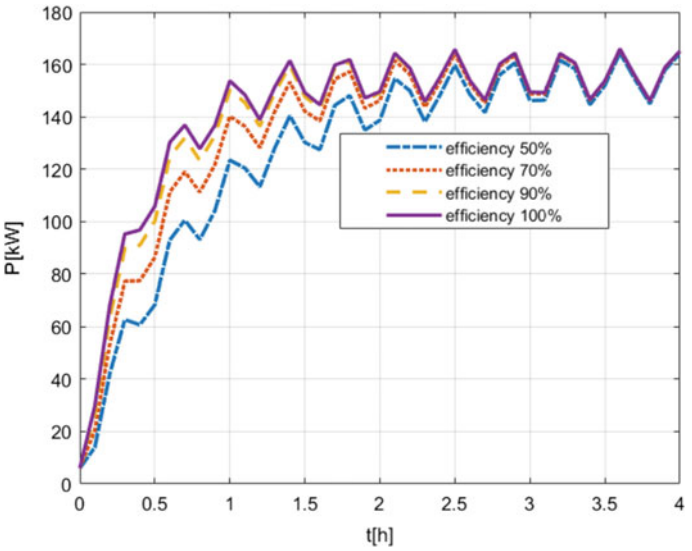


Fig. 14 Time dependency of the power injected into the grid from PEV1 in presence of renewable energy sources, for a battery discharging coefficient $\gamma_{dis} = 10.15$

- The application of (26) in the circumstances from above resulted in the following values of the charging coefficients: for PEV1 $\alpha_{C1,\,PEV1} = 3.2317$, $\alpha_{C2,\,PEV1} = 2.5347$, $\alpha_{C3,\,PEV1} = 3.29185$, $\alpha_{C1,\,PEV2} = 3.0297$, $\alpha_{C1,\,PEV2} = 1.929$, $\alpha_{C1,\,PEV1} = 2.1508$.

Following simulations involving charging the batteries of multiple vehicles from the same parking lot, without (see Fig. 18) and in presence of renewables (see Fig. 19), one can extract the power–time profile for charging, with a variation within 8 MW during the considered duration. Multiple scenarios can result in significant data, suitable for interpretation, stochastic calculations and finally in optimization functions.

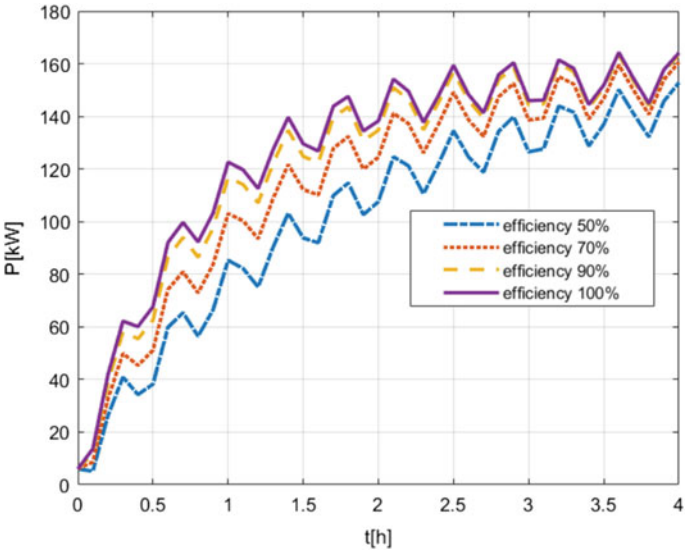


Fig. 15 Time dependency of the power injected into the grid from PEV1 in presence of renewable energy sources, for a battery charging coefficient $\gamma_{dis} = 5.0$

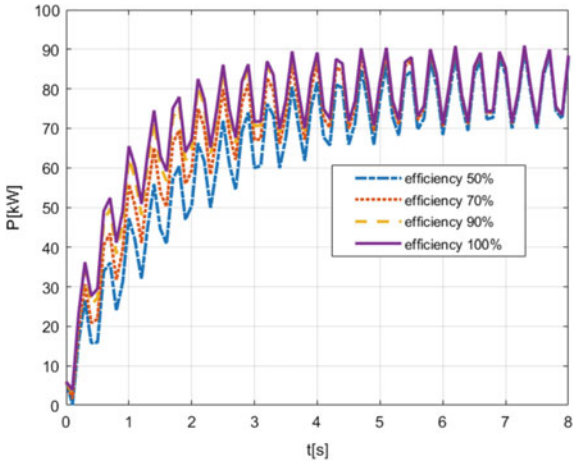


Fig. 16 Time dependency of the power injected into the grid from PEV2 in presence of renewable energy sources, for a battery charging coefficient $\gamma_{dis} = 8.75$

Although most of optimization functions cover domains like driving scheduling, emissions mitigation and revenues/taxes, the input from PEVs battery charging process characteristics, the interaction with SGPS, with and/or without compensated renewables proven very useful when commissioning elements of SGPS.

Power supply to the grid from a compound of vehicles from a parking lot, during the V2G operation, relies on discharging the batteries of these vehicles. Discharging process is faster than charging one [11], whereas the discharging process must

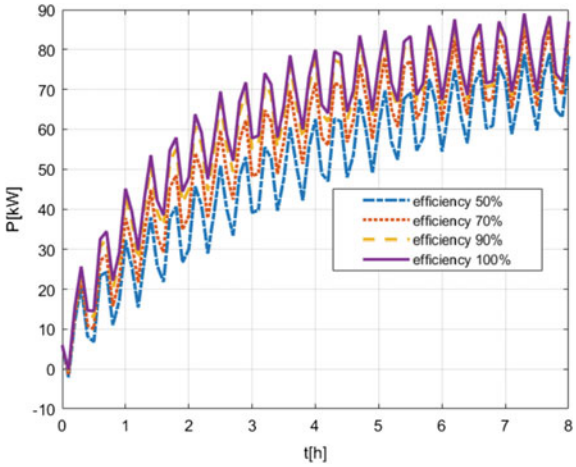


Fig. 17 Time dependency of the power injected into the grid from PEV2 in presence of renewable energy sources, for a battery charging coefficient $\gamma_{dis} = 4.0$

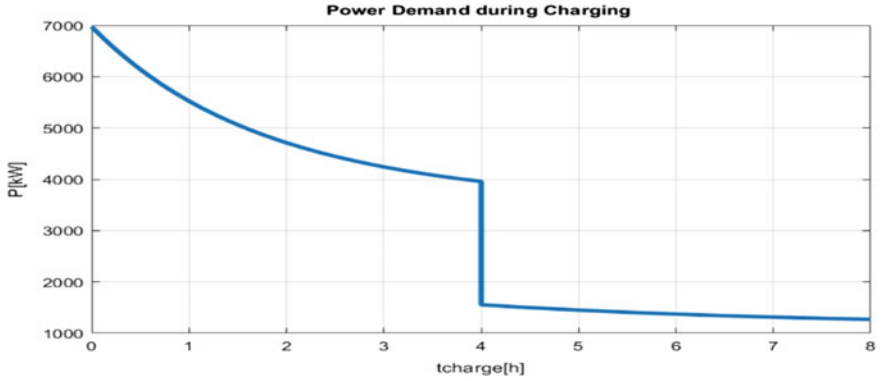


Fig. 18 Time dependency of the power demanded from the grid by the compounded PEV1 and PEV2 consumers from the parking lot, without involving renewable energy

encounter the low power threshold limit. Such a limit largely varies with respect to the type of vehicle and the next outcome of the PEV after providing energy to the SGPS during the power demand intervals (i.e., whether would live the parking lot for a drive or will stay for charging when available and for how long, and so for).

In this simulation regarding the battery discharging there were considered 10 PEV1s and 10 PEV2s, retrieving energy into SGPS through the same Level 2 chargers working as inverters.

- All 10 PEV1s and all 10PEV2s are fully charged when starting to supply the SGPS.

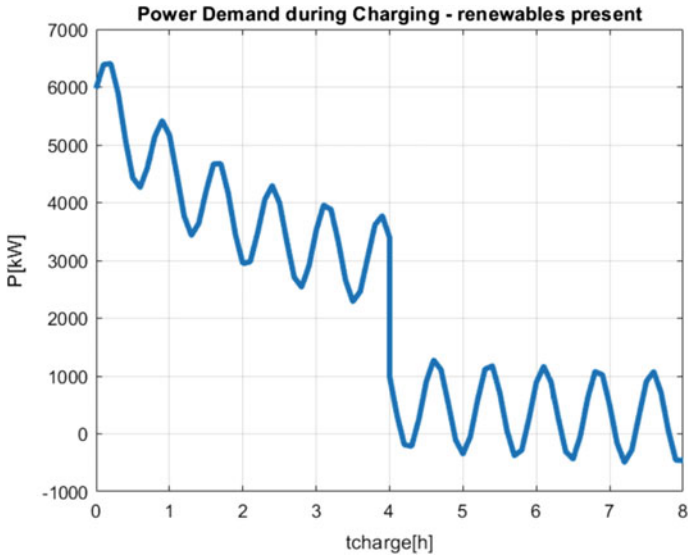


Fig. 19 Time dependency of the power demanded from the grid by the compounded PEV1 and PEV2 consumers from the parking lot, involving the presence of the fluctuating renewable energy sources

- The power level of the PEV1s battery drops from 100 to 5% (95% discharge) in 1/3 of its maximum charging time (see Table 4).
- The power level of the PEV2s battery drops from 100 to 10% (90% discharge) in 1/4 of its maximum charging time (see Table 4).
- The low power limit of the PEV1s is 4 kW, whereas for PEV2s is 2 kW.
- The application of (29) in the frame of the assumptions from above, results into the following discharging coefficients: $\gamma_{\text{dis,PEV1}} = 8.11$, respectively $\gamma_{\text{dis,PEV2}} = 8.70$.
- Because the battery discharging process is much faster than the charging one, we've considered four identical batches of vehicles performing V2G for a total duration equal to the maximum charging time of PEV2.

For this case of simulation, the intrinsic variability of the nature of the V2G model (Fig. 20) covers about 2 MW and has a cyclical nature. In a parking lot of vehicles demanding power and those who supply the SGPS coexist. To simulate such a regime, one superimposed the previous two scenarios, charging and discharging. The battery discharging is much faster, fact that explains the power-time profile is modulated by the profile of power demand, whereas presents a ripple following the cyclical character of the power-time profile recorded for V2G operation. Following the alternative character of the renewables of "wind generation" type, one can obtain a net power–time profile in presence of renewables (37), according to Fig. 21, or after compensating the renewables intermittency (33), according to Fig. 22.

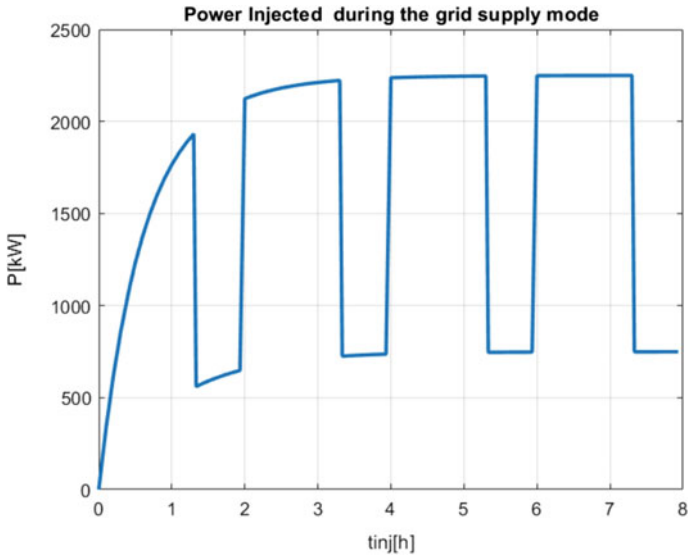


Fig. 20 Time dependency of the power injected into the grid by the compounded PEV1 and PEV2 suppliers from the parking lot, without involving renewable energy

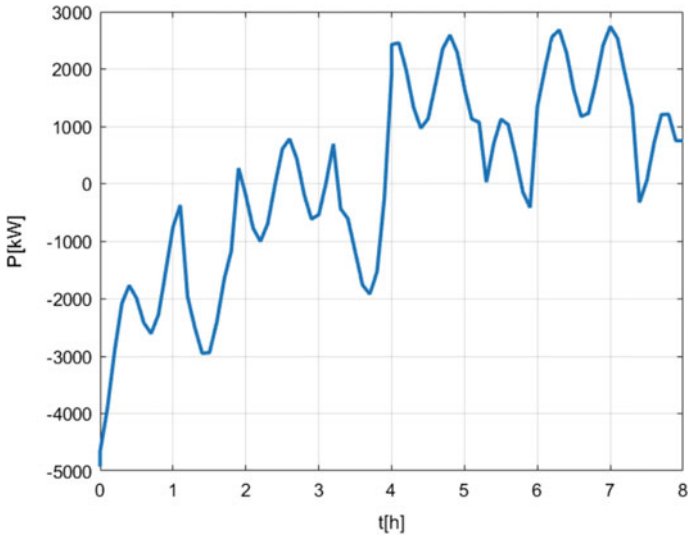


Fig. 21 Time dependency of the net power exchanged with the grid by the compounded PEV1 and PEV2 suppliers from the parking lot, involving the presence of fluctuating renewable energy

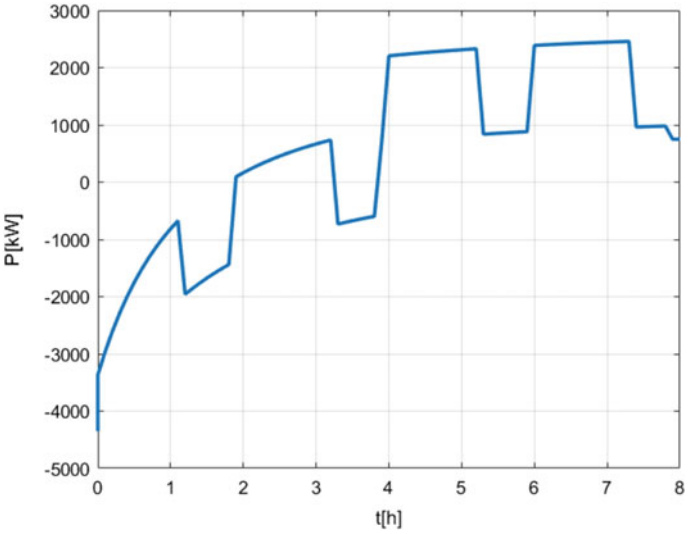


Fig. 22 Time dependency of the net power exchanged with the grid by the compounded PEV1 and PEV2 suppliers from the parking lot, involving the presence of compensated renewable energy

Although is highly unlikely to have a positive net power for long durations, the impact of the PEVs operation in V2G is beneficial for diminishing the overall energy consumption. The gain factors (35) and (36) became objective functions which can lead to optimization of parking lot power flow as function of time and provide information as input for two directions: the PEVs traffic schedules and revenue increase. Each of this direction operates in terms of own family of optimization functions, which can be minimized/maximized for optimal results [31].

Another study involves a system in a configuration similar the one from Fig. 1. The example originates from an application presented on the *MATLAB/Simulink portal* [32], treating a V2G system performing one of its ancillary duties: frequency regulation on a relatively small SGPS. All the SGPS components operate at a three-phase power supply rated for 25 kV. The connection to the SGPS for both, the wind farm and the photovoltaic required the presence of a standard step-up transformer following the output inverters, considered sinusoidal. There is no reference to harmonics in this section. A step-down transformer brings the voltage level to the EV, respectively, residential consumers. For simplification reasons, one can ignore the necessity of another intermediate step-down transformer to the residential consumers. The V2G participants are:

(a) One fossil-fuel-based (Diesel) generator with the nameplate data: $Sn = 30\ MW$, $V_n = 25\ kV$, $f_n = 60\ Hz$, driven by a 30 MW Diesel Engine
(b) One Wind Farm rated for: $SWF = 6\ MW$, *nominal wind speed 12.5 m/s, maximum wind speed 15 m/s*

(c) One Photovoltaic Farm rated for $SPV = 9\ MW$, *efficiency 11%, radiant surface 70,000* m^2
(d) Three Phase transformer rated for $S_T = 50\ MVA$, $V_1/V_2 = 25/0.46\ kV$
(e) The community has 1200 homes, with a ratio of 1:10 eV owners/(total households). The load contains a residential type of consumers as well as a low power induction machine, has a consumption profile and a given power factor. The induction machine, as the largest single consumer has the following data: $P_n = 200$ kW, $V_n = 600$ V, $f_n = 6$ Hz, $R_S = 0.0278$ pu, X_{lS} $(L_{lS}) = 0.1091$ pu, X_m $(L_m) = 2.974$ pu, $R_r = 0.0328$ pu, $X_{lr}(L_{lr}) = 0.1997$ pu.
(f) 120 New EVs, with a nominal power $PEV = 150\ kW$ each, and a charger efficiency of 95%.

The EVs operate under the V2G technology, and their batteries are in the charging mode when the power from SGPS is in excess, respectively in discharging mode when there is a deficit of power from SGPS. The interval under study comprised a full 24 h interval, whereas the profiles proposed in [32] by MathWorks and adopted here include all possible states for an EV. The profiles are:

- *Profile #1(42 cars)*: Going to work and charge the battery car there
- *Profile #2(25 cars)*: Going to work and charge the battery car after a more extended trip
- *Profile #3(15 cars)*: Going to work and charge the battery car after the arrival at home
- *Profile #4(25 cars)*: Stay at home
- *Profile #5(12 cars)*: Working night shift

The simulation covers 24 h. The solar radiance follows a normal distribution function with the maximum in the middle of the day, with partial shading at noon, whereas the wind has a very irregular profile, respectively one trip due to excessive speed at 10.00 pm. After running the simulation, the power distribution functions appear in Figs. 5, 6, 7 and 8. The shape of power indicates the effectiveness of the voltage regulation process (Figs. 23, 24 and 25).

The renewable energy sources have a cumulated installed power of 15 MW, equal to 50% of the nominal power of the diesel generator. The total power of the PEVs is

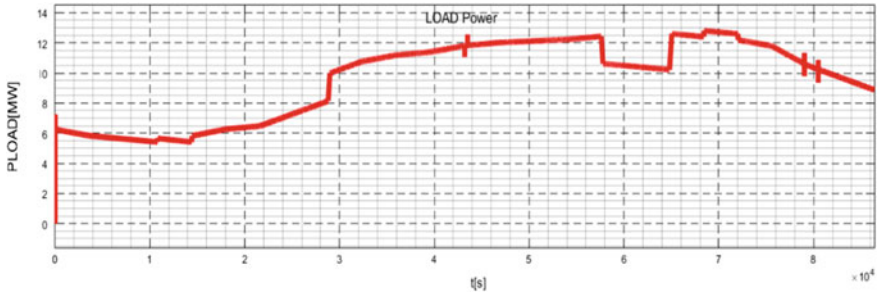


Fig. 23 Real Power at the load during an interval of 24 h (expressed in seconds)

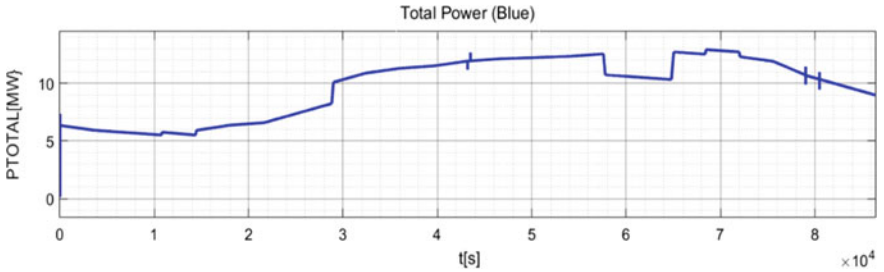


Fig. 24 Total Real Power during an interval of 24 h (expressed in seconds)

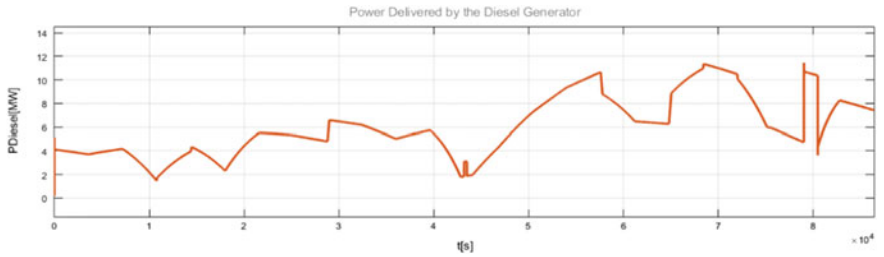


Fig. 25 Total real power delivered by the diesel generator during an interval of 24 h (expressed in seconds)

equal to 18 MW. However, the renewable energy harvested from both solar panels and wing generators is intermittent (see Fig. 26), whereas the nature of interaction between the PEVs and SGPS follows the profiles 1 through 5. All the consumers and generators determine a load power–time profile presented in (24), with all groups having an important participation to the total power, with respect to the principal power source, the Diesel generator.

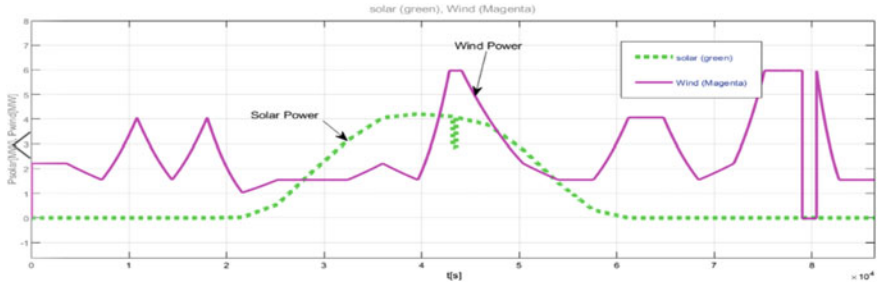


Fig. 26 Total real power delivered by the solar panels array and wind generators during an interval of 24 h (expressed in seconds)

6 Conclusion

The EVs, due to the V2G capability of bi-directional power transfer, can play the role of the consumer as one of the producers as well, capable of smoothening the power irregularities from wind and solar power sources. Besides the advantage of being the energy beneficiary, and the merit of being present when the SGPS needs, the EV battery plays the ancillary roles of power regulation, spinning reserves, and compensator for reactive power. Challenges faced by SGPS due to the presence of the power electronic converters involved in the EV charging/discharging process accompany the drawback of the life span shortening suffered by the EV batteries. The models presented in this chapter address the charging and discharging processes subjecting individual and multiple PEVs, through simulations, with and without the impact of the intermittent character of renewables. The primary outcome of the simulation regards the power–time profile in conditions of battery charging, supply to the grid, and mixed. The impact of the intermittent renewable sources reflects an increased power variability, indicating the need for compensation using stationary storage systems or accepting classical, fossil fuel-based energy sources at the point of common coupling. The latter are stable sources which smoothen the overall power–time profile. The net power determined through simulations becomes a candidate for "objective function" status, and its maximization becomes possible. The power or energy gain factors, derived from the net energy appear suitable as objective functions, and both include revenue elements. However, the construction procedure targeting all these objective functions must include stochastic algorithms for accuracy purposes. The last simulation, built-up derived from a MATLAB/Simulink existing application and could offer information about the electric power profile over 24 h and in different points of the integrated system. There are several treatments for the shortcomings regarding the EVs integration, and optimal scheduling for driving and charging are amongst them. However, the main reasons which make the V2G impractical for now are the reduced number of vehicles, unattractive compensation for the participating EV owners as well as the anxiety of losing the driving range whereas the EVs participate in V2G joint venture. However, many models indicated profitability for the distributor and aggregator. There are two distinct groups of standards following the subject of application: EV and SGPS, both addressing similar topics, yet from different referential.

References

1. Kempton W, Letendre SE (1997) Electric vehicles as power source for utilities. Transp Res D 2(3):157–175
2. Liu L, Kong F, Liu X, Peng Y, Wang Q (2015) A review on electric vehicles interacting with renewable energy in smart grid. Renew Sustain Energy Rev 51:648–661
3. Shaukat N, Khan B, Mehmood CA, Khan J, Farid U, Majid M, Anwar SM, Jawad M, Ullah Z (2018) A survey on electric vehicle transportation within smart grid system. Renew Sustain Energy Rev 81:1329–1349

4. Tan KM, Ramachandaramurthy VK, Yong JY (2016) Integration of electric vehicles in smart grid: a review on vehicle to grid technologies and optimization techniques. Renew Sustain Energy Rev 53:720–732
5. Francis Mwasilu F, Justo JJ, Kim EK, Do TD, Jung JW (2014) Electric vehicles and smart grid interaction: a review on vehicle to grid and renewable energy sources integration. Renew Sustain Energy Rev 34:501–516
6. Yang Z, Li K, Foley A, Zhang C (2014) Optimal scheduling methods to integrate plug-in electric vehicles with the power system: a review. Preprints of the 19th world congress the international federation of automatic control Cape Town, South Africa, August 24–29, 2014, pp 8595–8603
7. Xydas E, Marmaras C, Cipcigan Liana M, Jenkins N, Carroll S, Barker M (2016) A data-driven approach for characterising the charging demand of electric vehicles: a UK case study. Appl Energy 162:763–771
8. Green RC II, Wang L, Alam M (2011) The impact of plug-in hybrid electric vehicles on distribution networks: a review and outlook. Renew Sustain Energy Rev 15:544–553
9. Rezvani Z, Jansson J, Bodin J (2015) Advances in consumer electric vehicle adoption research: a review and research agenda. Transp Res Part D 34:122–136
10. Krupa JS, Rizzo Donna M, Eppstein Margaret J, Lanute BD, Gaalema Diann E, Lakkaraju K, Warrender Christina E (2014) Analysis of a consumer survey on plug-in hybrid electric vehicles. Transp Res Part A 64:14–31
11. Reddy GH, Goswami AK, Choudhury NBD (2018) Impact of plug-in electric vehicles and distributed generation on reliability of distribution systems. Int J Eng Sci Technol 21:50–59
12. Anastasiadis AG, Kondylis GP, Pilizakis A, Vokas G (2019) Effects of increased electric vehicles into a distribution network. Energy Procedia 157:586–593
13. Qian K, Zhou C, Yuan Y (2015) Impacts of high penetration level of fully electric vehicles charging loads on the thermal ageing of power transformers. Electr Power Energy Syst 65:102–112
14. Godina R, Rodrigues EMG, Paterakis NG, Erdinc Catalão JPS (2016) Impacts of high penetration level of fully electric vehicles charging loads on the thermal ageing of power transformers. Energy Convers Manag 120:206–216
15. Geske J, Schumann D (2018) Willing to participate in vehicle-to-grid (V2G)? Why not! Energy Policy 120:392–401
16. Rebecca Gough, Dickerson C, Rowley P, Walsh C (2017) 1Vehicle-to-grid feasibility: a techno-economic analysis of EV-based energy storage. Appl Energy 92:12–23
17. Bhandari V, Sun K, Frances Homans (2018) The profitability of vehicle to grid for system participants—a case study from the Electricity Reliability Council of Texas. Energy 153:278–286
18. Darlene S (2017) Critical elements of vehicle-to-grid (V2G) economics. Strategic partnership project Report NREL/TP-5400-69017 September 2017. National Renewable Energy Laboratory 15013 Denver West Parkway Golden, CO 80401 303-275-3000. www.nrel.gov
19. Zakariazadeh A, Jadid S, Siano P (2015) Integrated operation of electric vehicles and renewable generation in a smart distribution system. Energy Convers Manag 89:99–110
20. Michela Longo, Federica Foiadelli, Wahiba Yaïci (2019) Electric vehicles integrated with renewable energy sources for sustainable mobility book: new trends in electrical vehicle powertrains. IntechOpen 10:203–223
21. Fazelpour F, Vafaeipour M, Omid Rahbari O, Rosen MA (2014) Intelligent optimization to integrate a plug-in hybrid electric vehicle smart parking lot with renewable energy resources and enhance grid characteristics. Energy Convers Manag 77:250–261
22. Rahbari O, Vafaeipour M, Omar N, Marc A, Rosen MA, Hegazy O, Timmermans JM, Heibati S, Peter Van Den Bossche P (2017) An optimal versatile control approach for plug-in electric vehicles to integrate renewable energy sources and smart grids. Energy 134:1053–1067
23. Honarmand M, Zakariazadeh A, Shahram Jadid S (2014) Optimal scheduling of electric vehicles in an intelligent parking lot considering vehicle-to-grid concept and battery condition. Energy 65:572–579

24. Wang Z, Wang L, Dounis AI, Yanga R (2012) Integration of plug-in hybrid electric vehicles into energy and comfort management for smart building. Energy Build 47:260–266
25. Battistelli C, Baringo L, Conejo AJ (2012) Optimal energy management of small electric energy systems including V2G facilities and renewable energy sources. Electr Power Syst Res 92:50–59
26. Haddadian G, Khalili N, Khodayar M, Shahidehpour M (2016) Optimal coordination of variable renewable resources and electric vehicles as distributed storage for energy sustainability. Sustain Energy Grids Netw 6:14–24
27. Uddin K, Dubarry M, Glick MB (2018) The viability of vehicle-to-grid operations from a battery technology and policy perspective. Energy Policy 113:342–347
28. Peng C, Zou J, Lian L, Li L (2017) An optimal dispatching strategy for V2G aggregator participating in supplementary frequency regulation considering EV driving demand and aggregator's benefits. Appl Energy 190:591–599
29. Bishop JDK, Axon CJ, Bonilla D, Banister D (2016) Estimating the grid payments necessary to compensate additional costs to prospective electric vehicle owners who provide vehicle-to-grid ancillary services. Energy 94:715–727
30. Richardson DB (2013) Encouraging vehicle-to-grid (V2G) participation through premium tariff rates. J Power Sources 243:219–224
31. Chukwu UC, Mahajan SM (2014) V2G Parking lot with RV rooftop for capacity enhancement of a distribution system. IEEE Trans Sustain Energy 5(1):119–127
32. www.mathworks.com

Numerical Approaches of Biomass Plants Efficiency

Emil Diaconu, Alexandru Enescu, Horia Andrei, and Sorin Deleanu

Abstract Considering the ongoing process of diversification of the types of energy production, especially those from renewable sources, one of the goals of sustainable development has become the cleanest energy production, and the monitoring and modeling of the operating parameters is a topic of great interest for researchers in the field. Thus it became an imperative condition to promote the production of electricity from renewable sources in order to support the environmental protection movement but also to obtain an energy independence. The cogeneration process is one of the solutions for obtaining energy from renewable sources because it efficiently energizes the production system by obtaining thermal and electrical energy using the same quantity of fuel. Given that biomass is the primary energy source, it is clear that it is a clean energy source. This chapter presents a factory for the production of electricity and heat in Romania, which was put into operation in 2014. Its analysis will be done by mathematical modeling of the energy consumption (biomass and biogas) necessary for the operation. Also, the most important of a biomass-based power plant, precisely the electrical output of energy, the technological usage, and the electrical energy supplied to the national distribution grid are analyzed. The result consists of and an approximation model linking factors like the biogas consumption to the electrical energy output. In Sects. 2, 3 and 4 of this chapter, the authors

E. Diaconu (✉)
Faculty of Electrical Engineering, Electronics and Information Technology, University Valahia, Targoviste, Romania
e-mail: emy_diaconu@yahoo.com; emil.diaconu@valahia.ro

A. Enescu
Doctoral School of Engineering Sciences, University Valahia, Targoviste, Romania
e-mail: enescu_alex17@yahoo.com

H. Andrei
Department Doctoral School, University Valahia, Targoviste, Romania
e-mail: hr_andrei@yahoo.com

S. Deleanu
School of Applied Sciences and Technology, Northern Alberta Institute of Technology (NAIT), Edmonton, Canada
e-mail: sorind@nait.ca

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_23

describe the structure of a cogeneration power plant (CPP) and introduce an example through a case of study of the data acquisition system for the CPP's parameters and propose some numerical approaches for the evaluation of the CPP efficiency. The CPP analyzed in this chapter is in Suceava County, Romania, and when it first started production, in 2014, it was the biggest running on biomass cogeneration plant in the country. The chapter includes a description of the three distinct phases of the technological process of obtaining energy from biogas, starting with the type and quantities of the raw materials used and how much energy the CPP can produce. The data acquisition system is part of a sophisticated automated system called "DIANE," which permanently monitors, coordinates, and controls all the operations in the cogeneration power plant. Following the measurement of many parameters, the analysis focuses on electrical energy production due to each generator, biogas consumption of motors, domestic electricity consumption, and the power consumption required to operate the biogas station. The acquisition of the operating parameters from the last three years continued with the application of a numerical method of interpolation, based on the PYTHON software environment. Consequently, the authors obtained the relationship between the electrical energy output and the consumed biomass input parameters in the form of polynomial functions. The chapter ends with conclusions and many references on the topic of numerical approaches to the biomass plant technological process and overall efficiency.

Keywords Cogeneration power plant · Biomass · Biogas · Cogeneration engine · Efficiency · Numerical approaches model

Nomenclatures

Acronyms

CPP	cogeneration power plant
NPV	Net present value
IRR	Internal rate of return
PBP	Payback period
E	Engine
NPS	National Power System

1 Introduction

The commissioning of the first commercial electrical power plant in 1882, owned by Thomas Edison, also meant the emergence of the first cogeneration plant [1], because this facility supplied electricity to the grid but also used the secondary thermal energy

produced to deliver steam to the local businesses. However, it was not until 1978 when the United States of America Congress recognized the efficiency of the cogeneration plants and started encouraging utility companies to buy energy from such producers. Years later, The European Union has incorporated cogeneration into its energy policy through the Cogeneration Directive 2004/08/EC [2], a directive that aims to support cogeneration and establishes a method for calculating cogeneration capacities for each country. During and after Hurricane Sandy, cogeneration systems have played a key role in enabling hospitals, universities, schools and residential buildings to continue operations when the power grid in the worst-affected areas has fallen—proving that cogeneration plants are a viable and efficient choice when we want to make our energy infrastructure more resilient to extreme weather events. To highlight the role of the cogeneration systems, we can take into consideration the following example: every year, in the USA, energy is lost in the conventional power plants in the form of waste heat. The losses are higher than Japan's total energy consumption. The principle of cogeneration reduces this amount of wasted energy by almost half. A cogeneration plant generates electricity and heat at the same time: the residual heat produced at the generation of electricity finds an efficient utilization in heating systems and those for the preparation of household water. In this way, primary energy is used efficiently based on the principle of cogeneration.

Energy efficiency and global efficiency increase—in many cases, even over 100%. Therefore, a cogeneration plant can save 50% or more primary energy compared to a conventional heating solution without electricity generation. This fact also reduces CO_2 emissions and contributes to environmental protection. Many researchers approached the cogeneration subject and done many studies on it: Ramadan Hefny Ali et al. [3], Eduardo Jose Cidade et al. [4], Phairat Usubharatana and Harnpon Phungrassami [5], Jorge Islas et al. [6] or Joao Cardoso, Valter Silva and Daniela Eusebio [7] who presented an economic model based on NPV, IRR and PBP developed over the 25 years lifetime of the analyzed powerplant.

The objective of this chapter is to obtain mathematical models of polynomial type with high accuracy of the dependency between the input (consumed biomass) and the output (electrical energy). So, to prevent the purchase and use of excess biomass, thus increasing the efficiency of the CPP, we have to determine an accurate mathematical relationship between the energy produced and the biomass consumption.

The chapter has the following structure: Sect. 2 presents the structure of the biomass plants. In Sect. 3 the data acquisition system of the cogeneration power plant is presented. A mathematical approach to the CPP efficiency is presented in Sect. 4, where we determined an accurate mathematical relationship between the energy produced and the biomass consumption. Conclusions are drawn in Sect. 5, followed at the end of the chapter by the list of references.

2 Structure of Biomass Plants

Nowadays, it became an imperative condition to promote the production of electricity from renewable sources in order to support the environmental protection movement but also to obtain an energy independence.

The first clear act of legislation was The Directive 2001/77/EC of the European Parliament and the European Council [8] regarding the promotion of electricity produced from renewable energy sources in the internal market.

This act represents the first definite step of the European Union to meet the obligations to reduce greenhouse gas emissions to which they hired by ratifying the Kyoto Protocol.

Romania was among the first candidate countries to the European Union, which transposed the provisions of Directive 2001/77/EC into its legislation (GD no. 443/2003 [9], as amended from GD 958/2005 [10]). Subsequently, the Law 220/2008 was enacted to establish the system for promoting the production of energy from renewable energy sources, with subsequent amendments and completions, which established the level of national targets for the share of electricity produced from renewable energy sources in the final gross consumption of energy, for the years 2015 and 2020 of 35%, respectively 38%.

In this regard, the Romanian Government has implemented the provisions of European Directive 2001/77/EC on promoting the production of electricity from renewable energy sources, adapting in this respect the Romanian Energy Policy. Thus, Law 220/2008 [11] was adopted with the subsequent amendments and completions, regarding the promotion of renewable energy sources by applying the support scheme by granting Green Certificates. This scheme supports the implementation of electricity generation projects based on renewable energy sources, granting Green Certificates to electricity producers depending on the technology used to produce electricity from renewable sources.

Given the above, a Romanian company proposed the realization of such a project—the biomass cogeneration plant with an installed power of 2,974 MW el. The biogas resulted from the fermentation of the biomass. Biogas is the fuel used by the two internal combustion thermal engines, which equip the plant.

The technological process of cogeneration (producing energy with biogas generated from mixed crops and manure) can be described in 3 distinct parts:

- acquisition, storage, and preparation of the specific raw material utilized in the process of biomass production: the primary energy source (corn and animal manure) is obtained through a long term contract (15 years) with local farmers;
- obtaining the biogas from anaerobic fermentation process—takes place in two primary fermenters and a secondary fermenter, and the biogas is stored in a tank with the capacity of 5000 Nm^3;
- production process for both electrical and thermal energy represented by the biogas combustion in the cogeneration units.

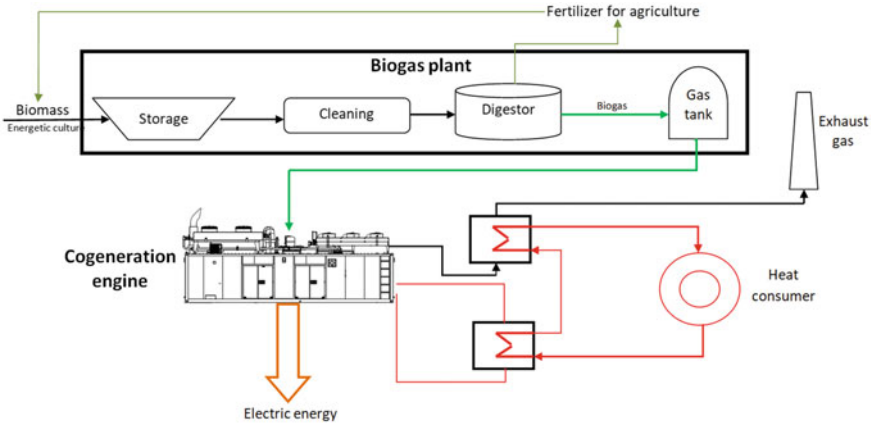


Fig. 1 Block diagram of the cogeneration plant

The biomass plant described and analyzed in this chapter is the largest biomass cogeneration plant put into operation in Romania [12]. The total investment was approximately 6.3 million Euros, financed from own funds combined with bank loans.

The technical solution adopted, for the production of electricity and heat in high-efficiency cogeneration from renewable sources, is based on the use of biogas obtained from the fermentation of biomass by two thermal engines of the JMS type 420 GS—B.LC having a power an installed electricity of 1,487 MW and an installed thermal power of 1,472 MW.

As shown in Fig. 1, the cogeneration power plant (CPP) structure consists of two main components: two high-efficiency cogeneration engines (modules) and the biogas plant.

High efficiency Cogeneration Engines (Modules)

This cogeneration plant has the advantage of producing thermal energy without additional costs.

The equipment from plant equipment includes two cogeneration modules, which utilize the biogas supplied by the fermentation of biomass (i.e., energy crops) as fuel has the advantage of producing thermal energy without additional costs. In this respect, CPP uses the residual energy from the production of electricity from the cooling circuits of the engines, thus producing both electrical and thermal energy simultaneously, fulfilling the conditions of high-efficiency cogeneration.

The technical features of the cogeneration modules appear highlighted in Table 1.

Table 1 Technical specification of cogeneration modules

Thermal engine type	JMS 420 GS-B.L v. B 25	
Electrical energy output	1415	kW
Recoverable thermal output (180° C)	1413	kW
Fuel consumption based on a LHV of 5 kWh/Nm^3	675	Nm^3/h
Electrical energy efficiency	41.9	%
Thermal energy efficiency	42.4	%
Total energy efficiency	84.3	%
Generator voltage	0.4	kV
Frequency	50	Hz
Fuel-gas pressure	120–180	mbar
Forward temperature	90	°C
Return temperature	70	°C
Engine noise level	97 dB at 1 m	

Redundancy required the choice of equipment configuration with two smaller cogeneration modules instead of a single larger one: to ensure the supply of electricity and heat in case of stopping an engine for repairs or scheduled maintenance.

The life expectancy for the engines is of 60,000 operating hours (~7–8 years), after that capital repairs become necessary. Following the repair, the engine's life expectancy doubles to 120,000 operating hours.

The biogas plant

The equipment components of the biogas station consist of the following essential components, which described below:

(a) The central part of the installation consists of two fermentation reactors. Inside, there are two fermenters installed, heated and thermally insulated each with a volume of approx. 3,393 m^3 and a post-fermentation basin with a capacity of approx. 3,393 m^3. Above the fermenters, a membrane roof is mounted, with one single foil for all 3 tanks. There is a total of nine electric mixers fitted, and in each fermenter/digester, with one solid fuel supply screw-mounted laterally.
(b) There are two basins, each with a volume of approx. 6,000 m^3.
(c) An earlier basin with a volume of approx. 49 m^3, covered with a rolling floor made of reinforced concrete, equipped with a pump for liquid animal waste (centrifugal pump).
(d) One separator.
(e) One collection basin for liquid waste from tanks (under construction).
(f) Two feeders of solid-fuel, each with a volume of approx. 65 m^3.
(g) Two buildings for technical equipment (made container type with technical equipment).
(h) One external gas storage tank.
(i) One gas burner.

(j) One gas cooler.

For supplying the biogas production station with solid fuels (corn silage and bovine manure), a solid fuel feeder exists with a capacity of approx.—65 m^3. The feeder intake comes twice a day by the company staff using a wheeled charger. The weighting cells, with on-site display, facilitate the optimum filling with substrate material, quantitatively and qualitatively.

The transport choke shape is as a tubular screw, submersible at the level of the substrate material in the fermenter. Such a solution ensures minimal loss of gas in the screw area, during the operation of the biogas plant operation. The control system of the cogeneration plant automatically takes over the feeding system, to ensure the optimum amount of substrate material permanently in the digester. The same amount of substrate as the fed one goes pumped through the distributor, reaching industrial building at the post-fermenter and then to the separator.

The substrate material from the basins comes extracted with the help of an evacuation station, transported by tankers and spread on agricultural surfaces.

To accelerate the production of biogas, control the dosage, and avoid the floating layers, the stirring of material takes place in both the digester and the post fermenter by using electric mixers at pre-established time intervals.

Both digesters and post-fermenters are gas-tight. The digesters and the post-fermenter are insulated and equipped with a stainless steel heating system, located inside, to control and obtain the desired temperature for the fermentation process. The cover of the digester tanks and the post-fermenter contains a membranous roof, resistant to the wind. The gas drives through the pipeline routes to ensure the mixing (and homogeneity) in the tanks. The membranous roofs and the external tank uniformed the oscillations in the production of gas and store the reserve for the cases of damage for a defined period. The biogas flows through a gas pipeline to the consuming units—the cogeneration modules. The digesters and the post-fermenter are heated and thermally insulated.

The biochemical fermentation process takes place in digesters and post-fermenters that are heated to approx. 40 °C. Different strains of bacteria gradually decompose the organic elements, and in the last phase of decomposition produce the biogas. The tanks are monolithic, made of reinforced concrete with a membrane roof. The temporary storage of resulting material in lagoons gets consideration for future use in agriculture. The captured residual biogas is then introduced into the gas network.

To obtain the required process temperature of 40 °C in digesters, thermal pipe circuits are mounted on the sidewall of the digesters and at the post-fermenter wall (V4A, DN 100). The heating of the tanks is done using the thermal energy in the form of the hot water produced in the cogeneration system by the thermal motors that constitute the cogeneration plant.

Transportation of collected liquids through pumping (water, water resulting from seepage into reservoirs.) in the retention tank on the spot continues with the transfer into a distribution node, in the connector structure at the top of the digesters.

The lighter organic mass relative to the water begins to decompose and sinks at the bottom. The unfermented material remains in the upper area of the digester. The heavy material arrives into the post-fermenter whereas pumped through a central station. There, the decomposition process begins again, with the difference that there are not fresh materials introduced in the post-fermenter. The same pumping process used in the digesters takes place in the post-fermenter.

From the separator, the collected material flows through the distributor in lagoons (basins). The decomposed material from lagoons, taken up with a pumping system is filling the tankers. The tankers ensure the transport of these fermentation wastes, used as agricultural fertilizer, into some agricultural regions. The decomposition proportion of the organic mass in digesters reaches 70–75%, to obtain the quality of the biogas required for the operations of the cogeneration plant under conditions of maximum efficiency. The total fermentation period is about 60 days. The necessity of a large biomass production imposes contact between bacteria and the substrate material, which is generally achieved by the potent mixture inside the reactor.

For biogas desulfurization, one utilizes a blower that circulates metered air through a network of pipes, first in the digesters and then in the post-fermenter as well as in the secondary fermenter.

The prescription regarding the dosage of the air quantity depends on the oxygen content of the biogas. Measuring the oxygen content of the respective tank requires a process analysis system. Then open the magnetic valves and send the gas directly to the tank. The amount of air is regulated so that it carries a volumetric flow of biogas of 4% (maximum 6%), produced during the same period.

The biogas plant operates continuously, stopping the process being possible only in exceptional conditions. Yet the cogeneration plant only works 16 h/day, during the daytime hours, due to economic reasons.

3 Data Acquisition System of Cogeneration Power Plants

There are continuous monitoring and coordination of the cogeneration power plant engines, executed by a dedicated control system [13]. The parameters measured by the data acquisition system are, for example, electricity produced by each engine, biogas consumption, internal electricity consumption (of the cogeneration and the biogas production plants), power consumption of the biogas station.

The monthly operating hours are referring to engine 1 and 2 (E1 and E2) appear in Figs. 2, 3, and 4.

Figures 5, 6 and 7 contain the illustration of the electricity monthly output [MWh] for each engine, respectively overall output, for each year taken into consideration.

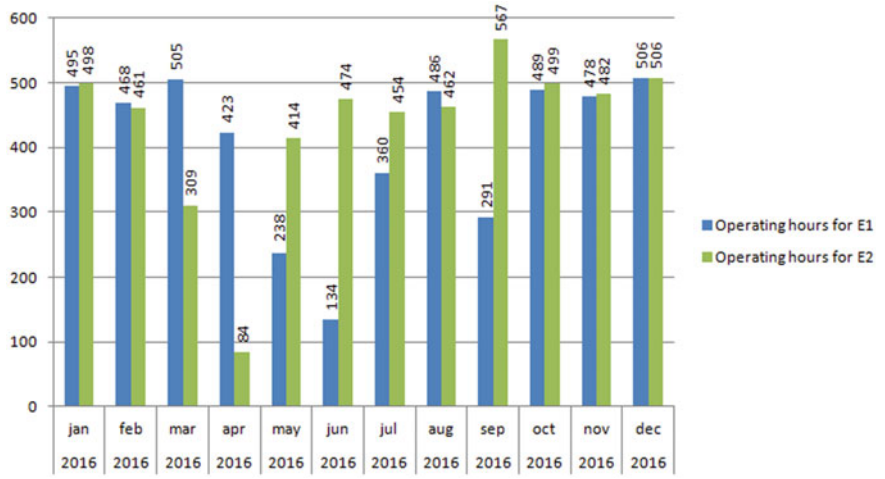


Fig. 2 Monthly operating hours for each engine, for year 2016

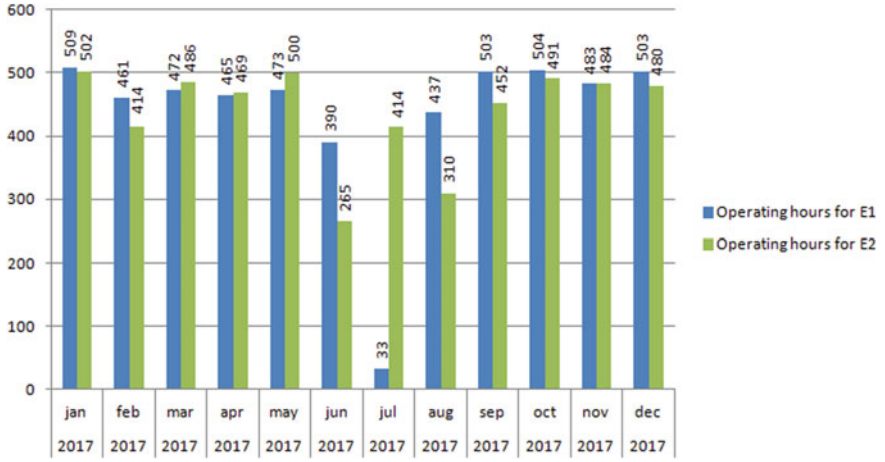


Fig. 3 Monthly operating hours for each engine, for year 2017

Figures 8, 9 and 10 display graphics represent the monthly electric energy output for the plant (E1 and E2 cumulated) for each year, and also the monthly electric energy supplied into the National Power System (NPS).

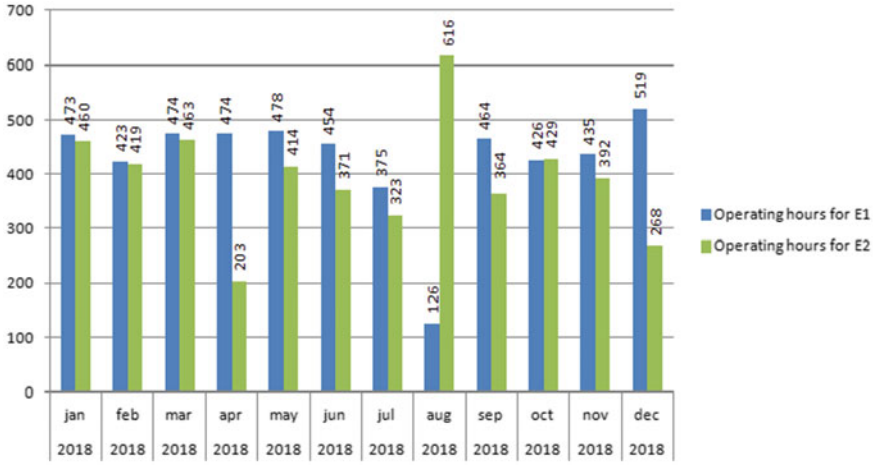


Fig. 4 Monthly operating hours for each engine, for year 2018

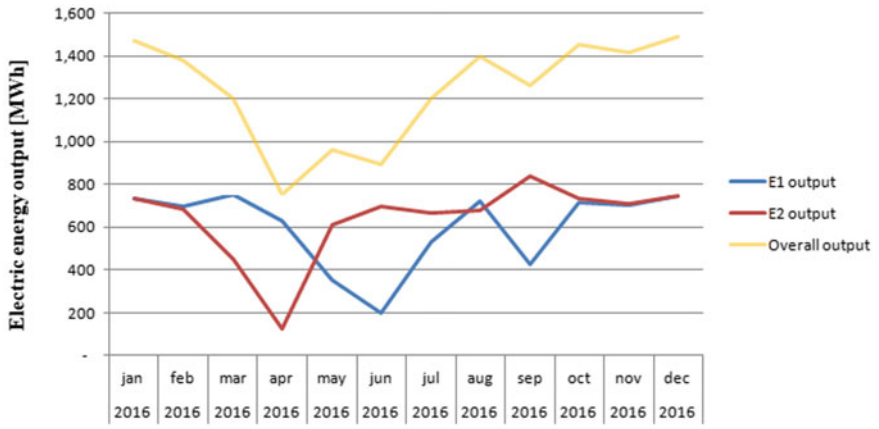


Fig. 5 Electrical Energy produced by E1 and E2 and the total electrical energy produced in 2016

The difference between the overall electric energy output of the engines and the electric energy supplied into the NPS counts for the technological consumption and the network losses, presented in Figs. 11, 12 and 13.

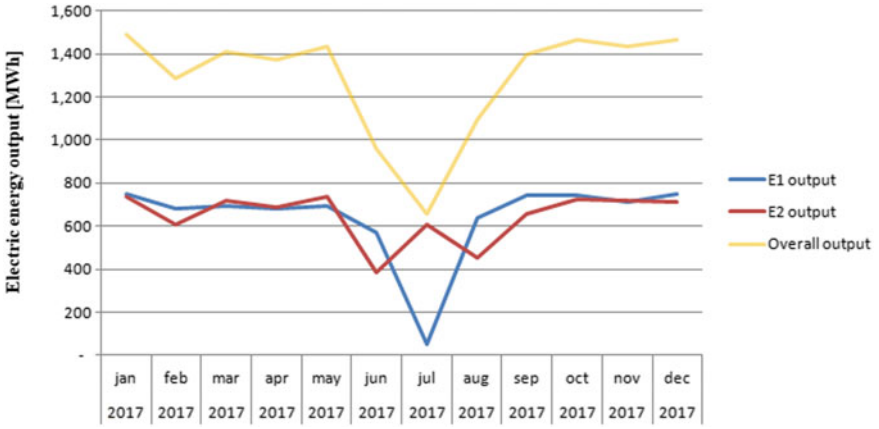


Fig. 6 Electrical energy produced by E1and E2 and the total electrical energy produced in 2017

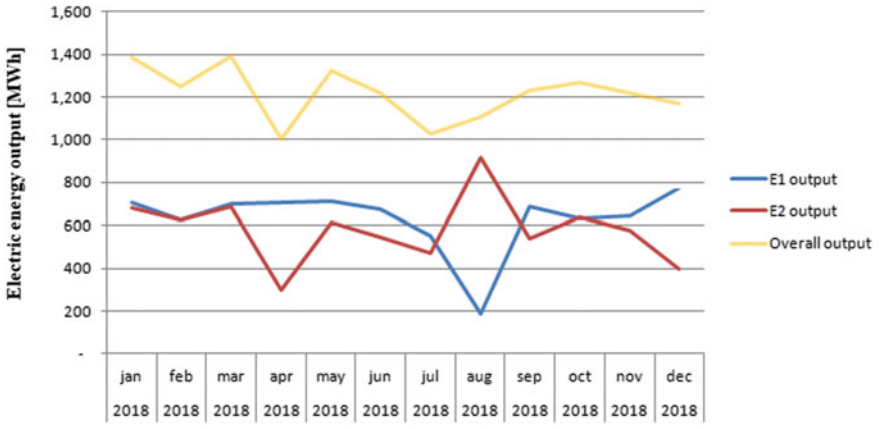


Fig. 7 Electrical energy produced by E1and E2 and the total electrical energy produced in 2018

For the production of the electricity mentioned above, the plant uses biogas obtained in its biogas plant. There are different raw materials available for biogas production, such as. One can give examples of plants from various energy crops and organic household or livestock manure. However, only the plant managers can decide to use forage maize and sugar beet or rye noodles, more comfortable to obtain and manipulate. An industrial weight scale measures the amount of biomass feeding the biogas plant.

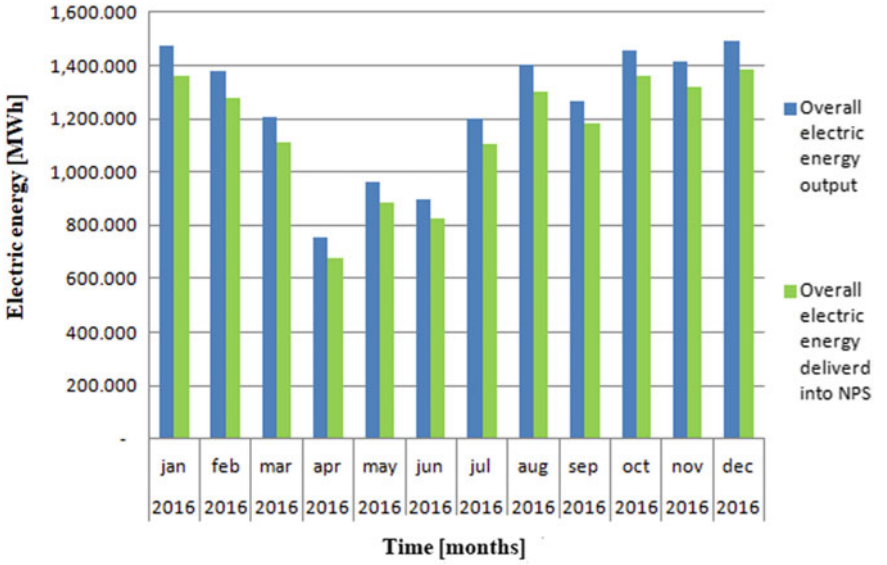


Fig. 8 Electrical energy produced and delivered into NPS in 2016

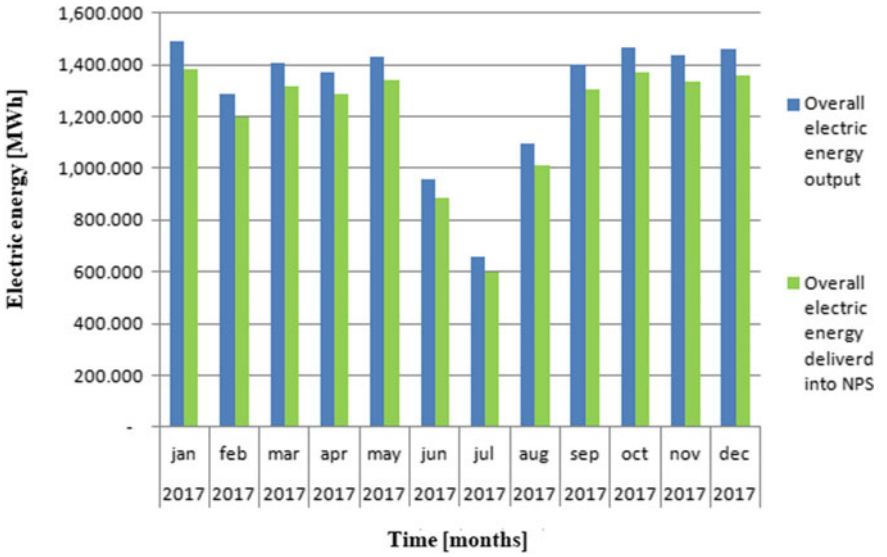


Fig. 9 Electrical energy produced and delivered into NPS in 2017

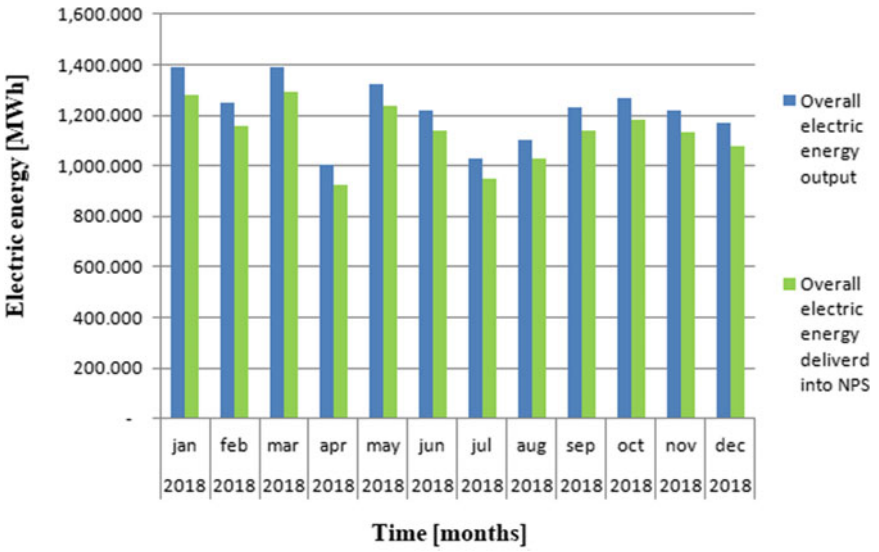


Fig. 10 Electrical energy produced and delivered into NPS in 2018

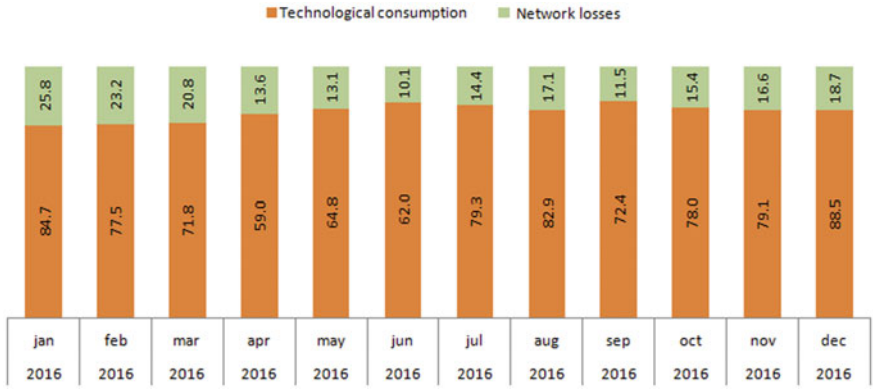


Fig. 11 Technological consumption of energy and network energy losses in 2016

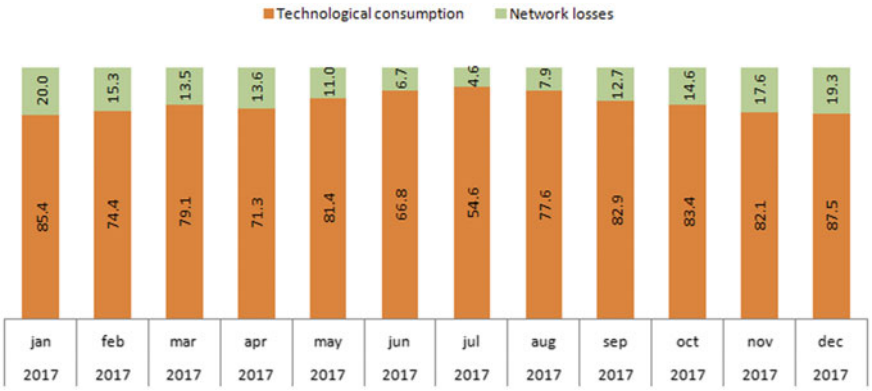


Fig. 12 Technological consumption of energy and network energy losses in 2017

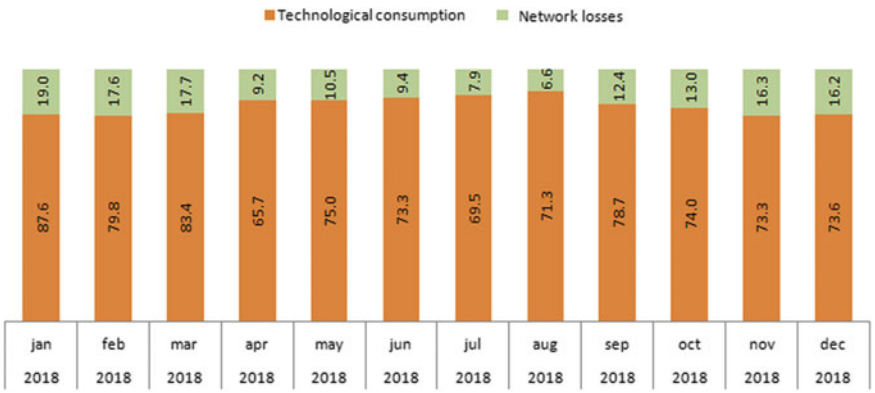


Fig. 13 Technological consumption of energy and network energy losses—2018

Figures 14, 15, and 16 display the biomass consumption of the plant for the three years of production analyzed.

When purchasing the raw material needed for biogas production, analyses are necessary to determine the calorific power of each type of fuel. One used the following relation to determine the energy contained in the raw material:

$$E_{total\ raw\ material} = Calorific\ power_{raw\ material\ 1} * Quantity_{raw\ material\ 1} + Calorific\ power_{raw\ material\ 2} * Quantity_{raw\ material\ 2} \quad (1)$$

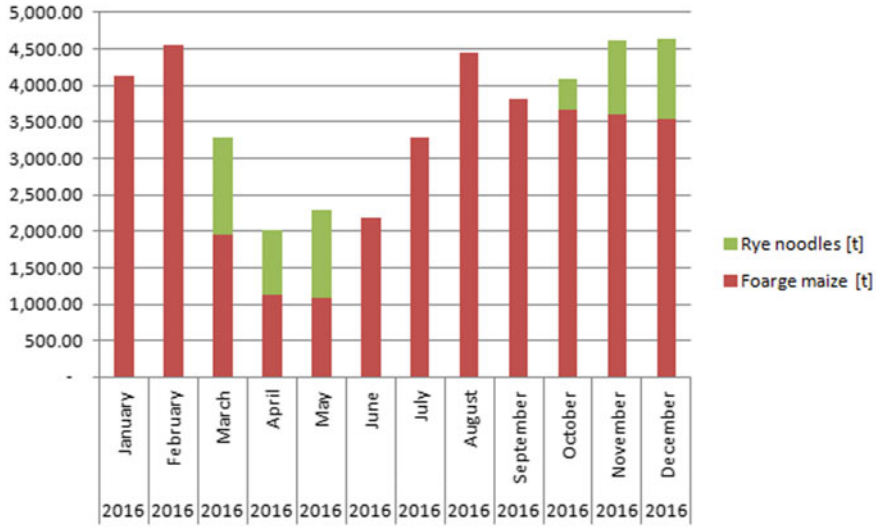


Fig. 14 Biomass consumption of the plant in 2016

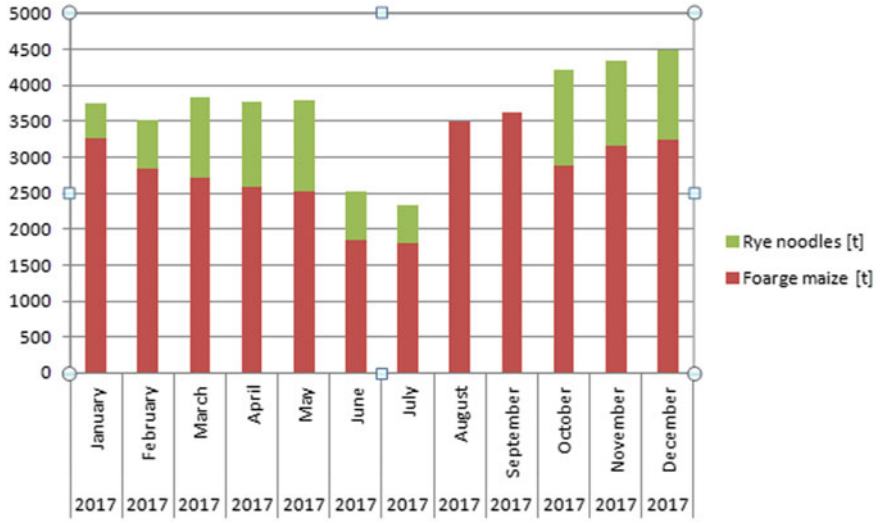


Fig. 15 Biomass consumption of the plant in 2017

For example, in May 2018, the Calorific power for the forge maize was, according to the analysis bulletin, 3.232 MWh/t and for the rye noodles 0.099 MWh/t. this way results

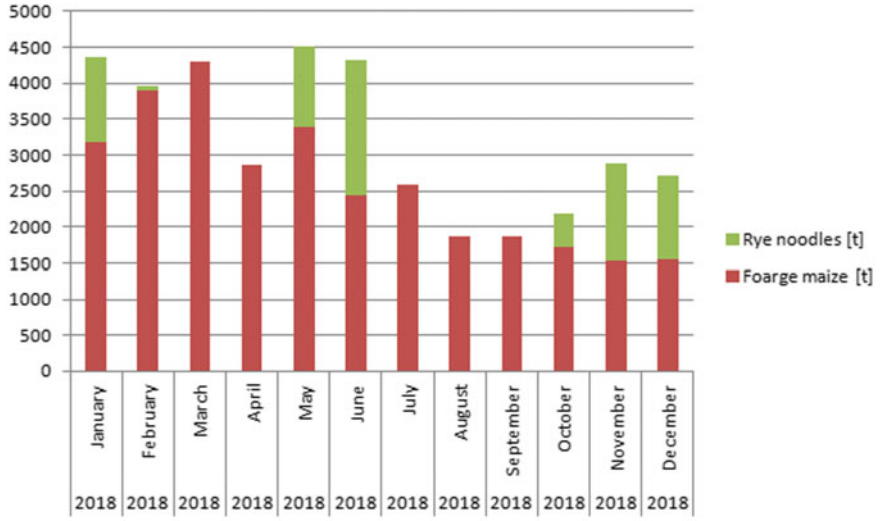


Fig. 16 Biomass consumption of the plant in 2018

$$E_{total\ raw\ material} = 3.232\,\text{MWh/t} * 3399.172\,\text{t} + 0.099\,\text{MWh/t} * 1103.9\,\text{t} = 11095.40\,\text{MWh} \tag{2}$$

Taken this into consideration, the total energy from raw materials used in the three years in the biomass cogeneration plant studied, is shown in Figs. 17, 18, and 19.

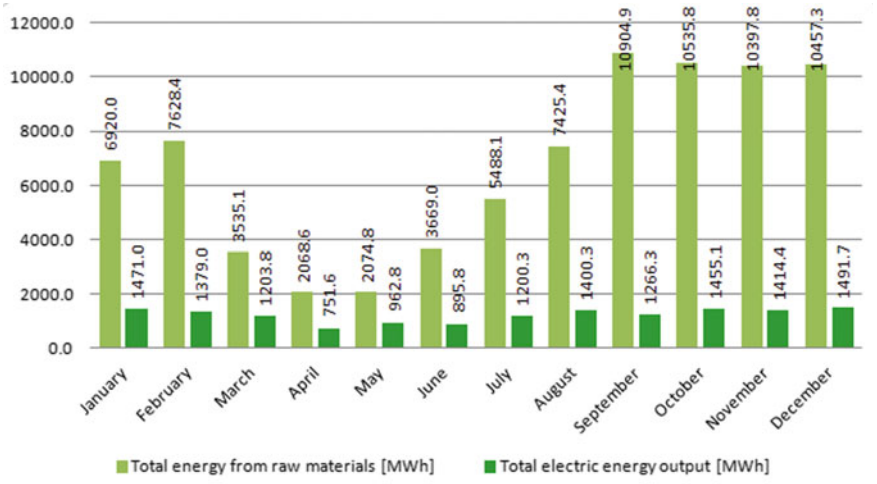


Fig. 17 Total energy from raw materials and total electric energy output, for year 2016

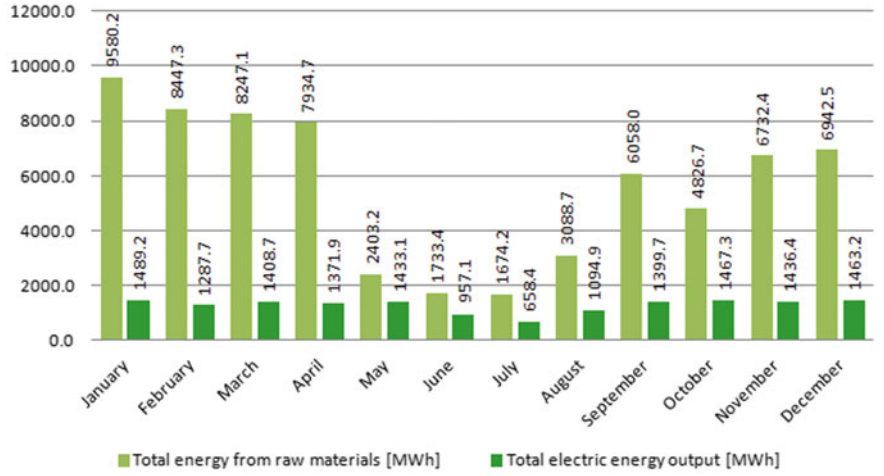


Fig. 18 Total energy from raw materials and total electric energy output, for year 2017

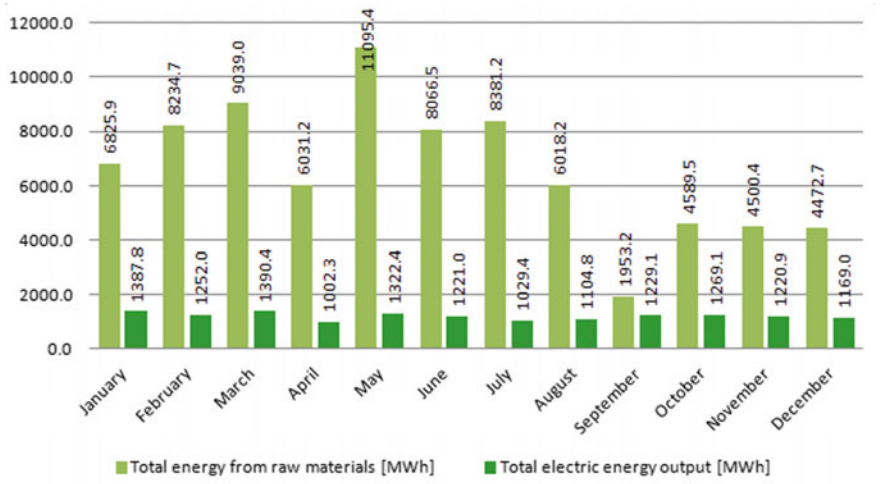


Fig. 19 Total energy from raw materials and total electric energy output, for year 2018

4 Numerical Approaches of CPP Efficiency

Analyzing the performance of the cogeneration plant, we observe that the engines operate at the parameters for which they were designed, except the stopping time, necessary for corrective or preventive maintenance. Thus, the only way we can improve the economic efficiency of the plant is by optimizing the fuel used for its operation.

Approaches of CPP efficiency by the quantity of biomass

Thus, to prevent the purchase and use of excess biomass, The PyCHARM Community 2019.1 program [14–16] was used to determine an accurate mathematical relationship between the energy produced and the biomass consumption [17].

The use of the data from above reproduced in Figs. 17, 18 and 19 and processed with the source code specially developed for their interpretation, resulted in a mathematical dependence presented in Fig. 20.

The best mathematical correlation between the real, respectively, the numerical model of the biomass consumption and the electrical energy generated in the boiler represents a 6th-degree polynomial function. The square error, in this case, is large with $R2 = 0.3705$, away from the ideal value $R2 = 1$.

Trying to improve the accuracy of the approximation function, we have divided the year into 4 quarters. Doing this, and willing to obtain a function which better approximates the quantity of biomass needed for each quarter, we could also order the values in increasing order of the amount of biomass used in each trimester (quarter).

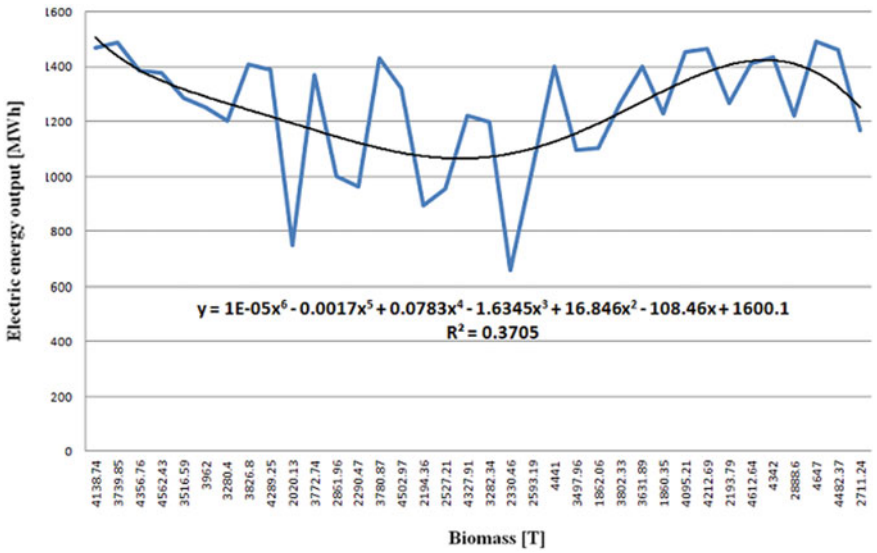


Fig. 20 The mathematical dependency between the consumption of biomass and the energy produced, 2016–2018

Table 2 Polynomial function order 1–7 for trim. 1

Function	Order	R-squared
$Poly1 = 1.12075543e-01x + 9.19075480e+02$	1	0.233129
$Poly3 = 2.89179625e-07x^3 - 3.66797890e-03x^2 + 1.54532874e+01x - 2.02310326e+04$	2	0.430852
$Poly3 = 2.89179625e-07x^3 - 3.66797890e-03x^2 + 1.54532874e+01x - 2.02310326e+04$		0.459028
$Poly4 = -3.98546993e-10x^4 + 6.56590173e-06x^3 - 4.05668409e-02x^2 + 1.11405341e+02x - 1.13346965e+05$	4	0.464148
$Poly5 = -4.84476832e-12x^5 + 9.46516927e-08x^4 - 7.36808664e-04x^3 + 2.85631735e+00x^2 - 5.51344208e+03x + 4.24002845e+06$	5	0.510650
$Poly6 = 7.52495742e-14x^6 - 1.77959206e-09x^5 + 1.74908583e-05x^4 - 9.14470812e-02x^3 + 2.68228254e+02x^2 - 4.18481000e+05x + 2.71306704e+08$	6	0.871552
$Poly7 = -1.50655062e-16x^7 + 4.23357875e-12x^6 - 5.08757280e-08x^5 + 3.38904630e-04x^4 - 1.35148952e+00x^3 + 3.22624323e+03x^2 - 4.26862749e+06x + 2.41470295e+09$	7	0.908909

Whereas applying the same interpolation algorithm, one obtained the results shown in Tables 2, 3 and 4, and respectively 5, and the graphics represented in Figs. 21, 22, 23, and 24.

Approaches of CPP efficiency by the quality of biomass

The quality of the biomass given in this particular case, by the calorific power of the raw materials, indicates the energy contained by the energy cultures used for powering the cogeneration engines.

To determine a mathematical link between the electricity produced in the cogeneration plant and the energy contained in the raw material, we have developed similar source code and used the same PyCHARM Community 2019.1 software.

Table 3 Polynomial function order 1–7 for trim. 2

Function	Order	R-squared
$Poly1 = 1.04864947e-01\ x\ + 6.86345737e+02$	1	0.033060
$Poly2 = -4.94649902e-04\,x^2 + 3.98798121e+00\ \ x - 6.85762521e+03$	2	0.141336
$Poly3 = 1.64966240e-06x^3 - 1.98844099e-02x^2 + 7.94561656e+01x - 1.04105260e+05$	3	0.289858
$Poly4 = 2.35707483e-10x^4\ - 2.06249801e-06x^3 + 1.93820590e-03x^2 + 2.27084859e+01x - 4.90349103e+04$	4	0.290148
$Poly5 = 4.13917164e-11x^5 - 8.11834613e-07x^4 + 6.34902535e-03\,x^3 - 2.47478540e+01x^2 - 4.80791012e+04\,x - 3.72424916e+07$	5	0.839966
$Poly6 = -8.84702678e-14x^6 + 2.12794662e-09x^5 - 2.12644030e-05x^4 + 1.12996301e-01x^3 - 3.36743325e+02\ x^2 + \ 5.33601456e+05x - 3.51230452e+08$	6	0.920772
$Poly7 = -4.92798532e-16x^7 + 1.35135852e-11x^6 - 1.58467412e-07x^5 + 1.03009247e-03x^4 - 4.00865121e+00\ x^3 + 9.33903799e+03\,x^2 - 1.20603765e+07x\ + 6.65990152e+09$	7	0.985518

The mathematical functions obtained this way appear successively in the Tables 6, 7, 8 and 9, and the graphs for the most accurate functions (the highlighted ones) are shown in Figs. 25, 26, 27, and 28.

5 Conclusion

Unlike the solar and wind sources of renewable energy that are intermittent and unpredictable because they depend on geographical and atmospheric conditions, the cogeneration plants utilize biomass as fuel and produce constant and efficient green energy. Cogeneration plants yield simultaneous electricity and heat using the same amount of biomass; in this way, there is an increase in the total efficiency of such a plant. In this regard, the analysis of the dependency between the amount of biomass as input and the electric and thermal energy parameters as the output of a cogeneration plant is a significant concern for users and researchers.

Table 4 Polynomial function order 1–7 for trim. 3

Function	Order	R-squared
$Poly1 = 1.04864947e-01\ x\ + 6.86345737e+02$	1	0.033060
$Poly2 = -4.94649902e-04\,x^2 + 3.98798121e+00\ \ x - 6.85762521e+03$	2	0.141336
$Poly3 = 1.64966240e-06x^3 - 1.98844099e-02x^2 + 7.94561656e+01x - 1.04105260e+05$	3	0.289858
$Poly4 = 2.35707483e-10x^4\ - 2.06249801e-06x^3 + 1.93820590e-03x^2 + 2.27084859e+01x - 4.90349103e+04$	4	0.290148
$Poly5 = 4.13917164e-11x^5 - 8.11834613e-07x^4 + 6.34902535e-03\,x^3 - 2.47478540e+01x^2 - 4.80791012e+04\,x - 3.72424916e+07$	5	0.839966
$Poly6 = -8.84702678e-14x^6 + 2.12794662e-09x^5 - 2.12644030e-05x^4 + 1.12996301e-01x^3 - 3.36743325e+02\,x^2 + 5.33601456e+05x - 3.51230452e+08$	6	0.920772
$Poly7 = -4.92798532e-16x^7 + 1.35135852e-11x^6 - 1.58467412e-07x^5 + 1.03009247e-03x^4 - 4.00865121e+00\ x^3 + 9.33903799e+03x^2 - 1.20603765e+07x\ + 6.65990152e+09$	7	0.985518

Thereby based on annual on-line data acquisition, one elaborated a numerical method of interpolation, finally implemented in Python environmental software and used to determine the dependency between the input of consumed biomass and the electrical energy output. The data separation into four periods has the justification of obtaining approximation functions of the polynomial type, having the highest accuracy. The data represents the most common biomass characteristics for every single year of study.

Table 5 Polynomial function order 1–7 for trim. 4

Function	Order	R-squared
$Poly1 = -1.11805977e-01x + 2.09514429e+03$	1	0.264080
$Poly2 = -1.35895524e-04x^2 + 1.57181579e+00x - 3.07406111e+03$	2	0.441242
$Poly3 = -3.87098904e-07x^3 - 7.08583488e-03x^2 - 4.30319357e+01x - 8.80734824e+04$	3)	0.648110
$Poly4 = 1.15779208e-09x^4 - 2.96663653e-05x^3 + 2.83594873e-01x^2 - 1.19845618e+03x + 1.88996527e+06$	4	0.711585
$Poly5 = 1.22634900e-11x^5 - 3.92436923e-07x^4 + 5.00946478e-03x^3 - 3.18803944e+01x^2 + 1.01133033e+05x - 1.27908657e+08$	5	0.921211
$Poly6 = -2.76903094e-14x^6 + 1.07872125e-09x^5 - 1.74743062e-05x^4 + 1.50648448e-01x^3 - 7.28915930e+02x^2 + 1.87654693e+06x - 2.00788593e+09$	6	0.970026
$Poly7 = 1.47032333e-16x^7 - 6.64575860e-12x^6 + 1.28566988e-07x^5 - 1.37990047e-03x^4 - 8.87353583e+00\ x^3 - 3.41859071e+04x^2 - 7.30538241e+07x - 6.67950423e+10$	7	0.974018

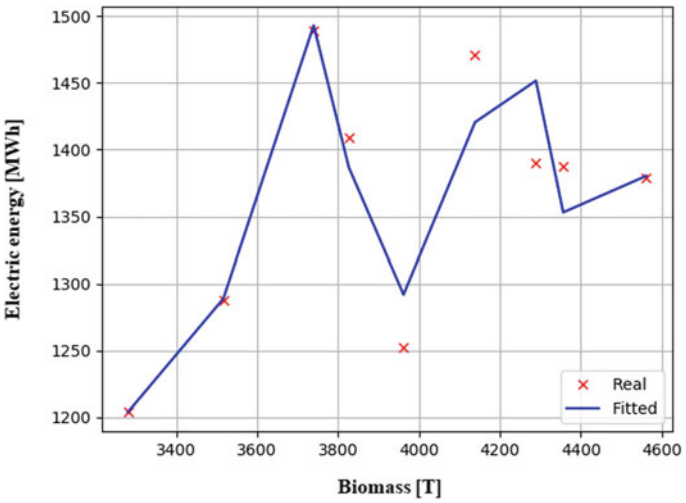


Fig. 21 Energy (MWh) versus Biomass (t) for the first quarter

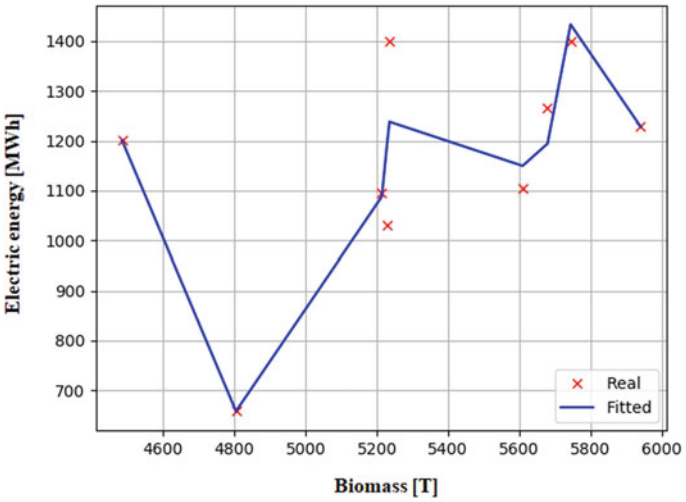


Fig. 22 Energy (MWh) versus Biomass (t) for the second quarter

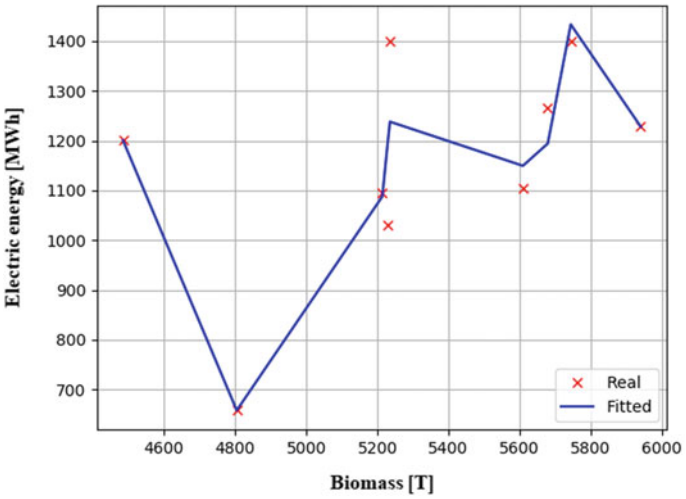


Fig. 23 Energy (MWh) versus Biomass (t) for the third quarter

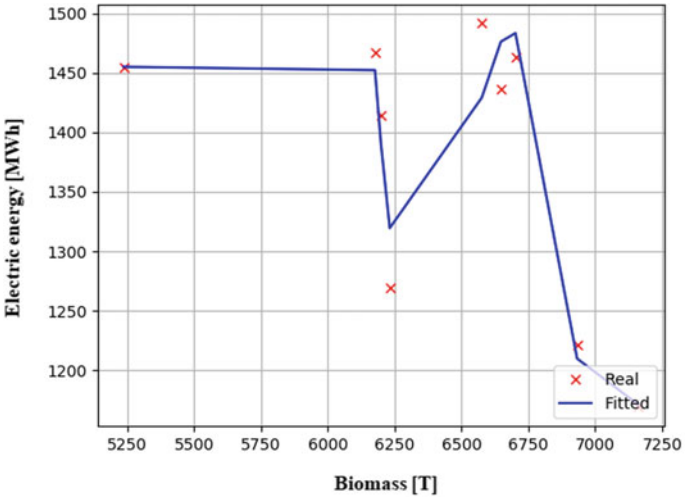


Fig. 24 Energy (MWh) versus Biomass (t) for the fourth quarter

Table 6 Polynomial function order 1–7 for the entire period

Function	Order	R-squared
$Poly1 = -9.3135x + 6505.1$	1	0.0012
$Poly2 = -6.0755x^2 + 215.48x + 5081.4$	2	0.0437
$Poly3 = 0.0985x^3 - 11.544x^2 + 297.53x + 4811.3$	3	0.0446
$Poly4 = -0.097x^4 + 7.2766x^3 - 183.89x^2 + 1760.8x + 1771.8$	4	0.1175
$Poly5 = -0.0149x^5 + 1.2819x^4 - 38.407x^3 + 463.76x^2 - 1857.8x + 7091.7$	5	0.2556
$Poly6 = 0.003x^6 - 0.3513x^5 + 15.51x^4 - 323.85x^3 + 3202.8x^2 - 12879x + 19478$	6	0.7081

Table 7 Polynomial function order 1–6 for 2016

Function	Order	R-squared
$Poly1 = 611.51x + 2783.9$	1	0.4275
$Poly2 = 150.83x^2 - 1349.3x + 7359.1$	2	0.6702
$Poly3 = -48.419x^3 + 1095x^2 - 6457.5x + 13968$	3	0.8873
$Poly4 = -6.4049x^4 + 118.11x^3 - 336.96x^2 - 1913.7x + 9971.7$	4	0.9181
$Poly5 = 3.2044x^5 - 110.55x^4 + 1350x^3 - 6759x^2 + 12115x + 529.48$	5	0.9762
$Poly6 = -0.2293x^6 + 12.148x^5 - 245.44x^4 + 2337.9x^3 - 10365x^2 + 18094x - 2787.9$	6	0.9782

Table 8 Polynomial function order 1–6 for 2017

Function	Order	R-squared
$Poly1 = -264.96x + 7361.3$	1	0.1152
$Poly2 = 196.03x^2 - 2813.3x + 13307$	2	0.7035
$Poly3 = 3.1389x^3 + 134.82x^2 - 2482.2x + 12879$	3	0.7048
$Poly4 = -11.532x4 + 302.98x3 - 2443.5x2 + 5699.1x + 5682.9$	4	0.8485
$Poly6 = -1.2356x^6 + 47.181x^5 - 688.88x^4 + 4830.7x^3 - 16889x^2 + 27592x - 8103.9$	5	0.8508
$Poly6 = 1.08x^6 - 41.581x^5 + 606.2x^4 - 4142.4x^3 + 13461x^2 - 20103x + 19720$	6	0.9151

Table 9 Polynomial function order 1–6 for 2018

Function	Order	R-squared
$Poly1 = -440.37x + 9463.1$	Linear	0.3994
$Poly2 = -63.502x^2 + 385.16x + 7536.8$	2	0.4769
$Poly3 = 28.904x^3 - 627.14x^2 + 3434.6x + 3591.4$	3	0.6163
$Poly4 = 5.149x^4 - 104.97x^3 + 524.04x^2 - 218.3x + 6804.4$	4	0.6522
$Poly5 = -1.0061x^5 + 37.847x^4 - 491.76x^3 + 2540.4x^2 - 4622.9x + 9769$	5	0.6625
$Poly6 = -1.2356x^6 + 47.181x^5 - 688.88x^4 + 4830.7x^3 - 16889x^2 + 27592x - 8103.9$	6	0.7682

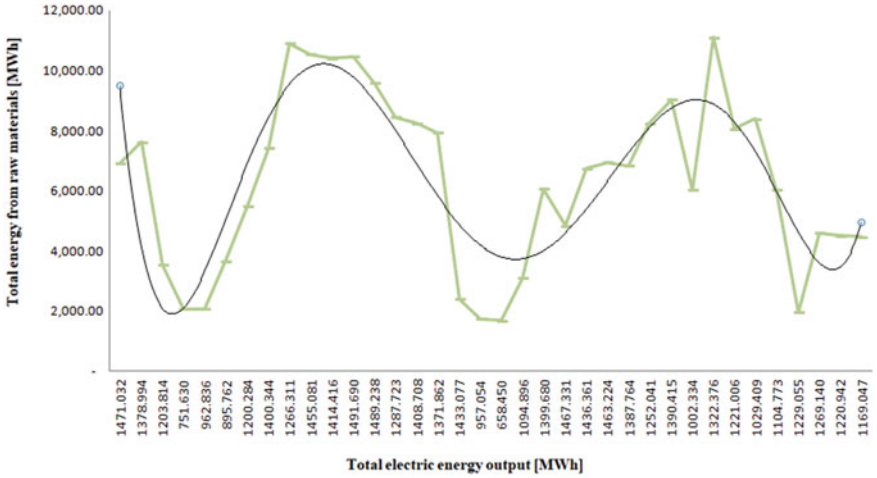


Fig. 25 The total electric energy output versus the energy contained in the raw material for 2016–2018

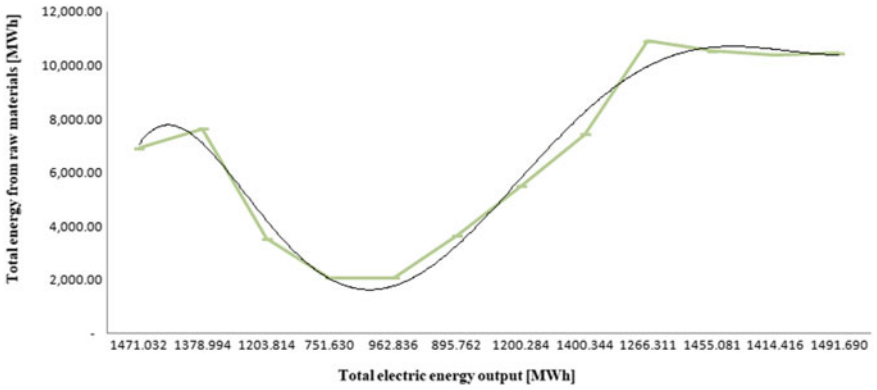


Fig. 26 The total electric energy output versus the energy contained in the raw material for 2016

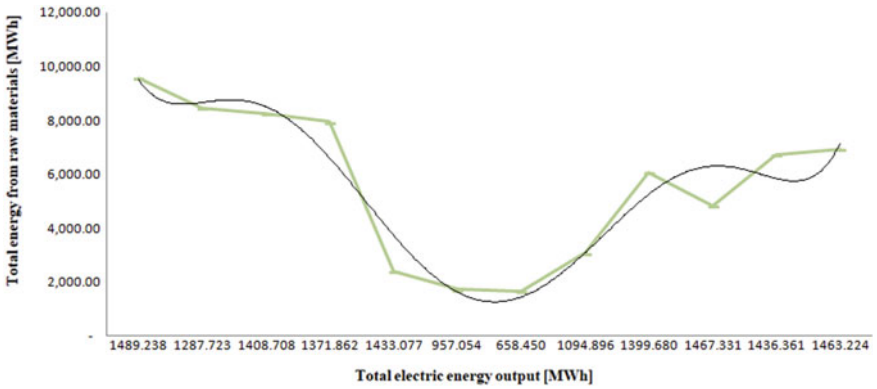


Fig. 27 The total electric energy output versus the energy contained in the raw material for 2017

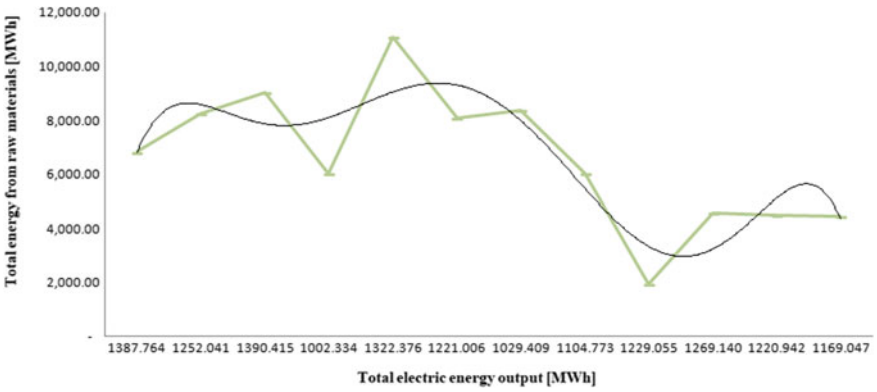


Fig. 28 The total electric energy output versus the energy contained in the raw material for 2018

References

1. https://history.com/topics…/thomas-edison
2. European Union—Cogeneration Directive 2004/08/EC. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:052:0050:0060:EN:PDF
3. Ali RH, Harby K, Maghrabie HM, Attalla M (2017) Exergy analysis of cogeneration power plant in paper industry. In: Proceedings of 4th international conference on energy engineering, faculty of energy engineering. Aswan University, Aswan, Egypt. https://www.researchgate.net/profile/Hussein_Maghrabie/publication/322146801_Exergy_Analysis_of_Cogeneration_Power_Plant_in_Paper_Industry/links/5a47dec4aca272d2945fd4a2/Exergy-Analysis-of-Cogeneration-Power-Plant-in-Paper-Industry.pdf
4. Cavalcanti EJC, de Souza GF, Lima MSR (2017) Exergy analysis of a cogeneration power plant. In: 24th ABCM Int congress of mechanical engineering, Curitiba, PR, Brazil. https://www.researchgate.net/profile/Eduardo_Cavalcanti3/publication/323214869_Exergy_Analysis_of_a_Cogeneration_Power_Plant/links/5b0da95a0f7e9b1ed7011d47/Exergy-Analysis-of-a-Cogeneration-Power-Plant.pdf

5. Usubharatana P, Phungrassami H (2018) Life cycle assessment for enhanced efficiency of small power plants by reducing air input temperature. Pol J Environ Stud 27(4):1–13. https://www.researchgate.net/profile/Harnpon_Phungrassami/publication/323564426_Life_Cycle_Assessment_for_Enhanced_Efficiency_of_Small_Power_Plants_by_Reducing_Air_Input_Temperature/links/5aaf7353a6fdcc1bc0bcbbb6/Life-Cycle-Assessment-for-Enhanced-Efficiency-of-Small-Power-Plants-by-Reducing-Air-Input-Temperature.pdf
6. Islas J, Manzini F, Masera O, Vargas V (2019) The role of bioenergy in the bioeconomy. Solid biomass to heat and power (Chapter 4). Academic Press, pp 145–177. https://www.sciencedirect.com/science/article/pii/B9780128130568000042
7. Cardosso J, Silva V, Eusebio D (2019) Techno-economic analysis of a biomass gasification power plant dealing with forestry residues blends for electricity production in Portugal. J Clean Prod 212:741–753. Elsevier, March 2019. https://www.sciencedirect.com/science/article/pii/S0959652618337569
8. Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32001L0077
9. Government Decision no. 443/10.04.2013 published in the Romanian Official Monitor no. 288/24.04.2013 on the promotion of electricity produced from renewable energy sources
10. Government Decision no. 958/2005 for the modification of the Government Decision no. 443/2003 regarding the promotion of the production of electricity from renewable energy sources and for the modification and completion of the Government Decision no. 1.892/2004 for establishing the system for promoting the production of electricity from renewable energy sources. https://lege5.ro/Gratuit/g4ydomby/hotararea-nr-958–2005-pentru-modificarea-hotararii-guvernului-nr-443-2003-privind-promovarea-productiei-de-energie-electrica-din-surse-regenerabile-de-energie-si-pentru-modificarea-si-completarea-hota
11. Law no. 220/27.10.2008 on establishing the system for promoting the production of energy from renewable energy sources. http://legislatie.just.ro/Public/DetaliiDocumentAfis/121467
12. Enescu A, Diaconu E (2018) Cogeneration plant on biomass-case study. Sci Bull Electr Eng Fac Vol 18(2):7–11. https://content.sciendo.com/view/journals/sbeef/18/2/article-p7.xml
13. Andrei H, Enescu A, Diaconu E, Ion V, Udroiu I (2019) Data acquisition and modeling of cogeneration power plant parameters. In: Proceedings of 11th symposium on advanced topics in electrical engineering (ATEE-IEEE), Bucharest, Romania. https://ieeexplore.ieee.org/document/8724738
14. Hu Q, Ma L, Zhao J (2018) Deep graph: a pycharm tool for visualizing and understanding deep learning models. In: Proceedings of 25th Asia-Pacific software engineering Conference (APSEC). Nara, Japan. https://ieeexplore.ieee.org/abstract/document/8719435
15. Oliphant TE (2007) Python for scientific computing. Comput Sci Eng 9(3):10–20 May–June 2007. https://ieeexplore.ieee.org/document/4160250
16. Lutz M (2001) Programming python, 2nd Edn. O'REILLY. https://books.google.ro/books?hl=en&lr=&id=c8pV-TzyfBUC&oi=fnd&pg=PR11&dq=Lutz+M+(2001)+Programming+python.+2nd+Edition&ots=n4aD5RSUVX&sig=THSyIfN-qHVuyZ22EtDteTr-8EE&redir_esc=y#v=onepage&q=Lutz%20M%20(2001)%20Programming%20python.%202nd%20Edition&f=false
17. Enescu A, Andrei H, Diaconu E, Ion V (2019) Numerical method for modeling the input-output characteristic in a cogeneration power plant. In: Proceedings Of 11th conference on electronics computers and artificial intelligence (ECAI-IEEE). Pitesti, Romania. http://ecai.ro/Documente/ECAI%202019%20program%20V2.pdf

Power and Energy Flow in Cvasi-Stationary Electric and Magnetic Circuits

Horia Andrei, Mihai Iordache, Paul Cristian Andrei, Marilena Stanculescu, Sorin Deleanu, and Lavinia Bobaru

Abstract Answering the question "what is the state in which conservative systems consume less power or energy?" is fundamental. Therefore, multitudinous works were dedicated to formulate the cvasi-stationary state of many domains such as physical sciences (mechanics, thermodynamics, electromagnetic), chemistry, life science (hydrology, meteorology, global climate) in power or energy terms. Based on the variational principles in this chapter specific functionals expressed in terms of power or energy for electric respectively magnetic circuits in cvasi-stationary state are defined. The matrix equations of electro-magnetic circuits formulated in terms of electric and magnetic potentials of nodes were used to calculate the power and energy functionals. Further used advanced numerical methods the existence of functional's minimum were demonstrated and by imposing the minimization conditions are obtained the first Kirchhoff's law for electric currents respectively magnetic flux. Several examples prove the theoretically and practically importance of the principles of minimum consumed power and energy mainly for understanding of the power and energy flow in electromagnetic systems.

H. Andrei (✉)
Department of Doctoral School, University Valahia, Targoviste, Romania
e-mail: hr_andrei@yahoo.com

M. Iordache · P. C. Andrei · M. Stanculescu · L. Bobaru
Department of Electrical Engineering, University Politehnica, Bucharest, Romania
e-mail: mihai.iordache@upb.ro

P. C. Andrei
e-mail: paul.andrei@upb.ro

M. Stanculescu
e-mail: marilena.stanculescu@upb.ro

L. Bobaru
e-mail: lavinia.bobaru@upb.ro

S. Deleanu
SAST Northern Alberta Institute of Technology, Edmonton, Canada
e-mail: sorind@nait.ca

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_24

Keywords Variational principles · Functionals · Quasi-stationary state · Electric and magnetic circuits · Active and reactive power · Principles of minimum consumed power and energy

Nomenclature

A. Acronyms

DOF	Degrees of Freedom
KCL	Kirchhoff Current Law
KVL	Kirchhoff Voltage Law
DC	Direct Current
AC	Alternating Current
NM	Nodal Method
PMP	Principle of Minimum Consumed Power
RLC	Resistor Inductance Capacitance
PMARP	Principle of Minimum Active and Reactive Power
ECAP	Electric Circuit Analysis Program
KMVL	Kirchhoff's Magnetic Voltage Law
KMFL	Kirchhoff Magnetic Flux Law
PMEM	Principle of Minimum Consumed Energy for Magnetic Circuits

B. Symbols/Parameters

$p_J(\mathbf{r},t)$	The volume density of the instantaneous electromagnetic power
$\mathbf{E}$	The vector of electric field strength
$\mathbf{J}$	The vector of electric conduction current density
dW_{em}	The variation of the electromagnetic field energy
$P_{\delta\Omega}dt$	The energy transferred through the domain boundary
dt	The time interval
W_m	The magnetic energy
W_e	The electric energy
$\mathbf{D}$	The vector of electric flux density
$\mathbf{B}$	The vector of magnetic flux density
ε	The absolute permittivity of the medium
μ	The absolute permeability of the medium
$\mathbf{S}$	The Poynting vector
$\mathbf{A}$	The magnetic potential vector
σ	The conductivity of conductors
Φ	The magnetic flux
$\Re$	The real set
H	The Hessian matrix

C^1	The continuous functions of class one
$\Delta\mathbf{v}$	The vector of branch voltages
$\mathbf{i}$	The vector of branch currents
$\mathbf{u}$	The vector of voltages at the branch resistances terminals
$\mathbf{e}$	The vector of voltage sources
V_i	The potential of node i
$\mathbf{V}$	The reduced vector of the nodes potentials
$\mathbf{C}$	The reduced branch-to-node incidence matrix
$\mathbf{G}$	The branch conductance diagonal matrix
$\Im(\mathrm{V})$	The power functional
P_{cons}	The power consumed by all the resistances
2-D	Two-dimensionales
$\Delta\underline{V}_k$	The k-branch complex voltage
$\mathbf{Y}$	The diagonal admittance matrix
G_k	The conductance of k-branch
B_k	The susceptance of k-branch functionals
$\Im_P$	The active power functional
$\Im_Q$	The reactive power functional
$\underline{\mathrm{S}}$	The complex power
$\underline{V}_i$	The complex potential of node i
Φ	The fascicular flux through circuit branch
R_m	The reluctance of the magnetic circuit branch
l	The length of magnetic material
A	The cross section area of magnetic material
Λ	The permeance
θ	The magnetomotive force
$\mathbf{V_m}$	The vector of the magnetic potential of circuit nodes
dim	The dimension of matrix or vector
$\mathbf{\Lambda}$	The branch permeance diagonal matrix
$\Im(\mathbf{V}_m)$	The magnetic energy functional
W_{m}	The consumed magnetic energy
n_R	The number of resistances
N	The number of nodes
K	The number of branches
$P_R(I)$	The power-current characteristic
$\underline{E}_1$; R_1, R_2;C_2; L_3;ω	The number and values of AC circuit parameters (complex voltage source, resistances, capacitance, inductance, angular velocity)
M	The number of turns
δ	The air-gap of magnetic circuit

1 Introduction

For many conservative systems as in thermodynamics, mechanics, hydrology, meteorology, electromagnetics is used the variational principle to formulate their quasi-stationary state or equilibrium regime in energy and power parameters [1–3]. In this respect specific power and energy functionals are defined and numerical methods are used to find their extremum points.

For example in the classical mechanics two categories of principle are employed: differential principles and variational (integral) principles [4]. First one as well as d'Alembert and Gauss principles inspect the mechanical parameters at a given time, whilst the variational principle like Maupertuis and Hamilton principles examine the mechanical parameters within a finite time interval and space in order to determine the parameters values that achieve particular integrals stationary.

In the classical thermodynamics specific thermodynamic potential are defined in order to analyze the equilibrium state and to measure the properties of materials [5, 6]. Pressure, temperature, volume and entropy are the thermodynamic parameters that can be studied using the thermodynamic potentials. If the entropy and volume of a closed system are kept constant, then the internal energy decreases to its minimum value at steady-state. Such being the case the second principle of thermodynamics is defined as the minimum energy principle.

In the case of intricate Earth system processes as hydrology, meteorology, global climate, the principles of minimum and maximum entropy production have been formulated to analyze the planetary energy balance [7, 8]. For linear system with permanent boundary conditions and which has several degrees of freedom (DOF) the minimum entropy principle is applied, to analyze the cases in which the disturbances of the system are far from its equilibrium state. Instead for non-linear systems with several degrees of freedom the maximum entropy principle is applied. In this case, many steady states can take place, and it is feasible to choose one of the steady state with maximum entropy production.

In the electromagnetic theory, if it is consider a domain Ω where exists electromagnetic field its energy can be turn into mechanical work, heat or other forms of energy. This energy is, on the one hand, transformed into other forms of energy, and the rest can leave the domain through its boundary. The energy conversion from the electromagnetic form in other forms of energy and vice versa is established at every point in domain by the conduction process law [9, 10]:

$$\mathrm{p_J} = \mathbf{E} \cdot \mathbf{J} \tag{1}$$

where $p_J(\boldsymbol{r},t)$ represents the volume density of the instantaneous electromagnetic power a scalar function of position and time, **E** and **J** are the vector of electric field strength respectively of electric conduction current density. In the domain Ω the following equality between energies is true:

$$\int_{\Omega} \mathbf{E} \cdot \mathbf{J} dv dt + dW_{em} + P_{\partial\Omega} dt = 0 \tag{2}$$

where $dW_{em}$is the variation of the electromagnetic field energy, and $P_{\delta\Omega} dt$ is the energy transferred through the domain boundary. In the time interval dt, relation (2) yields:

$$-\int_{\Omega} \mathbf{E} \cdot \mathbf{J} dv = \frac{dW_{em}}{dt} + P_{\partial\Omega} \tag{3}$$

The theorem of electromagnetic energy based on Maxwell's equations and relation (3) demonstrates the following relations [11]:

$$W_{em} = \int_{\Omega} \frac{\mathbf{D}^2}{2\varepsilon} dv + \int_{\Omega} \frac{\mathbf{B}^2}{2\mu} dv = W_e + W_m \tag{4}$$

$$W_e = \int_{\Omega} \frac{\mathbf{D}^2}{2\varepsilon} dv = \int_{\Omega} \frac{\mathbf{DE}}{2} dv = \int_{\Omega} \frac{\varepsilon \mathbf{E}^2}{2} dv \tag{5}$$

$$W_m = \int_{\Omega} \frac{\mathbf{B}^2}{2\mu} dv = \int_{\Omega} \frac{\mathbf{BH}}{2\mu} dv = \int_{\Omega} \frac{\mu \mathbf{H}^2}{2} dv \tag{6}$$

where W_m and W_e represents the magnetic respectively electric energy as component of the electromagnetic field energy, **D** and **B** represents the vectors of electric respectively magnetic flux density, ε and μ represents the absolute permittivity respectively permeability of the medium. Another conclusion of the above mentioned theorem states that the electromagnetic power transferred to the surroundings through domain boundary is given by:

$$P_{\partial\Omega} = \oint_{\Omega} (\mathbf{E} \times \mathbf{H}) n dS = \oint_{\Omega} \mathbf{S} n dS \tag{7}$$

where the vector $\mathbf{S} = \mathbf{E}$ x $\mathbf{H}$ is named the Poynting vector.

In previous works, the authors have chosen the potentials of nodes as variables whereas utilizing the matrix equation of the circuits. Further by imposing the minimization conditions of the power functionals is obtained the Kirchhoff Current Law (KCL).

In the second section, advanced numerical analysis is proposed to find the extreme point of power or energy functionals for electric and magnetic circuits in the quasi-stationary state. Lagrange multipliers and the variational method in Hilbert space have demonstrated the existence of the minimum of the functionals. The third section of

this chapter discusses the power flow in equilibrium conditions when the DC and AC circuit consumes minimum power, considering all classical powers (active and reactive).

Several examples implemented in PSPICE prove the theoretical principles of minimum consumed power statute in the previous section. It also shows that the transient regime of an electric circuit represents its passage between two quasi-stationary states with minimum power consumption and the co-existence of the fundamental theorem of maximum power transfer and the principles of minimum consumed power.

Based on the equivalence between the linear magnetic and electric networks, in section four, the minimum principle of consumed energy for magnetic circuits in the cvasi-stationary state is presented. Several examples prove the theoretical principle formulates by authors and put in evidence the applicability of this principle to the calculation of the energy and forces in electromagnetic types of equipment.

In conclusion, theoretically it can be stated that the proposed principles together with the Kirchhoff Voltage Law (KVL) determine an equivalent equations system to the classical one consisting of the KCL and KVL equations for DC and AC circuits. An analogous statement can also be concluded for the magnetic circuits in the cvasi-stationary state. On the other hand, from a practical point of view, the principles of minimum consumed power are very useful for the understanding of the power and energy flow in electromagnetic systems.

The chapter ends with a broad up-to-date list of references.

2 Advanced Numerical Analysis Applied to Determination of Power and Energy Functionals Extreme

2.1 Variational Method

In the classical analysis of electromagnetic field the variational equivalent formulation in the Hilbert space is used. Starting from a differential mathematical model the variational method establish a set of differential equations of the model complying with the cvasi-stationary conditions as indicated in Chap. 1 of Part I of this book.

Generally speaking the functional associated of the phenomenon depict by the scalar parameter *V(x, y, z)* is defined as [12]

$$\Im = \iiint\limits_{\Omega} f\left(x, y, z, V, \frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right) dxdydz + \iint\limits_{\Sigma} g(x, y, z) dS \qquad (8)$$

where f is a function specified by the know differential model of the phenomenon, $\partial V/\partial x$, $\partial V/\partial y$ and $\partial V/\partial z$ are the partial derivatives of the state quantity and g is a determined function on the boundary Σ of the domain Ω. The main idea of the

variational method associated to a phenomenon take into consideration the minimization of the expression (7) admitting that the differential equations of the model are verifying by the state parameters and its limit conditions.

For example in case of one-dimensional problem (1-D), the state parameter V depends only one coordinate, is defined in the domain $[x_1, x_2]$ and satisfies the limit conditions $V(x_1) = V_1$, $V(x_2) = V_2$, and the second integral of the relation (8) doesn't exist, then the functional associated of the phenomenon is expressed as:

$$\Im = \int_{\Omega} f(x, V, \frac{dV}{dx})dx \tag{9}$$

If it is consider $\tilde{V}(x)$ the approximate solution and is noted with $\delta V(x)$ the infinitesimal variation of the exact solution $\delta V(x)$ then the relation is true:

$$\tilde{V}(x) = V(x) + \delta V(x) \tag{10}$$

By imposed the stationarity condition and the minimum value of the functional (9) it is obtained for the functional variation $\delta\,\Im$ the relation:

$$\begin{aligned}\delta\Im &= \int_{x_1}^{x_2} \delta f \cdot dx = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial V}\delta V + \frac{\partial f}{\partial V'}\delta V'\right)dx \\ &= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial V}\delta V + \frac{\partial f}{\partial V'}\delta V'\right)dx = 0\end{aligned} \tag{11}$$

where $V' = \frac{dV}{dx}$ represents the derivative of V. Due to for the variation δf for a given value of the variable x in relation (10) is $\delta x = 0$, and then by using the parts integrating the last term of relation (11), results:

$$\begin{aligned}\int_{x_1}^{x_2} \frac{\partial f}{\partial V'}\delta V' dx &= \int_{x_1}^{x_2} \frac{\partial f}{\partial V'}\delta\left(\frac{dV}{dx}\right)dx = \int_{x_1}^{x_2} \frac{\partial f}{\partial V'}\frac{d}{dx}(\partial V)dx = \left[\frac{\partial f}{\partial V'}\delta V\right]_{x_1}^{x_2} \\ &- \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial V'}\right)dx\end{aligned} \tag{12}$$

The expression (11) becomes:

$$\delta\Im = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial V} - \frac{d}{dx}\left(\frac{\partial f}{\partial V}\right)\right]\delta V dx + \left[\frac{\partial f}{\partial V'}\delta V\right]_{x_1}^{x_2} = 0 \tag{13}$$

Considering the variation δV as an arbitrary one, than each term of relation (13) must be null, so:

$$\frac{\partial f}{\partial V} - \frac{d}{dx}\left(\frac{\partial f}{\partial V}\right) = 0 \tag{14}$$

and

$$\left[\frac{\partial f}{\partial V'}\delta V\right]_{x_1}^{x_2} = 0 \tag{15}$$

If the values of the state parameter *V(x)* at the two limits x_1 and x_2 of the domain are defined or, in other words, the Dirichlet conditions (forced limit conditions) are accomplished i.e.

$$\delta V(x_1) = 0 \text{ and } \delta V(x_2) = 0 \tag{16}$$

and then relation (15) is fulfilled. Otherwise in case of the state parameter doesn't satisfy Dirichlet forced limit conditions, then the following condition, named natural limit conditions, must be satisfied [13]:

$$\left[\frac{\partial f}{\partial V'}\right]_{x_1} = \left[\frac{\partial f}{\partial V'}\right]_{x_2} = 0 \tag{17}$$

The relation (15) is achieved also in the case in which the differential model implies at the two limits different conditions, namely at one natural limit condition and at the other one forced limit condition.

In the classical theory of the electromagnetic field the following functional of associated to the domain Ω and the volume bounded is defined as:

$$\Im = \int_{\Omega}\left[\left(\int_0^{E}\mathbf{D}\cdot\mathbf{E} - \int_0^{B}\mathbf{H}\cdot\mathbf{B}\right) + (\mathbf{J}\cdot\mathbf{A} - \rho_v V)\right]dxdydz \tag{18}$$

where $\mathbf{A}$ is the magnetic potential vector $\nabla \times \mathbf{A} = \mathbf{B}$ and V is the electric potential $\mathbf{E} = -\nabla V$. The minimization of the functional (18) implies the Maxwell's equations of the electromagnetic field, the physical properties of media, and the uniqueness conditions of the solution.

Thus the functional associated of the one-dimensional (1-D) electrostatic field is defined as:

$$\Im(V) = \int_{\Omega}\frac{\varepsilon}{2}\left[\left(\frac{\partial V}{\partial x}\right)^2 - \rho_v V\right]dx \tag{19}$$

The cvasi-stationary electric field the associated functional is defined as

$$\Im(V) = \int_{\Omega} \frac{\sigma}{2} \left[\left(\frac{\partial V}{\partial x} \right)^2 \right] dx \tag{20}$$

and the particular set of Maxwell 's equations available for linear, isotropic and homogenous media is

$$\nabla \times \mathbf{E} = 0 \tag{21}$$

$$\nabla \cdot \mathbf{J} = 0 \tag{22}$$

$$\mathbf{J} = \sigma \mathbf{E} \tag{23}$$

where σ is the conductivity of the conductors.

The cvasi-stationary magnetic field is governed by the particular set of Maxwell's equations

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{24}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{25}$$

$$\mathbf{B} = \mu \mathbf{H} \tag{26}$$

and is admit the associated functional

$$\Im(\Phi) = \int_{\Omega} \frac{\mu}{2} \left[\left(\frac{\partial \Phi}{\partial x} \right)^2 \right] dx \tag{27}$$

where Φ is the magnetic flux.

The functionals defined above (19), and (20), respectively (27) represent the power functionals for linear electric circuits, respectively energetic functional for linear magnetic circuits in cvasi-stationary state. In all these variational methods the electric and magnetic potentials of nodes are considered as variables in the algorithm of functionals minimization as will be further described in the following sections.

2.2 Lagrange's Method

Also the Lagrange method could be used in order to find the minimum or maximum of the function that defines the electric power and magnetic energy viewed from the perspective of an "objective-function" [14, 15]. If it is considered the objective-function

$$f(x,y) : U \to \Re, U \subset \Re^{2n} \tag{27}$$

of class C^1 and if it is assume that, between the scalars $x = (x_1, x_2, \ldots, x_n)$, and $y = (y_1, y_2, \ldots, y_n)$, exist m links

$$g_1(x,y) = 0, \ldots, g_m(x,y) = 0, g_i : U \to \Re, 1 \le i \le m \tag{28}$$

then in order to compute the minimum or maximum points $M(x_0, y_0)$ of function f the numerical method of Lagrange multipliers can be applied. Thereby the following function is defined

$$F = f(x,y) + \sum_{i=1}^{m} \lambda_i g_i(x,y) \tag{29}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are introduced as the Lagrange multipliers. In these conditions the extreme points $M(x_0, y_0)$ of function f represent the solutions of the non-linear system

$$\frac{\partial F}{\partial x_j} = 0, \frac{\partial F}{\partial y_k} = 0, g_i = 0, 1 \le j, k \le n, 1 \le i \le m \tag{30}$$

where the total number of unknown x, y, λ is $2n + m$. The sign of the square value (second order derivative) $d^2 f\big|_M$ decides the maximum or minimum nature of the extreme points $M(x_0, y_0)$. Practically a numerical procedure of the eigenvalues computation of f associate Hessian matrix i.e.

$$H = \left[\frac{\partial^2 f}{\partial x_j \partial y_k}(x_0, y_0) \right]_{1 \le j,k \le n} \tag{31}$$

yield information about the sign of square values: if the all the eigenvalues of Hessian matrix are positive or negative then the square value is positively or negatively defined, and implicitly the function f has a minimum or maximum at the

point $M(x_{o,}y_o)$. Because the matrix H is symmetrical, thus is has only real eigenvalues, consequently a critical point $M(x_{o,}y_o)$ it can't be a local extreme point for function f.

Let us examine a DC circuit (stationary state), with N nodes and L branches. For n_R variables (resistances), the consumed power (the objective function) $f(R,I)$: $\Re^{2n_R} \rightarrow \Re^+$, is defined as

$$f(R,I) = \sum_{i=1}^{n_R} R_i I_i^2 \tag{32}$$

where the resistances and the currents of branches $R = (R_1, R_2, \ldots, R_{n_R}), I = (I_1, I_2, \ldots, I_{n_R})$ are scalars and verify L Kirchhoff's current and voltage complete set of relations (links)

$$\begin{aligned} g_1 &= \sum_{l_k \in N_1} I_k = 0, \ldots, g_{N-1} = \sum_{l_k \in N_{N-1}} I_k = 0, \ldots, \\ g_N &= \sum_{l_k \in B_1} R_k I_k - E_k = 0, \ldots, g_L = \sum_{l_k \in B_{L-N+1}} R_k I_k - E_k = 0 \end{aligned} \tag{33}$$

There $g_j : \Re^{2n_R} \rightarrow \Re$, j = 1,..,L, and $B_{L\text{-}N+I} = L\text{-}N + 1$ are the independent loops of the circuit. In these assumptions it defines the function

$$F = f(R,I) + \sum_{j=1}^{L} \lambda_j g_j(R,I) = \sum_{i=1}^{n_R} R_i I_i^2 + \sum_{j=1}^{L} \lambda_j g_j(R,I) \tag{34}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_L$ are the unknown Lagrange's multipliers. The unique solution of the nonlinear system with $2n_R + L$ unknown

$$\frac{\partial F}{\partial R_i} = 0, \ \frac{\partial F}{\partial I_i} = 0, \ g_j = 0, i = 1, \ldots, n_R; j = 1, \ldots, L \tag{35}$$

coincides with an extreme point $M(x_{o,}y_o)$ of consumed power function (32) if the eigenvalues of the Hessian matrix are, in this point, real values and the same sign. If the sign is positively f has a maximum, otherwise the function f has a minimum.

This numerical procedure to determine the extreme point of function f is rather difficult because requires a lot of computing time and occupies a large memory space. This statement is explained by the fact that the method needs to calculate the m differentials of links relations, to solve a large nonlinear system, and to determine the square value of function f.

3 Equilibrium State of DC and AC Circuits and Minimum Power Flow. Examples

In classical theory of electric circuits of "content and co-content", the Hilbert space properties for solving the electromagnetic field and the theorem of the minimum power in the resistances for DC circuits are introduced [16–19].

Hereinafter the natural connection between the equilibrium state of DC and AC circuits and the minimum power flow is demonstrated in terms of appropriate power functionals defined for each category of circuit and the variational method is applied to examine the extreme point of functionals.

3.1 Principle of Minimum Consumed Power for DC Circuits and Variational Method

For a linear DC circuit that comprises N nodes and K branches, with general structure shown in Fig. 1, the K-dimensional vectors in $\Re^K$ of branch voltages $\Delta\mathbf{v}$ and currents $\mathbf{i}$, the voltages at the resistances terminals $\mathbf{u}$, and respectively the voltage sources $\mathbf{e}$ are defined as [20, 21].

$$\Delta\mathbf{v} = \begin{bmatrix} \Delta V_1 \\ \Delta V_2 \\ . \\ . \\ \Delta V_K \end{bmatrix}; \mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ . \\ . \\ I_K \end{bmatrix}; \mathbf{u} = \begin{bmatrix} U_1 \\ U_2 \\ . \\ . \\ U_K \end{bmatrix}; \mathbf{e} = \begin{bmatrix} \mathrm{E}_1 \\ \mathrm{E}_2 \\ . \\ . \\ \mathrm{E}_K \end{bmatrix} \tag{36}$$

The matrix relation

$$\Delta \mathrm{v} = \mathrm{C} \cdot \mathrm{V} \tag{37}$$

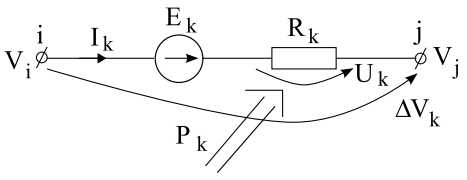


Fig. 1 General structure of a DC circuit branch

represents for k-branch the relation $\Delta V_k = V_i - V_j$, where V_i and V_j are the potentials of nodes i and j where the branch k is connected, $\mathrm{V} = \begin{bmatrix} V_1 \\ V_2 \\ . \\ . \\ V_{N-1} \end{bmatrix}$ is the reduced N-1, vector of the nodes' potential (there an arbitrary node is chosen with zero potential $V_N = 0$), and $\mathbf{C} = [c_{l,n}]$ is the reduced K x (N-1) branch-to-node incidence matrix. According to KVL it is obtain

$$\mathrm{i} = \mathbf{G}\mathrm{u} = \mathbf{G}(\Delta\,\mathrm{v} + \mathrm{e}) = \mathbf{G}(\mathrm{CV} + \mathrm{e}) \tag{38}$$

where $\mathbf{G} = diag(G_1, G_2, \ldots, G_K)$ is the branch conductance K x K- dimensional diagonal matrix**.** In the Hilbert space the power functional $\Im(\mathrm{V}) : \Re^{N-1} \rightarrow \Re$ is defined, by considering as variables N-1 potentials of nodes, as

$$\Im(\mathbf{V}) = \mathbf{u}^{\mathrm{T}}\mathbf{i} = P_{cons} \tag{39}$$

where the superindex T indicates the transposition. As is presented in Fig. 1 the same reference sense of the branch current and voltage the power functional (39) is equivalent with the definition of power consumed (P_{cons}) by all the resistances of the DC circuit. Taking into account relations (37) and (38) the power functional (39) can be expressed as

$$\begin{aligned} \Im(\mathbf{V}) &= \mathbf{u}^{\mathrm{T}}\mathbf{i} = (\Delta\mathbf{v} + \mathbf{e})^{\mathrm{T}}\mathbf{G}(\Delta\mathbf{v} + \mathbf{e}) = (\mathbf{CV} + \mathbf{e})^{\mathrm{T}}\mathbf{G}(\mathbf{CV} + \mathbf{e}) \\ &= \sum_{\substack{k=1,K \\ i,j=1,N-1 \\ i \neq j}} G_k(V_{k,i} - V_{k,j} + E_k)^2 \end{aligned} \tag{40}$$

From relation (40) results that always power functional is a quadratic form i.e.$\Im(\mathrm{V})\rangle 0$ and, consequently in the interval $(0, \infty)$, $\Im(\mathrm{V})$ has a minimum. Afterwards this minimum point corresponds to the solution of the system $\partial\Im/\partial\,\mathbf{V} = 0$. The first power functional derivative dependent on potential V_i can be written as

$$\frac{\partial \Im}{\partial V_i} = \frac{\partial}{\partial V_i}(\mathbf{CV} + \mathbf{e})^{\mathrm{T}}\mathbf{G}(\mathbf{CV} + \mathbf{e}) = 2\sum_{l_k \in n_i} c_{l_k,n_i} G_k(V_i - V_j + E_k) = 0 \tag{41}$$

where $k = 1, \ldots, K, i, j = 1, \ldots, N\text{-}1$, i ≠ j. The last equality in relation (41) represents even the formula of the nodal method (NM) expressed in node n_i

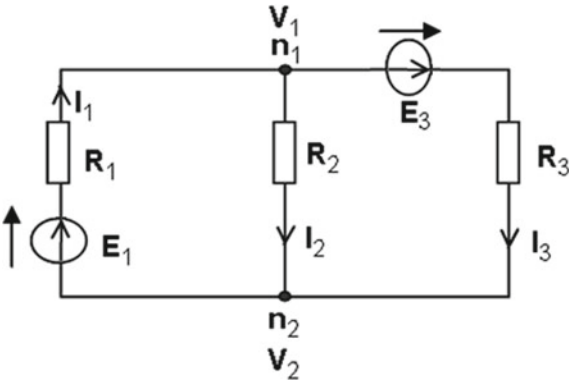


Fig. 2 DC circuit with three branches and two nodes

$$\sum_{l_k \in n_i} c_{l_k,n_i} G_k (V_i - V_j + E_k) = 0 \tag{42}$$

and, by using (38), we'll get

$$\sum_{l_k \in n_i} c_{l_k,n_i} I_k = 0, i = 1, \ldots, N-1 \tag{43}$$

that means even KCL for N-1 nodes of the circuit.

As a conclusion of the functional defined by (39) and from results obtained in (42) and (43) the Principle of Minimum Consumed Power for DC Circuits (PMP) can be stated in two equivalent forms: "In resistive DC circuits the condition of minimum consumed power in the resistances is consistent with the NM and KCL" or "In resistive DC circuit the branch currents and voltages have unique values such that the consumed power in all the resistances of the circuit is minimum" [22].

Example 1 For the DC circuit presented in Fig. 2, with $K = 3$, and $N = 2$, the structure is defined by the values $R_1 = 10\,\Omega$, $R_2 = 20\,\Omega$, $R_3 = 50\,\Omega$, $E_1 = 40\,V$ and $E_3 = 20$ V. The vectors of currents and voltages at the resistance terminals expressed in dependence with the potentials V_1 and V_2 of the nodes are written as. $\mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}$; $\mathrm{u} = \begin{bmatrix} \mathrm{V}_2 - \mathrm{V}_1 + \mathrm{E}_1 \\ \mathrm{V}_1 - \mathrm{V}_2 \\ \mathrm{V}_1 - \mathrm{V}_2 + \mathrm{E}_1 \end{bmatrix}$; $\mathbf{v} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$. For this DC circuit the branch-to-node incidence matrix is $\mathrm{C} = \begin{bmatrix} -1\,1\,1 \\ 1{-}1{-}1 \end{bmatrix}$ and the power functional (power consumed by resistances) constructed according to (40), is $\Im(\mathrm{V}_1, \mathrm{V}_2) = \mathrm{P}_{\mathrm{cons}}(\mathrm{V}_1, \mathrm{V}_2) = \mathrm{G}_1(\mathrm{V}_2 - \mathrm{V}_1 + \mathrm{E}_1)^2 + \mathrm{G}_2(\mathrm{V}_1 - \mathrm{V}_2)^2 + \mathrm{G}_3(\mathrm{V}_1 - \mathrm{V}_2 - \mathrm{E}_3)^2$. By imposed the minimum of power functional results $\frac{\partial \Im}{\partial \mathbf{v}} = \mathrm{C}\,\mathbf{i} = 0$ and thus imply KCL at each node n_1 and n_2:

$\frac{\partial \Im}{\partial V_1} = -G_1(V_2 - V_1 + E_1) + G_2(V_1 - V_2) + G_3(V_1 - V_2 - E_3) = 0$, involve KCL in n_1: $-I_1 + I_2 + I_3 = 0$;

$\frac{\partial \Im}{\partial V_2} = G_1(V_2 - V_1 + E_1) - G_2(V_1 - V_2) - G_3(V_1 - V_2 - E_3) = 0$, involve KCL in n_2: $I_1 - I_2 - I_3 = 0$.

Using the MAPLE software are obtained numerical values: $V_1 = 21.1765$ V, $V_2 = 0.0$ V and $\Im_{\min} = P_{cons,\min} = 91.7647$ W[23]. The 2-D dependence of power functionals of potentials of nodes is illustrated in Fig. 3. It is observe that the minimum point is (21.1765 V; 0.0 V; 91.7647 W).

For this DC circuit, a SCAP - Symbolic Circuit Analysis Program and MAPLE programs can be used to demonstrate that the functioning point of each resistance of the circuit does not represent the maximum absorbed power point [24, 25]. Based on the Thèvenin's theorem, the variations of the absorbed powers in each resistances of the circuit P_1, P_2, P_3 depending on the currents I_1, I_2, I_3 are calculated. The steps of the SCAP algorithm are the following:

(i) Calculation of branch currents (I_1, I_2, I_3) and voltages (U_{b1}, U_{b2}, U_{b3}), respectively the voltages at the resistances terminals in full symbolic form are:

$$I1 = \frac{E3\ R2 + E1\ R3 + E1\ R2}{R3\ R2 + R1\ R3 + R1\ R2} \quad I2 := \frac{E1\ R3 - 1.E3\ R1}{R3\ R2 + R1\ R3 + R1\ R2}$$

$$I3 = \frac{E1\ R2 + E3\ R2 + E3\ R1}{R3\ R2 + R1\ R3 + R1\ R2}$$

$$Ub1 = \frac{1.(E1\ R3 - 1.E3\ R1)R2}{R3\ R2 + R1\ R3 + R1\ R2} \quad Ub2 := \frac{1.(E1\ R3 - 1.E3\ R1)R2}{R3\ R2 + R1\ R3 + R1\ R2}$$

$$Ub3 = \frac{(E1\ R3 + 1.E3\ R1)R2}{R3\ R2 + R1\ R3 + R1\ R2}$$

$$UR1 = \frac{R1(E3\ R2 + E1\ R3 + E1\ R2)}{R3\ R2 + R1\ R3 + R1\ R2} \quad UR2 := \frac{(E1\ R3 - 1.E3\ R1)R2}{R3\ R2 + R1\ R3 + R1\ R2}$$

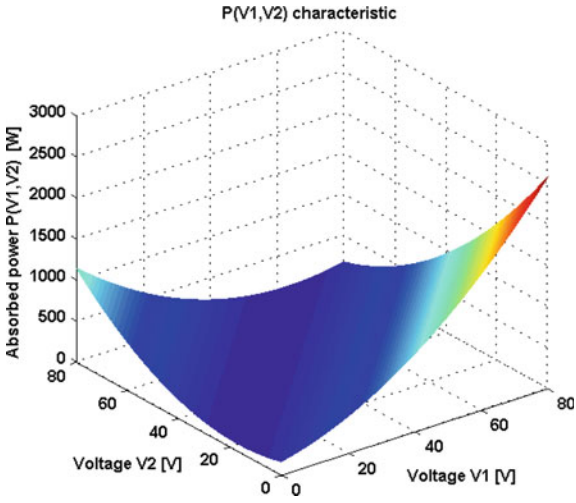


Fig. 3 2-D dependence of power functionals

$$\text{UR3} = \frac{\text{R3(E1 R2 + E3 R2 + E3 R2)}}{\text{R3 R2 + R1 R3 + R1 R2}}$$

(ii) Calculation of open voltages U_{Rk0}, $k = 1, 2, 3$ are computed by using the relation $U_{Rk0} = \lim_{R_k \to \infty} (U_{Rk})$ and it results:

$$\text{UR10} = \frac{\text{E3 R2 + E1 R3 + E1 R2}}{\text{R3 R2}} \quad \text{UR20} := \frac{\text{E1 R3} - \text{1.E3 R1}}{\text{R3 R1}}$$
$$\text{UR30} = \frac{\text{E1 R2 + E3 R2 + E3 R1}}{\text{R2 R1}}$$

(iii) Calculation of short-circuit currents $I_{k\text{sc}}$, $k = 1, 2, 3$, are computed by formula $I_{ksc} = I_k(R_k = 0)$, and results:

$$\text{I1sc} = \frac{\text{E3 R2 + E1 R3 + E1 R2}}{\text{R3 R2}} \quad \text{I2sc} = \frac{\text{E1 R3} - \text{1.E3 R1}}{\text{R1 R3}}$$
$$\text{I3sc} = \frac{\text{E1 R2 + E3 R2 + E3 R1}}{\text{R1 R2}}$$

(iv) Calculation of equivalent resistance at the nodes of each branch R_{0_k}, $k = 1$, 2, 3, is calculated as $R_{0\_k} = U_{Rk0}/I_{ksc}$. Then it is results:

$$\text{R0\_ 1} := \frac{\text{R2 R3}}{\text{R3 + R2}} \quad \text{R0\_ 2} := \frac{\text{R1 R3}}{\text{R3 + R1}} \quad \text{R0\_ 3} := \frac{\text{R2 R1}}{\text{R2 + R1}}$$

(v) Based on the Thèvenin's theorem the dependence between the consumed powers and the branch currents has the general formula

$$P_{Thev\_k} = (U_{Rk0} - R_{0\_k} I_k) I_k, k = 1, 2, 3$$

For for each branch the symbolic expressions are obtained:

$$\text{PThev\_ 1} := \frac{\text{(E1 R2 + E1 R3 + E3 R2 - R2 R3 I1)I1}}{\text{R2 + R3}}$$
$$\text{PThev\_ 2} := \frac{\text{(E1 R3} - \text{1.E3 R1} - \text{1.R1 R3 I1)I2}}{\text{R1 + R3}}$$
$$\text{PThev\_ 3} := \frac{\text{(E1 R2 + E3 R1 + E3 R2 - R2 R1 I3)I3}}{\text{R1 + R2}}$$

(vi) Used the theorem of maximum power transfer the powers delivered in the three resistors are:

$$\text{Pmax\_ 1} := \frac{\text{0.25000000(E1 R2 + E1 R3 + E3 R2)}^2}{\text{(R2 + R3)R2 R3}}$$

$$\text{Pmax\_ 2} := \frac{0.25000000(\text{E1 R3} - 1.\text{E3 R1})^2}{(\text{R1} + \text{R3})\text{R1 R3}}$$

$$\text{Pmax\_ 3} := \frac{0.25000000(\text{E1 R2} + \text{E3 R1} + \text{E3 R2})^2}{(\text{R1} + \text{R2})\text{R2 R1}}$$

(vii) According to the Thévenin theorem the current is calculated as $I_{Ik\_n} = U_{Rk0}/(R_{0\_k} + R_k)$ and for each branch have the expression:

$$\text{I1\_ n} := \frac{\text{E1 R2} + \text{E1 R3} + \text{E3 R3}}{\text{R1 R2} + \text{R1 R3} + \text{R2 R3}}$$

$$\text{I2\_ n} := \frac{\text{E1 R3} - 1.\text{E3 R1}}{\text{R1 R2} + \text{R1 R3} + \text{R2 R3}} \quad \text{I3\_ n} := \frac{\text{E1 R2} + \text{E3 R1} + \text{E3 R2}}{\text{R1 R2} + \text{R1 R3} + \text{R2 R3}}$$

For the above numerical values of circuit's parameters by using a MAPLE application it results:

I1_n: = 1.883 A *I2_n*: = 1.053 A *I3_n*: = 0.8253 A	*P1_n*: = 35.44 W *P2_n*: = 22.42 W *P3_n*: = 33.90 W	*I1sc*: 3.2000 A *I2sc*: = 3.600 A *I3sc*: = 7.000 A	*I1_max*: 1.600 A *I2_max*: 1.800 A *I3_max*: 3.500 A
PR_max: 36.78 W *PR_max*: 27.01 W *PR_max*: 81.66 W			

The power-current $P_R(I)$ characteristics are shown in Figs. 4, 5, and 6. It is remarkable to observe that the real consumed power in each resistance has a value lower than the maximum value.

$$P_{1\_n} = 35.44\ \text{W} < P_{R1\_\max} = 36.58\ \text{W};$$
$$P_{2\_n} = 22.42\ \text{W} < P_{R2\_\max} = 27.01\ \text{W};$$
$$P_{3\_n} = 33.90\ \text{W} < P_{R3\_\max} = 81.66\ \text{W}.$$

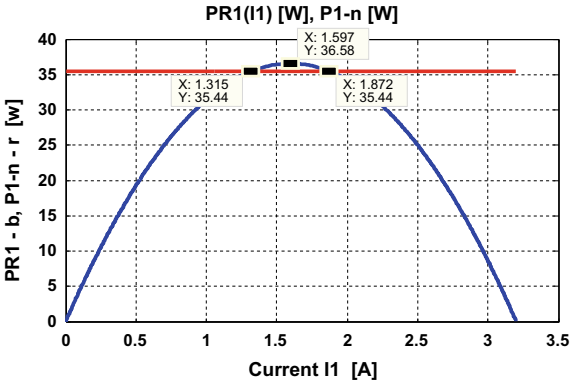


Fig. 4 The variation of P_{R1}

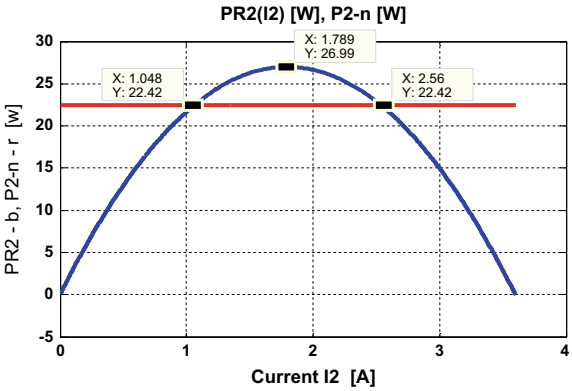


Fig. 5 The variation of P_{R2}

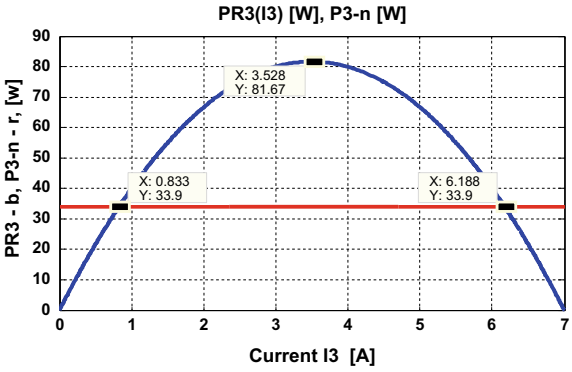


Fig. 6 The variation of P_{R3}

By used the ECAP - Electric Circuit Analysis Program software have been obtained the same values of voltage and current branches, and of consumed and generated power. The program is presented below:

Input file *ex1_cap24.nln.*
3
2
2 1 R1E1 r = 10,0 e = 40,0
1 2 R2 r = 20.0
1 2 R3E3 r = 50.0 e = 20.0
UNKNOWNS V1
EQUATION SYSTEM
+ (+G1 + G2 + G3)*V1 = + E1*G1-E3*G3
NODE POTENTIALS
V1 = 21.176471 V
V2 = 0 V
BRANCH CURENTS AND VOLTAGES
U1 = -21.176471 VI1 = 1.882353 A

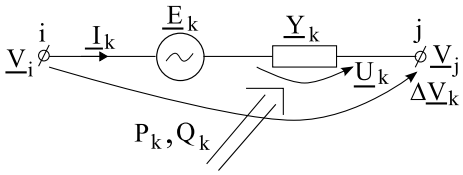


Fig. 7 Structure of AC branch

U2 = 21.176471 VI2 = 1.058824 A
U3 = 21.176471 VI3 = 0.823529 A
BALANCE OF THE POWERS
Generated power: = 91.764706 W
Consumed power: = 91.764706 W

3.2 Principle of Minimum Consumed Power for AC Circuits and Variational Method

A linear AC circuit in cvasi-stationary state, which includes N nodes and K branches whose general structure shown in Fig. 7 it contains in its structure passive RLC admittances and voltage sources. By analogy with the relation (36) the same quantities expressed as K-dimensional vectors in complex set C^K are described below [26, 27]

$$\Delta\underline{\mathbf{v}} = \begin{bmatrix} \Delta\underline{V}_1 \\ \Delta\underline{V}_2 \\ . \\ . \\ \Delta\underline{V}_K \end{bmatrix}; \underline{\mathbf{i}} = \begin{bmatrix} \underline{I}_1 \\ \underline{I}_2 \\ . \\ . \\ \underline{I}_K \end{bmatrix}; \underline{\mathbf{u}} = \begin{bmatrix} \underline{U}_1 \\ \underline{U}_2 \\ . \\ . \\ \underline{U}_K \end{bmatrix}; \underline{\mathbf{e}} = \begin{bmatrix} \underline{\mathrm{E}}_1 \\ \underline{\mathrm{E}}_2 \\ . \\ . \\ \underline{\mathrm{E}}_K \end{bmatrix} \tag{44}$$

where the k-branch complex voltage is $\Delta\underline{V}_k = \underline{V}_i - \underline{V}_j$.

Based on the interconnection properties of AC circuits branches the following matrix relations are true

$$\Delta\underline{\mathrm{v}} = \mathrm{C} \cdot \underline{\mathrm{V}} \tag{45}$$

and KVL

$$\underline{\mathrm{i}} = \underline{\mathrm{Y}}\underline{\mathrm{u}} = \underline{\mathrm{Y}}(\Delta\underline{\mathrm{v}} + \underline{\mathrm{e}}) = \underline{\mathrm{Y}}(\mathrm{C}\underline{\mathrm{V}} + \underline{\mathrm{e}}) \tag{46}$$

where the diagonal admittance matrix $\mathbf{Y} = diag\left(\underline{Y}_1, \underline{Y}_2, \ldots, \underline{Y}_K\right)$, each $\underline{\mathrm{Y}}_k = G_k - jB_k$, $k = 1,K$. Fort inductive branch the sign of susceptance B_k it is considered

positive respectively for capacitive branch it is negative. Under these conditions two functionals are defined: the active power $\Im_P : \Re^{2(N-1)} \rightarrow \Re$, and the reactive power $\Im_Q$: $\Re^{2(N-1)} \rightarrow \Re$, expressed as

$$\Im_P = \mathrm{Re}[\underline{\mathbf{u}}^{\mathrm{T}}\underline{\mathbf{i}}^*] = P_{cons} \tag{47}$$

$$\Im_Q = \mathrm{Im}[\underline{\mathbf{u}}^{\mathrm{T}}\underline{\mathbf{i}}^*] = Q_{cons(gen)} \tag{48}$$

where the superindex * denotes the conjugate complex operator. Taking into account the reference sense adopted for the AC circuit branch of Fig. 7 and the definition of complex power $\underline{\mathrm{S}} = \underline{\mathbf{u}}^T\underline{\mathbf{i}}^* = P + jQ$, then the power functional (47) represents the *active power consummated by all the resistances of the circuit*, while the power functional (49) represents the *reactive power consummated (*or *generated), by all the reactive elements of the circuit*.

Let's suppose that the N-1 potentials of nodes are variables and the voltage sources are constant, and for node i respectively for k-branch are expressed as

$$\underline{V}_i = \mathrm{Re}[\underline{V}_i] + j\mathrm{Im}[\underline{V}_i] = x_i + jy_i,\ i = 1, \ldots, N-1 \tag{49}$$

$$\underline{\mathrm{E}}_k = \mathrm{Re}[\underline{\mathrm{E}}_k] + j\mathrm{Im}[\underline{\mathrm{E}}_k] = \mathrm{a}_k + j\mathrm{b}_k,\ k = 1, \ldots, K \tag{50}$$

Then the functionals (47) and (48) can be expressed as

$$\begin{aligned}\Im_P(x, y) &= \mathrm{Re}[\underline{\mathbf{u}}^{\mathrm{T}}\underline{\mathbf{i}}^*] = \mathrm{Re}[(\Delta\underline{\mathbf{v}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\Delta\underline{\mathbf{v}} + \underline{\mathbf{e}})^*] \\ &= \mathrm{Re}[(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^*] \\ &= \sum_{\substack{k=1,K \\ i,j=1,N-1 \\ i\neq j}} G_k[(x_i - x_j + \mathrm{a}_k)^2 + (y_i - y_j + \mathrm{b}_k)^2]\end{aligned} \tag{51}$$

$$\begin{aligned}\Im_Q(x, y) &= \mathrm{Im}[\underline{\mathbf{u}}^{\mathrm{T}}\underline{\mathbf{i}}^*] = \mathrm{Im}[(\Delta\underline{\mathbf{v}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\Delta\underline{\mathbf{v}} + \underline{\mathbf{e}})^*] \\ &= \mathrm{Im}[(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^*] \\ &= \sum_{\substack{k=1,K \\ i,j=1,N-1 \\ i\neq j}} B_k[(x_i - x_j + \mathrm{a}_k)^2 + (y_i - y_j + \mathrm{b}_k)^2].\end{aligned} \tag{52}$$

From relation (51) results that because always $G_k\rangle 0$, then $\Im_P$ is always strictly positive (quadratic form)$\Im_P(x, y))0$. Therefore the active power functional $\Im_P$ has a minimum in the definition set $(0, \infty)$, and, consequently, the resistances of the circuit consume minimum active power. The minimum point of active power functional is fixed by the fulfillment of the conditions $\partial\Im_P/\partial\ x_i = 0$ and $\partial\Im_P/\partial\ y_i = 0$, for $i = 1, \ldots N-1$, which can be expressed synthetically in relation to $\underline{V}_i$ in the following

form

$$\frac{\partial \Im_P}{\partial x_i} = \frac{\partial}{\partial x_i}\mathrm{Re}[(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^*] = 2\sum_{l_k \in n_i} c_{l_k,n_i} G_k(x_i - x_j + \mathrm{a}_k) = 0 \tag{53}$$

$$\frac{\partial \Im_P}{\partial y_i} = \frac{\partial}{\partial y_i}\mathrm{Re}[(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^*] = 2\sum_{l_k \in n_i} c_{l_k,n_i} G_k(y_i - y_j + \mathrm{b}_k) = 0 \tag{54}$$

for $k = 1, \ldots, K$, $i, j = 1, \ldots, N-1$, $i \neq j$, and where $c_{\mathrm{lk,ni}}$ are the coefficients of reduced branch-to-node incidence matrix.

Afterwards analyzing relation (52) the value of $\Im_{\mathrm{Q}}$ might be: (i) $\Im_{\mathrm{Q}} > 0$ ($\Im_{\mathrm{Q}}$ is a quadratic form) if all the branches of AC circuit are inductive, then the extreme point of $\Im_{\mathrm{Q}}$ is a minimum, and, consequently, the reactive power consumed is minimum; (ii) $\Im_{\mathrm{Q}} < 0$ ($-\Im_{\mathrm{Q}}$ is a quadratic form) if all the branches of AC circuit are capacitive. In this case, by multiplication with (-1), the sign of the reactive power functional has changed into a positive and can be formulated as the reactive power produced (generated) has a minimum; (iii) $\Im_{\mathrm{Q}} = 0$ represents the particular case of resonance condition, in which the AC circuit provides a null contribution to the consumed or generated reactive power.

The first derivative of reactive power functional in terms on real and imaginary part of potential $\underline{V}_i$ can be written as

$$\frac{\partial \Im_Q}{\partial x_i} = \frac{\partial}{\partial x_i}\mathrm{Im}[(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^*] = 2\sum_{l_k \in n_i} c_{l_k,n_i} B_k(x_i - x_j + \mathrm{a}_k) = 0 \tag{55}$$

$$\frac{\partial \Im_Q}{\partial y_i} = \frac{\partial}{\partial y_i}\mathrm{Im}[(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^{\mathrm{T}}\underline{\mathbf{Y}}^*(\mathbf{C}\underline{\mathbf{V}} + \underline{\mathbf{e}})^*] = 2\sum_{l_k \in n_i} c_{l_k,n_i} B_k(y_i - y_j + \mathrm{b}_k) = 0 \tag{56}$$

for $k = 1, \ldots, K$, $i, j = 1, \ldots, N$-1, $i \neq j$. The minimum of active and reactive power functionals results from the system of $4(N-1)$ equations formed by relations (53), (54), (55), and (56) as follows

$$\frac{\partial \Im_P}{\partial x_i} = \sum_{l_k \in n_i} c_{l_k,n_i} G_k(x_i - x_j + \mathrm{a}_k) = 0; \quad \frac{\partial \Im_P}{\partial y_i} = \sum_{l_k \in n_i} c_{l_k,n_i} G_k(y_i - y_j + \mathrm{b}_k) = 0$$
$$\frac{\partial \Im_Q}{\partial x_i} = \sum_{l_k \in n_i} c_{l_k,n_i} B_k(x_i - x_j + \mathrm{a}_k) = 0; \quad \frac{\partial \Im_Q}{\partial y_i} = \sum_{l_k \in n_i} c_{l_k,n_i} B_k(y_i - y_j + \mathrm{b}_k) = 0 \tag{57}$$

where $k = 1, \ldots, K$, $i, j = 1, \ldots, N-1$, and $i \neq j$. If in the equations system (57) the relations $\partial \Im_P / \partial\, y_i = 0$ and $\partial\, \Im_Q / \partial\, x_i = 0$ multiplied by j respectively –j are added up, then results

$$\sum_{l_k \in n_i} c_{l_k, n_i} \underline{Y}_k (\underline{V}_i - \underline{V}_j + \underline{E}_k) = 0 \tag{58}$$

that means, for $i = 1, \ldots, N-1$, the equations of nodal method (NM) for all the N-1 circuit' nodes. By using (46) relation (58) becomes

$$\sum_{l_k \in n_i} c_{l_k, n_i} \underline{I}_k = 0 \tag{59}$$

that represents the KCL equations.

Consequently, the Principle of Minimum Active and Reactive Power (PMARP) can be stated in two equivalent forms: "In linear AC circuits the conditions of minimum active consumed power and minimum reactive consumed (or produced) power are consistent with the NM and KCL" or "In linear AC circuit the branch currents and voltages have unique values such that the consumed active power and the consumed (or produced) reactive power in all the admittances of the circuit is minimum" [22].

Example 2 For the AC circuit shown in Fig. 8, with $K = 3$, and $N = 4$, the branches contain passive linear elements (resistor, capacitor and inductance) and a voltage source. The circuit parameters have the numeric values: $\underline{E}_1 = 100.0$; $R_1 = 10.0\ \Omega$; $R_2 = 20.0\ \Omega$; $C_2 = 1.0\text{e}{-}04$ F; $\mathrm{L}_3 = 2.0\text{e}{-}04$ H and $\omega = 314.0$ rad/s. Let be the AC circuit presented in Fig. 8 has $K = 3$, and $N = 4$. The real and imaginary parts of potentials of nodes are considered as variables, and can be expressed as $\underline{V}_1 = x_1 + jy_1$, $\underline{V}_2 = x_2 + jy_2$, respectively $\underline{V}_3 = x_3 + jy_3$.

The branches currents, the admittances' voltages, and the potential of nodes are described by the vectors:

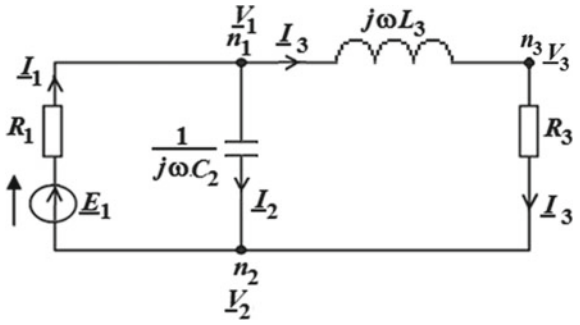


Fig. 8 The AC circuit

$$\underline{\mathbf{I}} = \begin{bmatrix} \underline{I}_1 \\ \underline{I}_2 \\ \underline{I}_3 \\ \underline{I}_4 \end{bmatrix}, \underline{\mathbf{u}} = \begin{bmatrix} \underline{V}_2 - \underline{V}_1 + \underline{E}_1 \\ \underline{V}_1 - \underline{V}_2 \\ \underline{V}_1 - \underline{V}_3 \\ \underline{V}_3 - \underline{V}_2 \end{bmatrix}; \underline{\mathbf{V}} = \begin{bmatrix} x_1 + jy_1 \\ x_2 + jy_2 \\ x_3 + jy_3 \end{bmatrix}.$$

For this AC circuit the matrices C and $\underline{\mathbf{Y}}$ are expressed as:

$$C = \begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & -1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \underline{\mathbf{Y}} = \begin{bmatrix} G_1 & 0 & 0 & 0 \\ 0 & j\omega C_2 & 0 & 0 \\ 0 & 0 & -\frac{j}{\omega L_3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

By using relations (45) and (46) and assuming that $\underline{E}_1 = a + jb$, where a, b are real constants, then the complex power functional attached to overall complex consumed power by the passive elements of circuit is defines as

$$\begin{aligned} \underline{\Im}_S = \underline{S}_{cons} = \underline{\mathbf{u}}^T \cdot \underline{\mathbf{I}}^* = {} & \frac{1}{R_1}(-x_1 + x_2 + a + j(-y_1 + y_2 + b))^2 \\ & + j\omega C_3(x_1 - x_2 + j(y_1 - y_2))^2 + \frac{j}{\omega L_3}(x_1 - x_3 + j(y_1 - y_3))^2 \\ & + \frac{1}{R_3}(-x_2 + x_3 + j(-y_2 + y_3))^2. \end{aligned}$$

The active power consumed by the resistors is expressed by the functional below:

$$\begin{aligned} Fr := {} & -(\omega^2 C2R1L3R3x1y1 - 2\omega^2 C2R1L3R3x1y2 - 2C2R1L3R3x2y1 \\ & + 2\omega^2 C2R1L3R3x2y2 + 2\omega L3R3x2x1 - 2\omega L3R3x2a \\ & + 2\omega L3R3x1a + \omega L3R3y2y1 + 2\omega L3R3y2b - 2\omega L3R3y1b \\ & + 2R1\omega L3x3x2 + 2R1\omega L3y3y2 + \omega L3R3b^2 - R1\omega L3x3^2 \\ & - R1\omega L3x2^2 + R1\omega L3y3^2 + R1\omega L3y2^2 - 2R1R3x1y1 + 2R1R3x1y3 \\ & + 2R1R3x3y1 - 2R1R3x3y3 - \omega L3R3x2^2 - \omega L3R3x1^2 - \omega L3R3a^2 \\ & + \omega L3R3y2^2 + \omega L3R3y1^2)/(2R1\omega L3R3) \end{aligned}$$

Afterwards the reactive power consumed (or generated) is expressed by the functional below:

$$\begin{aligned} Fi := {} & -(\omega^2 C2R1L3R3x1^2 + \omega^2 C2R1L3R3x2^2 - \omega^2 2C2R1L3R3y1^2 \\ & - \omega^2 2C2R1L3R3y2^2 - 2\omega^2 C2R1L3R3x2x2 + 2\omega^2 C2R1L3R3y2y1 \\ & - 2\omega L3R3ay1 - 2\omega L3R3x1b + \omega L3R3x2b - 2\omega L3R3x1y2 \\ & + 2\omega L3R3x2y2 - 2\omega L3y3x2y1 + 2\omega L3R3ay2 + 2\omega L3R3x1y1 \\ & + 2\omega L3R3ab + 2R1\omega L3x3y3 - 2R1\omega L3x3y2 - 2R1\omega L3x2y3 \end{aligned}$$

$$+ 2R1\omega L3x2y2 + R1R3y3^2 - R1R3x1^2 + R1R3y1^2 - R1R3x3^2$$
$$+ 2R1R3x1x3 - 2R1R3y1y3)/(2R1\omega L3R3)$$

To determine the minimum points of the active and reactive power functionals the solutions of the system which contains $4 \times 3 = 12$ equations are computed, and results:

$$\frac{\partial F_R}{\partial x_1} = 0, \frac{\partial F_R}{\partial y_1} = 0, \frac{\partial F_R}{\partial x_2} = 0, \frac{\partial F_R}{\partial y_2} = 0, \frac{\partial F_R}{\partial x_3} = 0, \frac{\partial F_R}{\partial y_3} = 0,$$
$$\frac{\partial F_i}{\partial x_1} = 0, \frac{\partial F_i}{\partial y_1} = 0, \frac{\partial F_i}{\partial x_2} = 0, \frac{\partial F_i}{\partial y_2} = 0, \frac{\partial F_i}{\partial x_3} = 0, \frac{\partial F_i}{\partial y_3} = 0$$

which, for the numerical values of the AC circuit shown in Fig. 8 have the solution:

Soluation :=
{x1 = 0.021099770, x2 = 63.860285 + x3, x3 = 0., y1 = 0.10026065 + y3, y2 = 13.439344, y3 = 0.}

Then: $\underline{V}_{1\_min} = 0.0211 + j0.10026$ and $\underline{V}_{2\_min} = -63.8603 + j13.4393$, and for these values of potentials of nodes the minimum active and reactive power are defined as consumed.

4 Equilibrium State of Linear Magnetic and Minimum Energy Flow. Example

The basic strategy of variational method applied to linear electric DC and AC circuits exposed above can be extended to linear magnetic circuit in quasi-stationary state. This is possible because between magnetic and electric circuits exists a well-known analogy, which makes the construction of the functional and the analysis of its minimum to be done in a similar way.

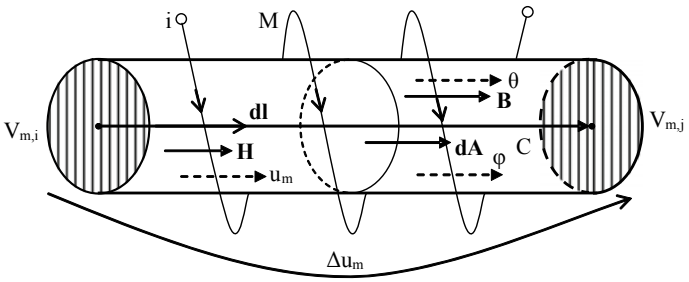


Fig. 9 Structure of magnetic branch

For magnetic circuit branch presented in Fig. 9 a magnetic field generator is considered with M turns crossed by the current i. The magnetic voltage Δu_m between the terminals i and j can be calculated by using the Ampère's theorem [11, 28].

$$\Delta u_m = R_m \varphi - \theta = \varphi / \Lambda - \theta \tag{60}$$

where φ is the fascicular flux through circuit branch; $R_m = l/\mu A$ is the reluctance of the magnetic circuit branch depending on the magnetic material properties: l the length, A the cross section area and μ the permeability of the linear and homogenous medium, then always $R_m > 0$; $\Lambda = 1/R_m$ is the permeance defined as the inverse of the reluctance; $\theta = Mi$ is the magnetomotive force. The relation (60) is also called Kirchhoff's Magnetic Voltage Law (KMVL) by analogy with KVL. Thereby the analogous magnetic circuit branch defined by Eq. (60) is shown in Fig. 10.

Likewise to the matrices defined in the two previous sections for a linear magnetic circuit in cvasi-stationary state with K branches and N nodes, the branch magnetic voltages of branches $\Delta u_m = V_{m,i} - V_{m,j}$ defined as the difference between the magnetic potentials of branches' nodes, the fascicular fluxes, the magnetic voltages of reluctances, and the magnetomotive forces are defined as K-dimensional vectors in real set [29]

$$\Delta \mathbf{u}_m = \begin{bmatrix} \Delta u_{m,1} \\ \Delta_{m,2} \\ . \\ . \\ \Delta u_{m,K} \end{bmatrix}; \quad \varphi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ . \\ . \\ \varphi_K \end{bmatrix}; \mathbf{u}_m = \begin{bmatrix} u_{m,1} \\ u_{m,2} \\ . \\ . \\ u_{m,K} \end{bmatrix}; \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ . \\ . \\ \theta_K \end{bmatrix} \tag{61}$$

By using the matrix $\mathbf{C} = [c_{l,n}]$ that is the K x $(N-1)$-dimensional reduced branch-to-node incidence matrix and if an arbitrary magnetic potential of circuit's nodes is chosen as null $V_{m,N} = 0$, then it can be written that $\Delta \mathbf{u}_m = \mathbf{C} \cdot \mathbf{V}_m$, where $\mathbf{V_m}$ is the vector of the magnetic potential of circuit nodes, with $\dim \mathbf{V_m} = N-1$.

The matrix equation expressing dependence between the fascicular fluxes of branches and magnetic potentials of nodes has the following expression

$$\varphi = \boldsymbol{\Lambda} \mathbf{u}_m = \boldsymbol{\Lambda}(\Delta \mathbf{u}_m + \boldsymbol{\theta}) = \boldsymbol{\Lambda}(\mathbf{C}\mathbf{V}_m + \boldsymbol{\theta}) \tag{62}$$

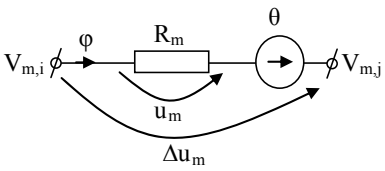


Fig. 10 The analogous magnetic circuit branch

where the branch permeance matrix is $\mathbf{\Lambda} = diag(\Lambda_1, \Lambda_2, \ldots, \Lambda_K)$, $\mathbf{V_m}$ is the $(N-1)$-vector of the magnetic potential of circuit nodes ($V_{m,N} = 0$), and Ohm's Law for magnetic circuits is $\Delta \mathbf{u}_m = \mathbf{C} \cdot \mathbf{V}_m$.

The functional $\Im(\mathbf{V}_m) : \Re^{N-1} \rightarrow \Re$ expressed as [30]

$$\Im(\mathbf{V}_m) = \frac{1}{2}\mathbf{u}_m^{\mathrm{T}}\varphi \tag{63}$$

represents the magnetic energy consumed by the reluctances of the magnetic circuit in cvasi-stationary state. If the -1 magnetic potentials of circuit nodes are considered as variables, the energetic functional can be expressed as

$$\Im(\mathbf{V}_m) = \frac{1}{2}\mathbf{u}_m^{\mathrm{T}}\varphi = \frac{1}{2}(\Delta\mathbf{u}_m + \boldsymbol{\theta})^{\mathrm{T}}\mathbf{\Lambda}(\Delta\mathbf{u}_m + \boldsymbol{\theta}) = \frac{1}{2}\sum_{\substack{k=1,K \\ i,j=1,N-1 \\ i\neq j}} \Lambda_k(V_{m,i} - V_{m,j} + \theta_k)^2 \tag{64}$$

From (64) results that $\Im(\mathbf{V}_m)\rangle 0$ (i.e.$\Im(\mathbf{V}_m)$ is a quadratic form for any value of the magnetic potentials of the nodes $\mathbf{V}_m$). As consequently the extreme point of the energetic functional $\Im(\mathbf{V}_m)$ is obtained by imposing the condition $\partial\Im/\partial\,\mathbf{V}_m = 0$. It is results

$$\frac{\partial\Im}{\partial V_{m,i}} = \frac{\partial}{\partial V_{m,i}}\frac{1}{2}(\mathbf{C}\mathbf{V}_m + \boldsymbol{\theta})^{\mathrm{T}}\mathbf{\Lambda}(\mathbf{C}\mathbf{V}_m + \boldsymbol{\theta}) = \frac{\partial}{\partial V_{m,i}}\frac{1}{2}\sum_{\substack{k=1,K \\ i,j=1,N-1 \\ i\neq j}} \Lambda_k(V_{m,i} - V_{m,j} + \theta_k)^2$$

$$= \sum_{l_k \in n_i} c_{l_k,n_i}\Lambda_k(V_{m,i} - V_{m,j} + \theta_k) = 0 \tag{65}$$

where $k = 1, \ldots, K$, $i, j = 1, \ldots, N-1$, $i \neq j$, and $V_{m,N} = 0$. The last equality of (65) represents exactly the equations of NM for magnetic circuit's i.e.

$$\sum_{l_k \in n_i} c_{l_k,n_i}\Lambda_k(V_{m,i} - V_{m,j} + \theta_k) = 0 \tag{66}$$

Similarly to the electric circuits, if it is rewrite (66) by using (62), we'll get

$$\sum_{l_k \in n_i} c_{l_k,n_i}\varphi_k = 0, i = 1, \ldots, N-1 \tag{67}$$

so these equations mean the Kirchhoff Magnetic Flux Law (KMFL).

Considering the definition (64) and the relations (66) and (67) the Principle of Minimum Consumed Energy for Magnetic Circuits (PMEM) in cvasi-stationary state can be stated in two equivalent forms: "In linear magnetic circuit the circumstance of

minimum consumed energy in the branches reluctances (permeances) is equivalent with the NM and KMFL" or "In linear magnetic circuits the branch fascicular fluxes and magnetic voltages have unique values such that the consumed energy in the reluctances (permeances) is minimum" [22].

Example 3 Let us consider the electrical transformer excited by currents i_1 and i_2, which is presented in Fig. 11a. It is assumed that the transversal area A it is the same everywhere, the ferromagnetic material is linear with the relative permeability μ_r, the two excitation coils has M_1 and M_2 turns, and the width of the air-gap is δ. From the geometrical dimensions indicated in Fig. 11a the permeances Λ_1, Λ_2 and Λ_3 of the magnetic circuit can be calculated. The analogous magnetic circuit with $K = 3$ branches and $N = 2$ nodes of this electrical transformer in cvasi-stationary state is illustrated in Fig. 11b.

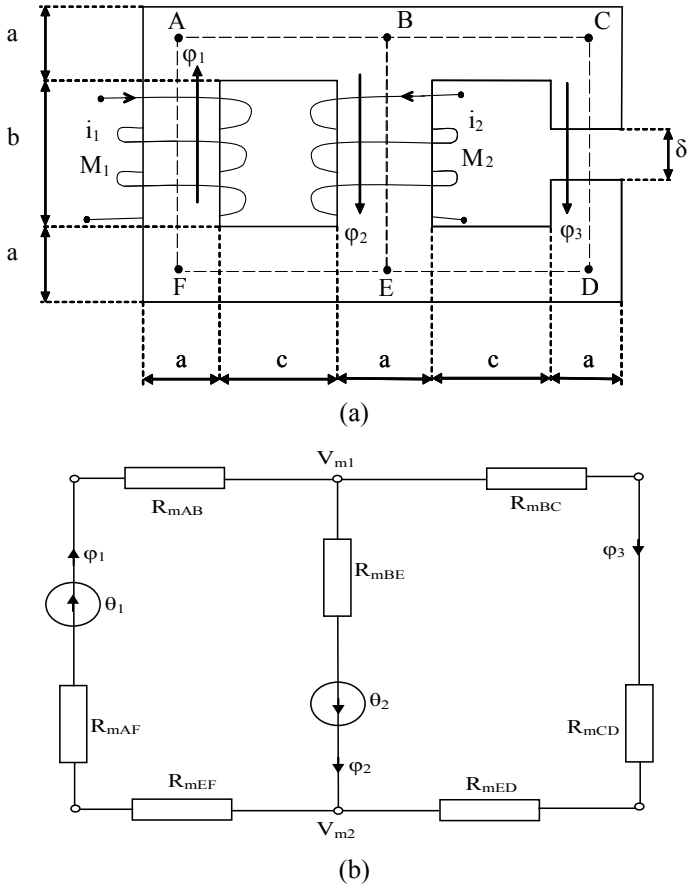


Fig. 11 **a** The electrical transformer excited by two currents; **b** Analogous magnetic circuit of the electrical transformer

By considering the magnetic potential $V_{m,1}$ as variable and $V_{m,2} = 0$, then the energetic functional equivalent to the magnetic energy W_{m} consumed by the circuit's permeances becomes

$$\Im(V_{m,1}) = W_m(V_{m,1}) = \frac{1}{2}\Big[\Lambda_1\big(-V_{m,1}+\theta_1\big)^2 + \Lambda_2\big(V_{m,1}+\theta_2\big)^2 + \Lambda_3\big(V_{m,1}\big)^2\Big]$$

where the magnetomotive forces are $\theta_1 = M_1 i_1$ respectively $\theta_2 = M_2 i_2$.

By imposing the conditions (65) is obtained

$$\begin{aligned}\frac{\partial \Im}{\partial V_{m,1}} &= \frac{\partial}{\partial V_{m,1}}\frac{1}{2}\Big[\Lambda_1\big(-V_{m,1}+\theta_1\big)^2 + \Lambda_2\big(V_{m,1}+\theta_2\big)^2 + \Lambda_3\big(V_{m,1}\big)^2\Big] \\ &= -\Lambda_1\big(-V_{m,1}+\theta_1\big) + \Lambda_2\big(V_{m,1}+\theta_2\big) + \Lambda_3\big(V_{m,1}\big) = 0\end{aligned}$$

where the last equality represents the NM expressed in node 1 of the analogous magnetic circuit. Afterwards results KMFL in node 1:

$-\varphi_1 + \varphi_2 + \varphi_3 = 0.$

5 Conclusion

The variational properties are applied to compute the extreme points of power and energy functionals for electric and magnetic circuits in stationary and cvasi-stationary state. Advanced numerical methods prove that the power and energy functionals have a minimum point thus the consumed power and energy by passive elements of electric and magnetic circuits is minimum. The matrix expressions of minimum conditions for power and energy functionals together with KVL for electric circuit and KMVL for magnetic circuit imply, concurrent, the equations of NM and of KCL respectively KMFL. Also the electric and magnetic potential of nodes were chosen as variables, because only the currents, magnetic fluxes, voltages and magnetic voltages of the branches are uniquely determined. The presented examples demonstrate each of theoretical principles PMP, PMARP and PMEM enunciated.

References

1. Dewar RC (2003) Information theory explanation of the fluctuation theorem, maximum entropy production, and self-organization criticality in non-equilibrium stationary states. J Phys A36:631–641
2. Ozawa H, Ohmura A, Lorenz RD, Pujol T (2003) The second law of thermodynamics and the global climate system—a review of the maximum entropy production principle. Rev Geophys 41:1018

3. Rodriguez-Iturbe I, Rinaldo A (1994) Fractal river basins: chance and self-organization. Cambridge University Press, New York, USA
4. Lanczos C (1970) The variational principles of mechanics, 4th edn. Dover Publication, New York, USA
5. Kleidon A, Fraedrich K, Kunz T, Lunkeit F (2003) The atmospheric circulation and states of maximum entropy production. Geophys Res Lett 30:2223
6. Kleidon A, Lorenz RD (2004) Non-equilibrium thermodynamics and the production of entropy: life, Earth, and beyond; Heidelberg. Springer Verlag, GE
7. Kleidon A (2002) Testing the effect of life on earth's functioning: how gaian is the earth system? Clim Change 52:383–389
8. Bejan A (2000) Shape and structure, from engineering to nature. Cambridge University Press, Cambridge, UK
9. Chua LO (1969) Introduction to nonlinear network theory. Part I. Mc Graw Hill Book Company, New York, USA
10. Vasiliu M, Hantila IF (2006) Electromagnetics, bucharest. Editura Electra, Romania
11. Clerk Maxwell J (1954) A treatise on electricity and magnetism, 3rd ed., vol 1. over Publication, Inc. New York
12. Clemente-Gallardo J, Scherpen JMA (2003) Relating lagrangian and hamiltonian formalism of LC circuits. IEEE Trans Circuits Syst 50(10):1359–1363
13. Desoer CA, Kuh ES (1969) Basic circuit theory. Mc Graw-Hill, New York
14. Kalman D (2009) Leveling with Lagrange: an alternate view of constrained optimization. Math Mag 82(3):186–196
15. Andrei H, Spinei F, Andrei P, Rohde U, Silaghi M, Silaghi H (2009) Evaluation of hilbert space techniques and lagrange's method for the analysis of dissipated power in DC circuits. In: Proceedings of IEEE -ECCTD'09, Antalya, pp 862–865
16. Penfield P, Spence R, Duinker S (1970) Tellegen's theorem and electrical networks. Research monograph no. 58. Massachusetts. M.I.T. Press, USA
17. Ionescu V (1958) Hilbert space applications to distorted waveform analysis (in Romanian: Aplicatii ale spatiilor Hilbert la studiul regimului deformant). Electrotehnica 6:280–286
18. Mocanu CI (1979) Electric circuits theory (in Romanian: teoria circuitelor electrice), Bucharest. Editura Didactica si Pedagogica, Romania
19. Fireteanu V, Popa M, Tudorache T, Levacher L, Paya B, Neau Y (2004) Maximum of energetic efficiency in induction through-heating processes. In: Proceedings of HES symposium, Padua, Italy, pp 80–86
20. Stern TE (1996) On the equations of nonlinear networks. IEEE Trans Circuits Theory CT-13(1), 74–81
21. Andrei H, Andrei PC (2013) Matrix formulations of minimum dissipated power principles and nodal method of circuits analysis, IEEE-advanced topics in electrical engineering—ATEE, 23–25 May. Bucharest, Romania
22. Andrei H, Chicco G, Spinei F (2011) Minimum dissipated power and energy—two general principles of the linear electric and magnetic circuits in quasi-stationary regime, pp 130–205, chapter 5 of the book Advances in energy research: distributed generations systems integrating renewable energy resources, Nova Science Publishers, New York, 2011
23. https://www.maplesoft.com
24. Iordache M (2015) Symbolic, numeric—symbolic and numeric simulation of analog circuits—user guides. MATRIX ROM, Bucharest
25. Iordache M, Dumitriu L (2014) Computer-aided simulation of analoque circuits—algorithms and computational techniques, vol I and II. Editura POLITEHNICA Press, Bucharest
26. Bellevitch V (1968) Classical network theory. Holden-Day
27. Andrei H, Andrei PC, Oprea G, Botea B (2014) Basic equations of linear electric and magnetic circuits in quasi-stationary state based on principle of minimum absorbed power and energy. In: Proceedings IEEE-ISFEE. Bucharest, 28–29 Nov 2014, pp 1–6

28. Andrei H, Andrei PC, Mantescu G (2014) Matrix formulation of minimum absorbed energy principle and nodal method of magnetic circuits analysis. In: Procedings of IEEE-14th international conference on optimization of electrical and electronic equipment–OPTIM. Brasov, 22–24 May 2014
29. Andrei H, Spinei F (2007) An extension of the minimum energy principle in stationary regime for electric and magnetic circuits. Rev Roum Sci Tech Ser Electrotech Energ, Tome 52:419–427
30. Lafontaine J (2015) An introduction to differential manifolds. Springer

Numerical Methods for Analysis of Energy Consumption in Drying Process of Wood

Livia Bandici, Simina Coman, and Teodor Leuca

Abstract This chapter presents general aspects regarding the electromagnetic field in radio frequency and microwaves, the thermal field, mass problems in radio frequency drying, and the authors' contributions to the numerical analysis of high frequency drying. The material used in numerical simulations is wood. Wood temperature control is very important due to cracks and loss of mechanical properties during drying. The properties of dielectric materials are very important when studying the interaction that takes place between the electromagnetic field energy and the material. The need to process and obtain products that meet market demands has led to the development of process modeling software that are aimed at simulating processes and physical phenomena as precisely and realistically as possible. In the research center (Center for Research and Technological Engineering in Electromagnetic Energy Conversion—CCITCEE) a Fortran software complex that couples electric, thermal, and mass and motion problems called FEM-BEM.3D-RF was developed. In the present study the drying process of wood was numerical simulated using FEM-BEM.3D-RF in radio frequency field and Comsol Multiphysics in microwave field.

Keywords Wood · Drying · Radiofrequency · Microwaves

L. Bandici · T. Leuca
Department of Electrical Engineering, University of Oradea, Oradea, Romania
e-mail: lbandici@uoradea.ro

T. Leuca
e-mail: tleuca@uoradea.ro

S. Coman (✉)
Department of Computers and Information Technology, University of Oradea, Oradea, Romania
e-mail: scoman@uoradea.ro

T. Leuca
Academy of Romanian Scientists, Bucharest, Romania

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_25

Nomenclatures

A. Acronyms

FEM	Finite Element Method
BEM	Boundary Element Method
RF	Radiofrequency
MW	Microwaves
CCITCEE	Center for Research and Technological Engineering in Electromagnetic Energy Conversion

B. Symbols/Parameters

ε^*	Complex Dielectric Constant
ε'	Dielectric Constant
ε''	Dielectric Loss Factor
$tg\sigma$	Loss Angle Tangent
U	Moisture Content of Wood
m_{water}	Quantity of Water in the Wood
m_{wood}	Quantity of Dried Wood
m_{wet}	Initial Mass of the Wood
m_{dried}	Mass of the Wood after Drying
U_A	Anode Voltage
TE_{10}	Electric Transverse Mode
C_p	Specific Heat
P	Density
k	Wave Factor

1 Introduction

The interaction of the electromagnetic field with dielectric materials at high frequencies allows the development of a large range of scientific, industrial and medical applications [1].

The part that describes the interaction among the electromagnetic field and loads is critical when we want to design high frequency applicators, because the dielectric properties of the processed material will become an important part in the applicator's functioning [2, 3].

Various installations or devices such as microwave ovens, medical therapeutical applicators, industrial microwave sensors, RF or MW plasma applicators, and many more falls into this general category.

RF and MW heating are limited to a set of frequencies reserved for scientific, medical and industrial purposes. In addition to these permitted frequencies, it is

necessary to consider the possibility of interfering with other radio waves that may affect the economic efficiency of the electro thermal installations used [4].

This chapter presents issues concerning RF and MW heating of wood. The first part of the chapter describes theoretical terms about phenomena that stand at the base of the processing and description of dielectric properties of wood. The second part consists of a general description of RF and MW installations used for experimental determinations. The final part of the chapter includes results of the numerical modeling using a program in which the electromagnetic field is coupled with the thermal field and mass problems. The results of the simulation with the Comsol program are also presented.

1.1 The Current State of Dielectric Processing in the Radiofrequency Field

In the case of RF heating/drying, the temperature and moisture gradient increases in the material starting from the inside to the outside. The thermal energy is created by the electromagnetic field that will penetrate the material to be processed.

With the development of MW technologies, RF technologies have also been developed, but these have remained anonymous due to the microwave oven that has drawn attention to microwaves [5, 6].

Although the large public does not know radio frequency processing very well, it has been widely used in the industry since the 1950s, with a lot of electro thermal equipment for the processing of dielectric materials being developed.

In Japan, in 1952, the company founded by Denki Kogyo—DKK, first develops a vacuum tube oscillator, necessary for RF generators, and three years later after research and experience it begins to produce and market RF equipment [7].

After the 1990s, the domain of the dielectric processing in a radio frequency field developed further, so that today RF technologies are extremely numerous, and from the existing companies on the market we can mention: PSC Company, The Nemeth Group Incorporated, High Frequency Electronics Ltd etc. [8, 9].

Most RF processing technologies can be found in various industrial applications of heating/drying ceramics, glass fiber, textile fibers, wood etc.

1.2 The Current State of Dielectric Processing in the Microwave Field

The benefits of microwaves are renowned, starting with the fact that the dielectric materials are immediately penetrated by the electric field by diffusing thermal energy into the dielectric subjected to heating. There are heating methods that depend on

the transfer of thermal energy from the surface to the interior of the load, which negatively influences the quality of the finished product [10, 11].

Raytheon Company and General Electric Company were the first American companies to study MW heating, Raytheon Company developing and manufacturing equipment at a frequency of 2.45 GHz, and General Electric Company making equipment at a frequency of 915 MHz. In 1968, Midea Company held a monopoly in high frequency field installation manufacturing in China [12].

Since the 2000s, the development of equipment for processing in MW field has developed more and more, especially for industrial applications.

2 General Aspects of Processing of Dielectrics in a High Frequency Field

High-frequency technologies are often used as alternatives to heat transfer in dielectric materials by convection, conduction, and radiation methods. The major advantage of high frequency field processing is that the material is heated in the whole volume because of the fact that the energy is absorbed into the load and transferred by a contact surface [13].

2.1 Radiofrequency Heating

This type of heating is characterized by the phenomenon of potential diffusion of the thermal field in the dielectric material and is used in cases where the conventional methods of heating do not meet the conditions imposed by the speed and the uniformity of the temperature in the material.

In the case of a heating device that uses RF field the generator is going to create an alternative electric field between two electrodes. The dielectric is located inside the applicator, where the alternative electrical energy determines a continuous reorientation of the polar molecules in the material subjected to heating, and, because of the friction, the material is heated throughout the volume [6].

This heating can be explained by the movement of the electrical charges under the action of the energy dissipated in the material molecule (electronic polarization) and at the boundaries between two heterogeneous media (ion polarization).

Under the action of the electric field, the material molecule having positive and negative charges tends to deform. As the polarity of the electrodes is reversed, the charges of the molecules are attracted in opposite directions and by the successive changes of the directions the phenomenon of heating appears.

2.2 Microwave Heating

MW heating is, basically, similar to radio frequency heating and consists of the absorption of the electromagnetic wave by the material undergoing processing.

The distribution of energy in the material corresponds to the stationary wave regime, which can cause heterogeneous heating (it can be avoided by causing a brazing in the heat sources), but in some situations with this disadvantage, heating in the microwave field is the only one that can be taken into consideration.

By MW heating inverse thermal gradients are created in the dielectric material, the temperature advancing from the inside to the outside of the material, compared to the classical methods whose disadvantages are reflected in the non-uniform heating and in the long period of time required to reach the imposed temperatures [14].

This heating effect occurs, on the one hand, due to the polarization of the charged particles of the material by the electric field (losses through hysteresis), and, on the other hand, due to the free charge conduction.

2.3 Dielectric Properties. Complex Dielectric Constant

Dielectric materials are characterized by their ability to stock energy. When the intensity of the electric field increases beyond certain limits, dielectric materials lose their insulating qualities.

The mechanism of moving electrons under the action of an external electric field is named polarization. Four fundamental classes of polarization [15, 16] are known: electronic, atomic, orientation, and heterogeneous polarization. The orientation and heterogeneous polarization contribute to high frequency processing to which the complex dielectric constant is added. The behavior of a dielectric material under the action of the high frequency field is described by the complex dielectric constant ε^*, expressed by its real and imaginary components [15]:

$$\varepsilon^* = \varepsilon' - j\varepsilon'' \tag{1}$$

The loss factor includes all dissipative effects due to dielectric losses and the Joule effect. The relation of the loss factor is [15, 17]:

$$tg\sigma = \frac{\varepsilon''}{\varepsilon'}. \tag{2}$$

2.4 Material Parameters of Wood

High frequency electromagnetic field processing is influenced by frequency, temperature, moisture and density [17].

The dielectric properties that characterize the dielectric material under the action of a high frequency electromagnetic field are the complex dielectric constant ε^*, which defines the propagation speed of the electromagnetic wave through a dielectric material as well as its ability to accumulate energy and the $tg\sigma$, which characterizes the ability of the load to accumulate thermal energy [18, 19]. The ratio of the $tg\sigma$ to the ε^* describes the capability of the load to transform electromagnetic energy from the outside and dissipate it throughout the volume in the form of heat [19, 20].

Because the wood is very important in the industry, the physical and thermal properties in high frequency field processing were studied, as to get superior quality finished products [17, 21–24].

From the literature studied: [3, 25–29] the following parameters that influence the dielectric properties of the wood can be described: moisture, temperature, frequency, density, thermal conductivity, specific heat, and penetration depth [17].

2.4.1 Moisture Content of Wood

The humidity of the wood represents actually the amount of water contained in wood when it is to be processed. The moisture content of wood U can be expressed by the ratio of the mass of water in the wood to the mass of the dried wood [29]:

$$U = \frac{m_{water}}{m_{wood}} \cdot 100\,[\%] \tag{3}$$

where: m_{water} is the quantity of water in the wood and m_{wood} is the quantity of the dried wood.

For a given piece of wood, the moisture content can be calculated with the relation:

$$U = \frac{m_{wet} - m_{dried}}{m_{dried}} \cdot 100\,[\%] \tag{4}$$

where: m_{wet} is the initial mass of the wood and m_{dried} is the mass of the wood after drying.

The density of wood ρ is an important aspect that must be taken into account when determining its dielectric properties, as this parameter can cause significant variations of the properties. The thermal conductivity of the wood is directly influenced by the density, moisture, and temperature at which the heat transfer is made. Temperature and moisture strongly influence the specific heat values of the wood, which increase with the increase of each of the respective parameters [30–32].

2.5 Radiofrequency Electrothermal Installations

In general, a radio frequency field dielectric material processing installation has three basic components: the RF generator, the impedance adapter, and the applicator (Fig. 1).

Radiofrequency generators are divided into two main categories:

– radiofrequency generators controlled by power amplifiers (with self-excitation);
– power generators in which the dielectric material is part of the resonant circuit.

Radiofrequency applicators are mechanical devices that ensure optimal transfer of high frequency electromagnetic energy from the generator to the material being processed. From the point of view of industrial use there are three major types of RF applicators:

a. Through-field applicators are most often used in the continuous processing of dielectrics. They consist of conveyor belts with negligible dielectric losses passing through the electrodes (Fig. 2).
b. Stray-field applicators are mainly used for processing thin materials, such as paper and fabrics. The intensity of electric field at the surface of the material can be easily adjusted by the distance between the electrodes and the dielectric material (Fig. 3).
c. Staggered through-field applicators are used for processing large and flat products. Examples of such products are baking cereal cookies and cakes, drying or gluing wood etc (Fig. 4).

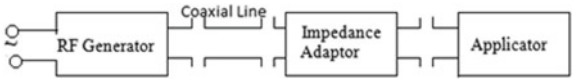


Fig. 1 Block chart of a radiofrequency heating device

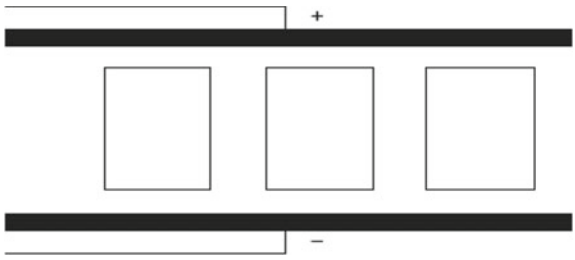


Fig. 2 Through-field applicator [37]

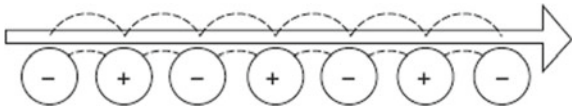


Fig. 3 Stray-field applicator type [37]

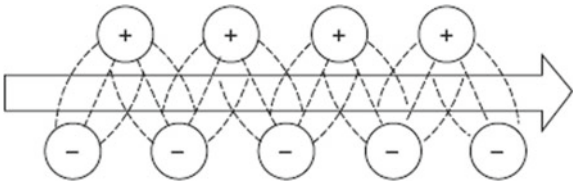


Fig. 4 Staggered through-field applicator [37]

In the CCITCEE there is an RF staggered through-field type laboratory installation, presented in Fig. 5.

This installation is equipped with a 13.56 MHz RF generator, which can be used at anode voltages up to 2000 V according to Table 1, and the total power of the installation is 1.5 Kw [17].

The characteristics of the installation are presented below:

– The installation is equipped with two pentode power tubes.

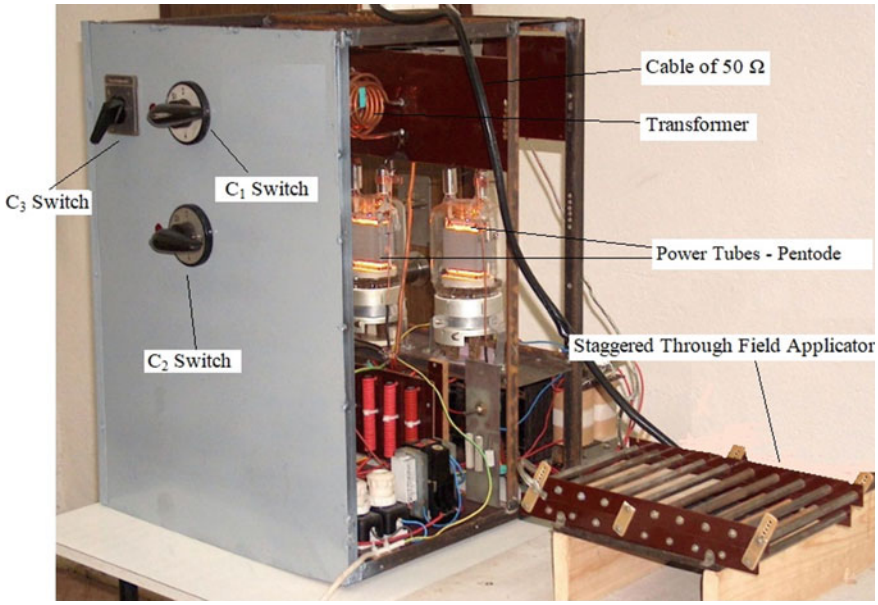


Fig. 5 RF staggered through-field installation [34]

Table 1 The values of the anode voltages of the 13.56 MHz RF installation [17, 22]

Values of the anode voltage U_A [V]								
1150	1240	1340	1380	1490	1550	1610	1690	1820

– The RF generator is serviced by the filament supply block represented by two transformers: the low voltage transformer (220–12 V) and the high voltage transformer (over 220 V).
– The RF voltage that can appear on the filament is switched off, as the case may be, by the two capacitors or by the four shock coils on the installation. Once the rectification and filtration have taken place, the voltage can reach 2000 V due to the filter block.
– The connection between the generator and the applicator is made with a 50 Ω RF cable.
– Using C_1 and C_2 switches up to nine stages of anode voltage can be obtained, presented in Table 1.

The operation of the installation firstly establishes the desired anode voltage level through switches C_1 and C_2, and then, after placing switch C_3 in the first position, the installation is supplied with a voltage of 220 ± 5% and at a frequency of 50 Hz.

2.6 Microwave Electrothermal Installations

A microwave heating system, from a functional point of view, has three basic components: the microwave generator, the waveguide, and the applicator.

At one end, the microwave generator generates the electromagnetic wave at a fixed or variable level of power, and at the other end, the applicator is so constructed that it achieves the connection between the most efficient electromagnetic energy, often referred to as a charger.

Between the generator and the applicator, there must be components that enhance the transmission and protect the generator from reflected waves. Figure 6 presents the block diagram of a microwave installation.

Microwave generators can be composed of a simple oscillator coupled to the measuring line, and the most complex ones allow the application of a calibrated and adjustable variable frequency signal.

The microwave applicator is the assembly where the heating process is happening through the interaction of the microwave energy with the dielectric material. If the heating is concentrated in a small loading area, then much more energy will be needed to heat the entire product to the demanded temperature than if the heating system is uniform.

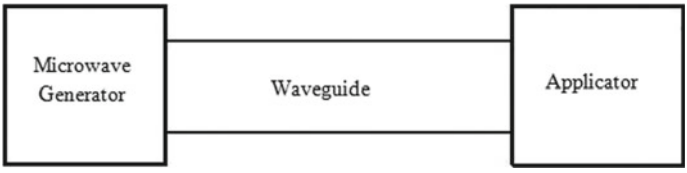


Fig. 6 Block diagram of the microwave electrothermal installation

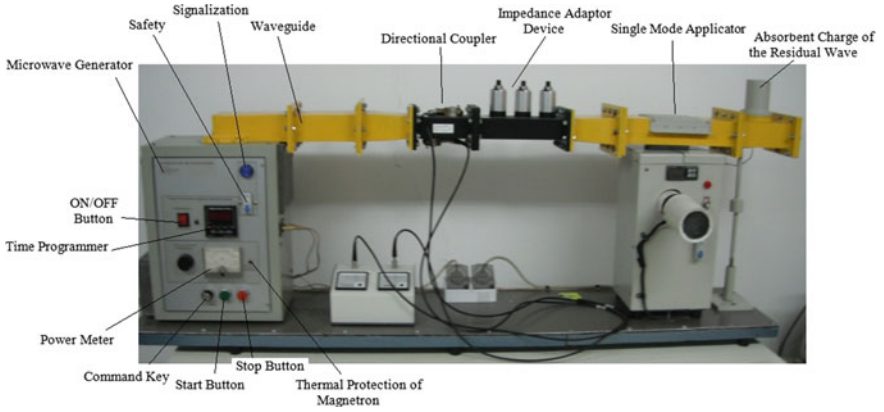


Fig. 7 Single mode applicator for drying dielectrics [34]

The most important categories of applicators are single mode resonant cavities and multimode applicators.

In the CCITCEE, there is a singlemode applicator, this microwave installation operates at a voltage of 220 ± 5%, a frequency of 50 Hz, and has the following main components: a generator with adjustable power up to 800 W, the waveguide, the applicator, the circulator, the directional coupler, and the impedance device.

The stand presented in Fig. 7 has the following characteristics:

- The directional coupler has the role of protecting the magnetron by determining the energy direction towards the dielectric material;
- Through the variation of the three plungers of the impedance adapter the shape of the field and the load impedance are being adjusted so that the enclosure is resonant;
- The applicator is the single mode resonant cavity operating in the TE_{10} electric transverse mode;
- The generator frequency is 2450 MHz, and can produce a variable power between 0 and 800 W.

3 Numerical Modeling of Dielectric Heating in RF and MW Field

Computer simulation has become an essential part of science and engineering. The purpose of choosing a simulation program is to be able to simulate models that describe the real processes as accurately as possible.

Due to the relationship between the theoretical and practical engineering aspects, the numerical methods of analysis have developed so much over time that the differential equation that mathematically reflects a certain phenomenon has been transformed into a system of algebraic equations, easily to be determined using the computer [14].

3.1 3D Calculation Programs for Numerical Analysis of the Drying Process Using Electromagnetic Field Coupling—Thermal Field—Mass Problems—Movement Problems

Based on the 2D calculation programs that use the FEM method and the FEM-BEM hybrid method, developed by other researchers from the CCITCEE was developed a 3D numerical computing software package written in Fortran.

The 3D numerical computing software package is based on the equations of the FEM-BEM hybrid method and is composed of a program that can simulate the drying process in the RF field when the dielectric is stationary, FEM-BEM.3D-RF, and a program with which the drying process in the radio frequency field can be simulated when the dielectric material is moving, FEM-BEM.3D-RF.

These programs consist of a main program that calls for several subroutines, in which were defined, one at a time, the data that refer to the applicator, the dielectric material, the FEM equations inside the dielectric material, the BEM equations on the (outer) borders of the dielectric material, and the thermal field equations, realizing the joint of the electromagnetic and the thermal field, with mass and movement problems.

The electromagnetic field in anamagnetic regime inside the wood piece is analyzed using FEM.

The diffusion of the thermal field is given by FEM, using the same mesh as in the electric field problem, while the time discretization has been done using the trapezoidal method [24].

The developed 3D software package has the following algorithm:

- reads the data of the applicator and of the dielectric material;
- reads the mesh of the electrodes and the mesh of the dielectric material, indicating the nodes on the border, calculating the volume of the tetrahedrons on the border of the dielectric and the surfaces on the borders of the electrodes;
- builds the matrix equation between the potentials of the nodes on the dielectric boundary and the derivatives after the normal on the sides of the dielectric;
- calculates the FEM submatrices of the electromagnetic and thermal problems associated with the nodes on the border;
- calculates the FEM-BEM matrix on the border of the dielectric material, considering the initial values of the properties defined for the dielectric material;
- calculates the losses produced by the electromagnetic field and the values associated with the nodes of these losses;

– calculates the material parameters during drying and the FEM matrix of the thermal problem, including the null elements, corresponding to the stationary regime term;
– calculates the submatrix corresponding to the nodes on the border in the thermal problem, including the null elements and the FEM matrix of the thermal diffusion, the time discretization being done by the trapezoidal method.

3.1.1 Numerical Results Obtained for Drying Immobile Dielectric Materials

The 3D programs, defined above, were used to simulate the drying of oak wood having the electrical and thermal properties taken from specialized literature [17, 20, 29].

The structure, for which the FEM-BEM.3D-RF program was used to simulate drying in the RF field when the dielectric is stationary, considers a staggered through-field applicator with 5 electrode pairs (the dimensions are according to [33]). The geometry of the structure for which drying in the radiofrequency field and the thermal field was simulated is presented below.

Figure 8 presents the geometry and the discretization network of the considered structure. Figure 9 shows the temperature distribution in the dielectric at 1150 [V] [17].

Simulations with the developed FEM-BEM. 3D-RF for program have been made for all the anode voltages at which the RF laboratory installation can operate. The program FEM-BEM.3D-RF. for calculates the maximum values of the temperature having an ambient temperature of 25 [°C]. The variations of the maximum and average temperatures in the dielectric with respect to the drying time for different anode voltages, as well as the variations of moisture during the drying process are shown in Figs. 10, 11, and 12.

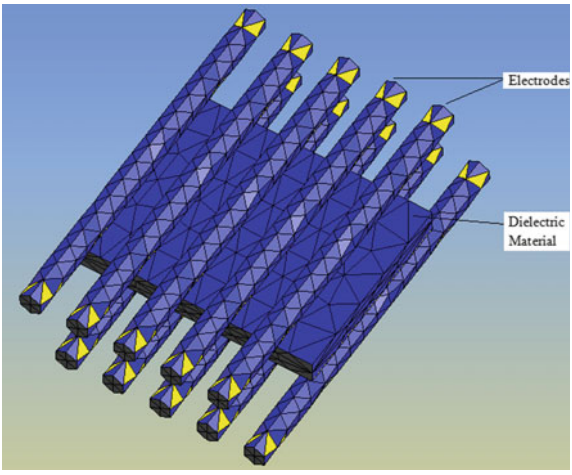


Fig. 8 The geometry and network discretization of the considered model

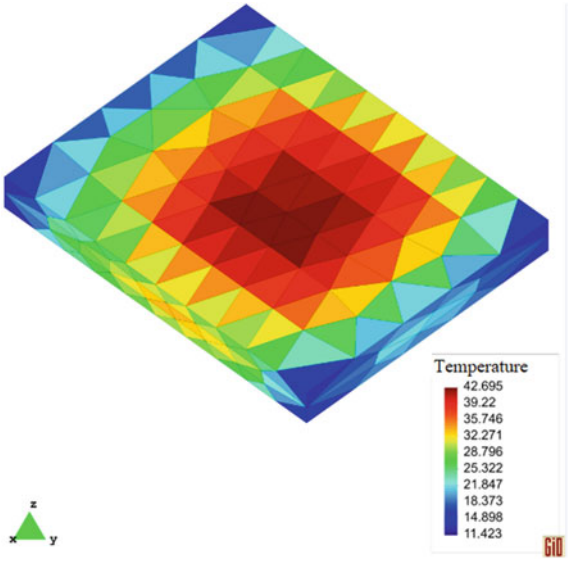


Fig. 9 Temperature distribution in the volume of the dielectric at an anode voltage of 1150 V

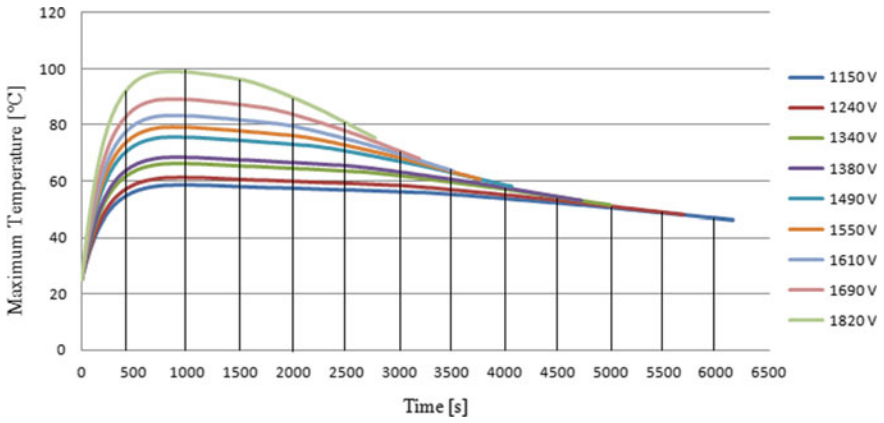


Fig. 10 Temperature variation in volume of the dielectric at different anode voltages

By analyzing the results obtained, it is found that, at high voltage values, the moisture of the dielectric decreases faster, respectively the drying time decreases. For all the analyzed cases the temperature does not exceed 80 [°C], and the moisture of the wood reaches 10 [%].

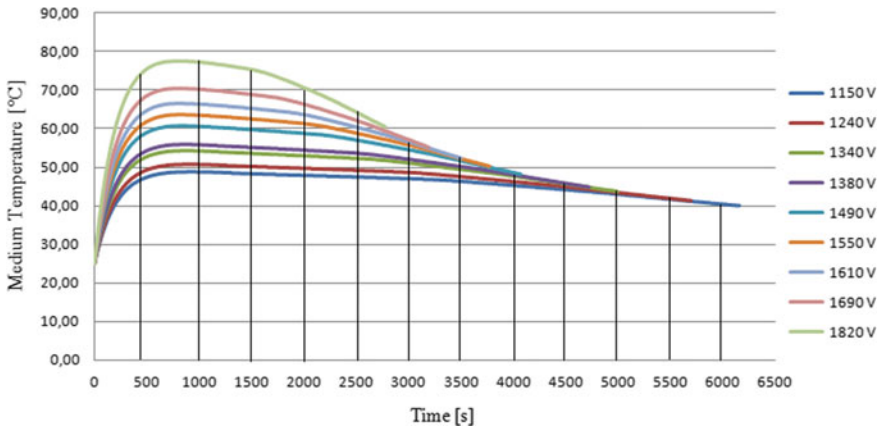


Fig. 11 Average temperature variation in volume of the dielectric at different anode voltages

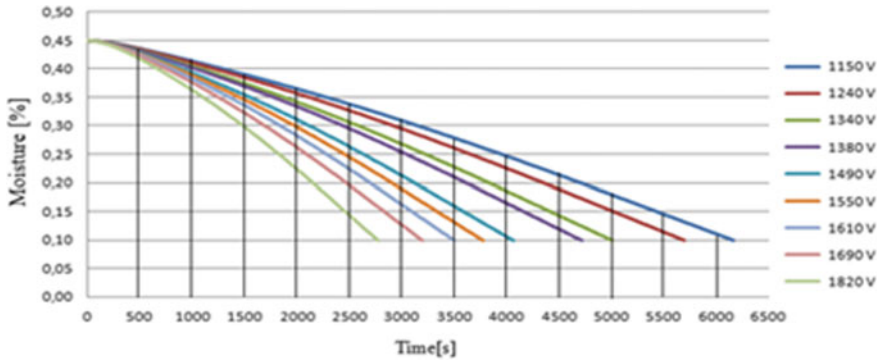


Fig. 12 Moisture variation in volume of the dielectric at different anode voltages

3.1.2 Numerical Results Obtained for Drying Dielectric Materials in Motion

Simulations of drying in the radiofrequency field of the oak wood and with the FEM-BEM.3D-RFmove.for program were performed for the situation in which the dielectric is in motion, having as its applicator the previously used structure and the same parameters [17, 20, 29].

Next are presented the geometry of the applicator and of the dielectric, for each position that the dielectric has in the applicator during drying, as well as the temperature of the dielectric, for each position of the dielectric in the applicator [17].

Figures 13, 15, 17, 19 and 21 show the geometry of the applicator and the different positions of the dielectric for which the temperature values were tracked and noted. Figures 14, 16, 18, 20 and 22 present the temperature distribution depending on the position of the dielectric inside the applicator.

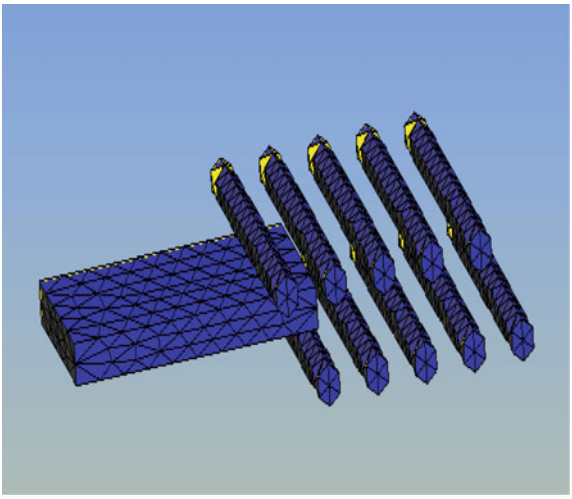


Fig. 13 Geometry and discretization network at the entrance of the dielectric into the applicator

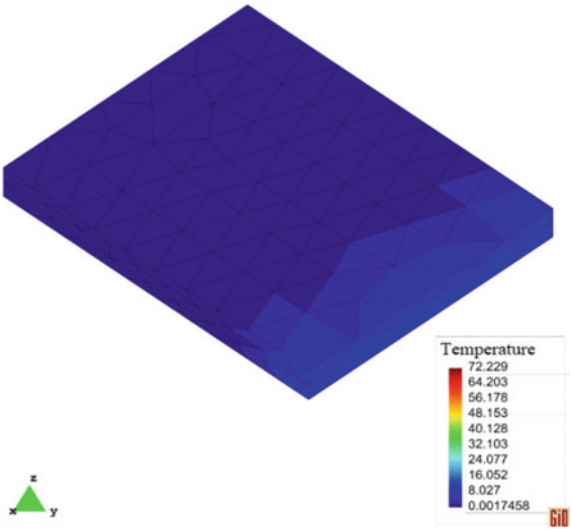


Fig. 14 Distribution of the temperature of the dielectric when entering the applicator

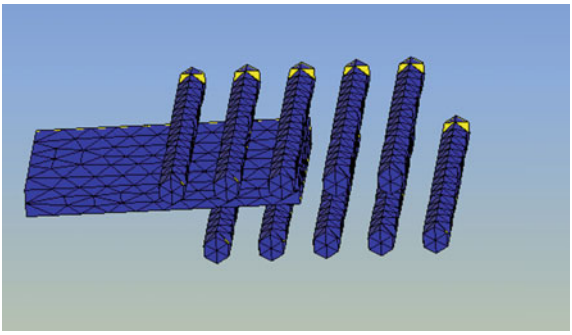


Fig. 15 Geometry and discretization network when the dielectric enters the first half of the applicator

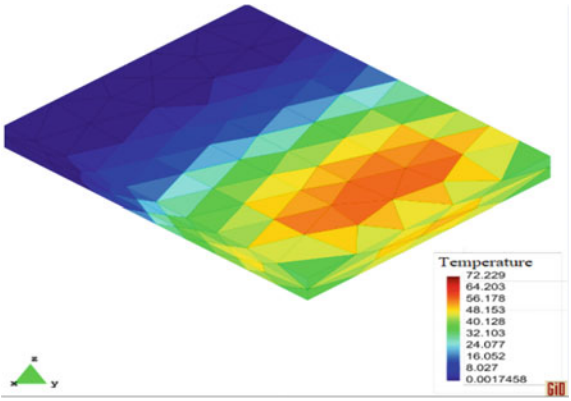


Fig. 16 Temperature distribution when the dielectric enters the first half of the applicator

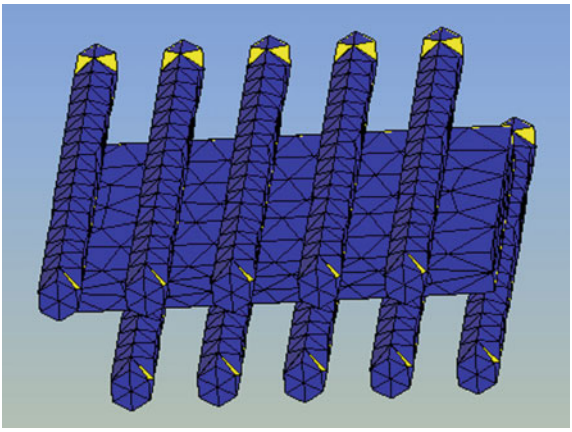


Fig. 17 Geometry and discretization network when the dielectric is inside the applicator

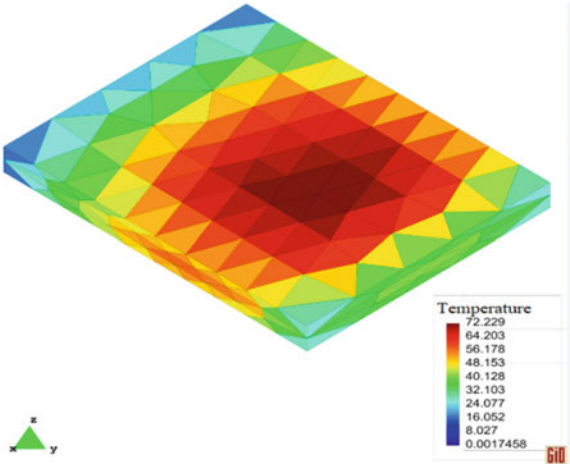


Fig. 18 Temperature distribution when the dielectric is inside the applicator

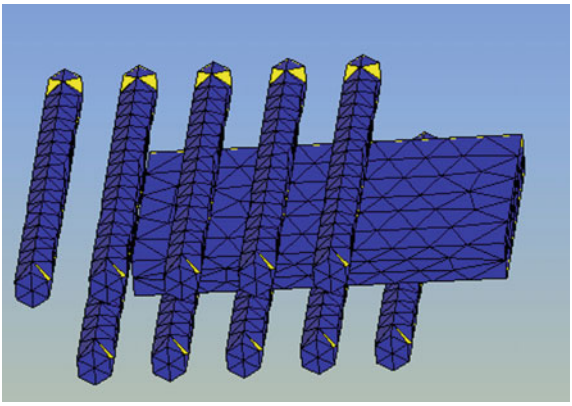


Fig. 19 Geometry and discretization network when the dielectric is in the second half of the applicator

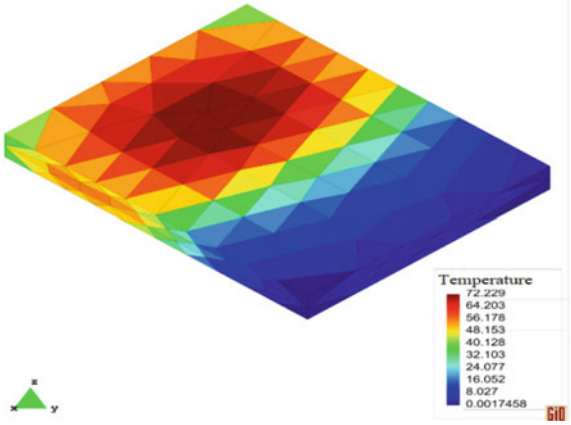


Fig. 20 Temperature distribution when the dielectric is in the second half of the applicator

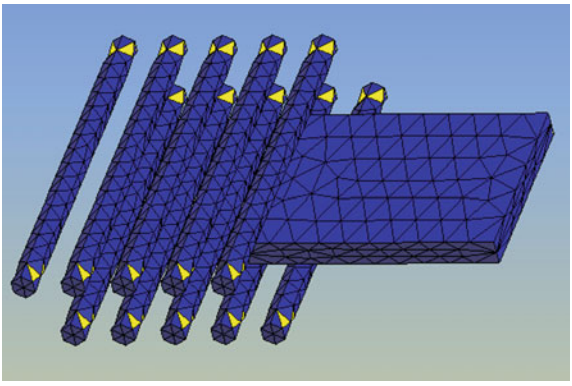


Fig. 21 Geometry and discretization network when the dielectric is at the exit of the applicator

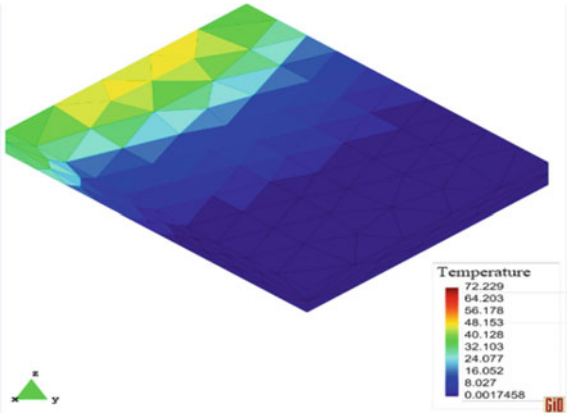


Fig. 22 Temperature distribution when the dielectric is at the exit of the applicator

When the dielectric is in motion, the distribution of the thermal field is not uniform on its surface. So it is recommended to use an external heat source. This method will reduce the heating time and obtain a better distribution of the thermal field on the dielectric surface.

The variation of the maximum and average temperatures in dielectric with respect to the drying time for different anode voltages, as well as the variation of moisture during drying when the dielectric is in motion are presented in Figs. 23, 24, and 25.

Figure 23 shows the variation of the maximum dielectric temperature in motion for different voltage values. By analyzing the graph we find values of the maximum temperature of almost 100 [ºC] at a voltage of 1820 [V] and a time of 4000 [s]. For lower values of the anodic voltage the maximum temperature recorded values between 60 and 80 [ºC]. The variation of the average temperature recorded values up to 70 [ºC] at a voltage of 1820 [V], for the other values of the voltage remaining at values between 35 and 50 [ºC]. As can be seen in Fig. 25 the moisture value recorded

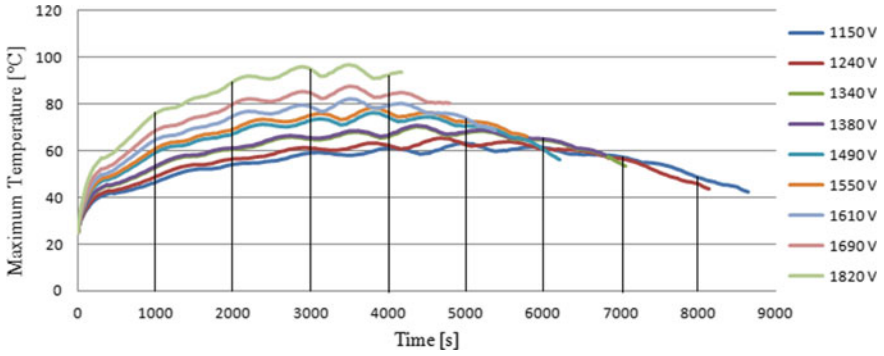


Fig. 23 Temperature variation in volume of the dielectric in motion at different anode voltages

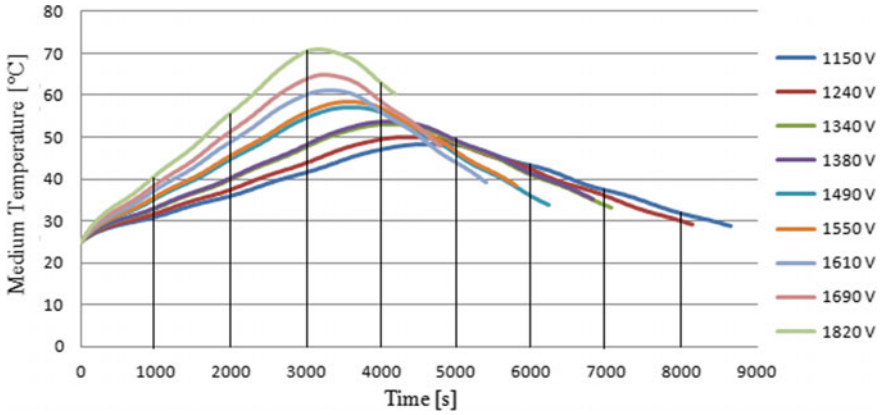


Fig. 24 Average temperature variation in volume of the dielectric in motion at different anode voltages

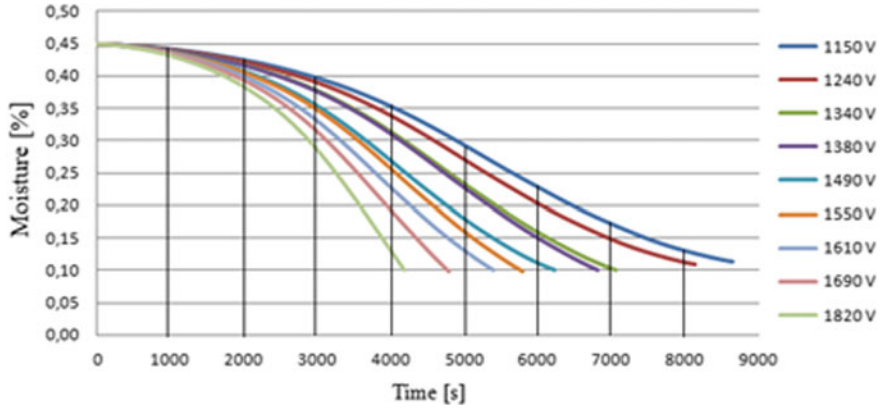


Fig. 25 Moisture variation in volume of the dielectric in motion at different anode voltages

the fastest decrease to 10 [%] when using the voltage of 1820 [V] and the time of 4000 [s]. For voltage values of 1490 [V] a time of 6000 [s] was necessary to reach the moisture value of 10 [%], while for a voltage value of 1150 [V] a processing time of over 8000 [s] was required.

Therefore, by analyzing the results obtained, as it was found previously, for high voltage values, the moisture of the dielectric decreases faster, respectively the drying time decreases.

Comparing the results obtained from the numerical analysis for the situations where the dielectric is stationary or in motion, it is found that the values of moisture and temperature of oak wood are slightly different [17].

The results obtained from the simulation of the drying process of the oak wood in the RF field, in either case, show an increase in the temperature of the dielectric and

a decrease in time required for the moisture to drop from 45 [%]–10 [%] (moisture at which wood is considered dried), due to the increase in anode voltage [17].

The results obtained from the simulation show an increase in time required to dry the wood when the dielectric is in motion compared to the stationary one, at the same values of the anode voltage.

3.2 Numerical Modeling of Heating in the Electromagnetic Field Using the Comsol Multiphysics Software

Comsol is one of those complex numerical modeling programs with a friendly interface that allows modeling of all the important physical aspects of their design. It uses numerical analysis with the method of finite elements, offering simulation software and a solver for a large scale of applications in the field of engineering and physics, in particular allowing the simulation of coupled phenomena.

For the numerical simulation of heating the fir wood was used the microwave drying installation from the CCITCEE [34, 35]. The purpose of the model is to determine the temperature in the dielectric material, the absorbed power and the electric field distribution according to the input data [17].

Input data refers to different values of microwave power and processing time, the other data remaining constant. Numerical simulations were performed using different values of processing time:

- 600 [s] with values of the microwave power of 10, 20, 30 40, 50, 60, 70, 80, 90, and 100 [W];
- 1200 [s] with values of the microwave power of 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 [W];

During the numerical simulations it was assumed that the dielectric is in a static state.

Three domains have been defined to perform the numerical simulation, having the dimensions according to [17, 36].

Once the geometry was completed the dielectric properties were defined for each of the previously described domain: for the cavity and waveguide domain the interior was defined as Air and the exterior boundaries as Copper. The dielectric was characterized by the following dielectric properties related to fir wood having humidity of approx. 62%: $\varepsilon' = 5$, $\varepsilon'' = 1.4$, $C_p = 2500$ [J/kgK], density $\rho = 760$ [kg/m^3], $k = 0.21$ [W/mK] [11, 36].

The discretization of the geometry was achieved as follows: Extremely Fine mesh—very refined—was used for the field called dielectric and type mesh Normal for the domains defined as cavity and waveguide, as can be seen in Fig. 26.

Figure 27 presents the electric field in the geometry, highlighting the minimum and maximum points specific to the propagation mode TE_{10}, realizing 5 sections with deformation effect along the XY axis [23]. In Fig. 28 the electric field distribution

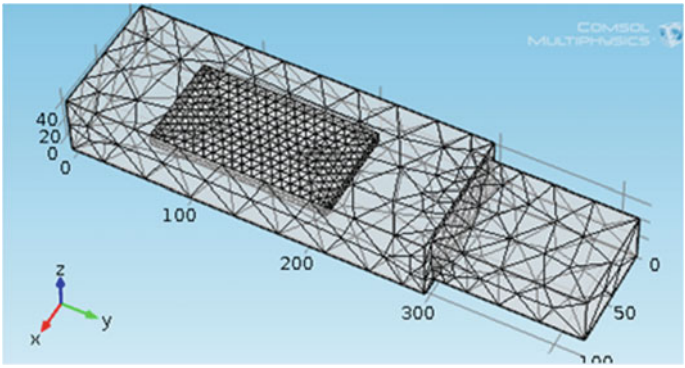


Fig. 26 Geometry and network discretization

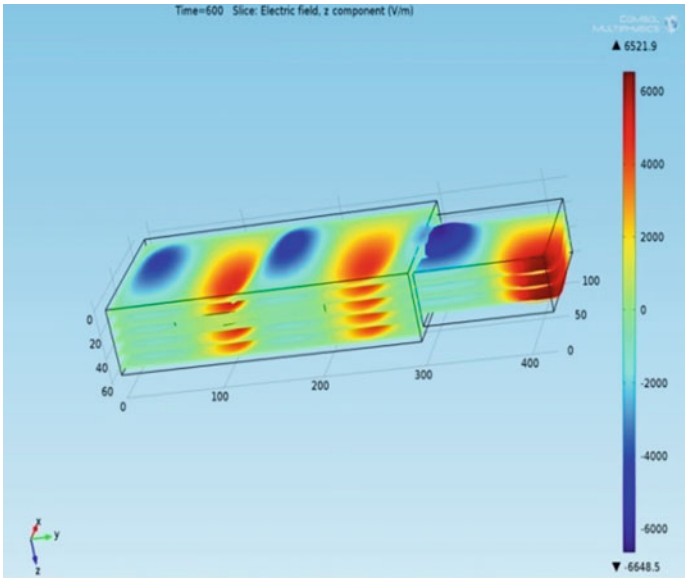


Fig. 27 Electric field throughout the geometry

through all geometry, cavity, and dielectric, using the Transparency setting from Comsol can be seen.

In order to highlight the electric field distribution through the cavity, guide, and load, the Slice setting was used, so in Fig. 29, the electric field distribution in the geometry through 30 sections along the ZX axis is presented [17, 23].

In Fig. 30, using the Arrow Volume setting, the power flow direction is presented, and its absorption can be observed near the dielectric.

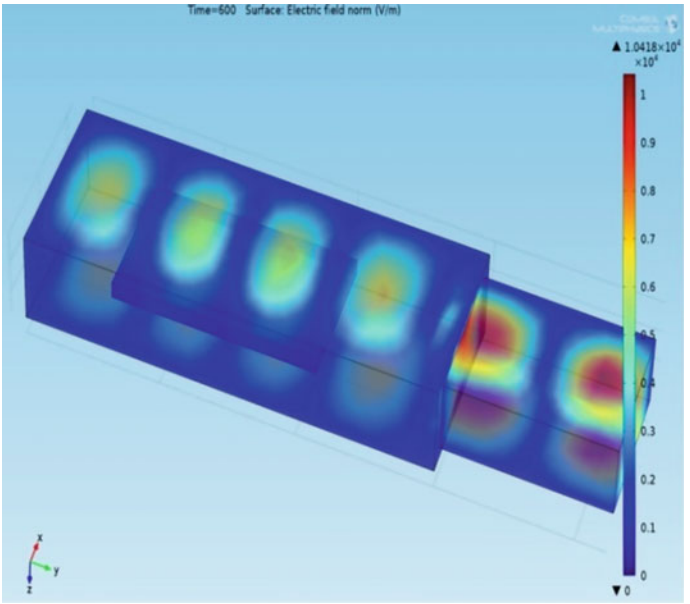


Fig. 28 Electric field in the geometry and dielectric

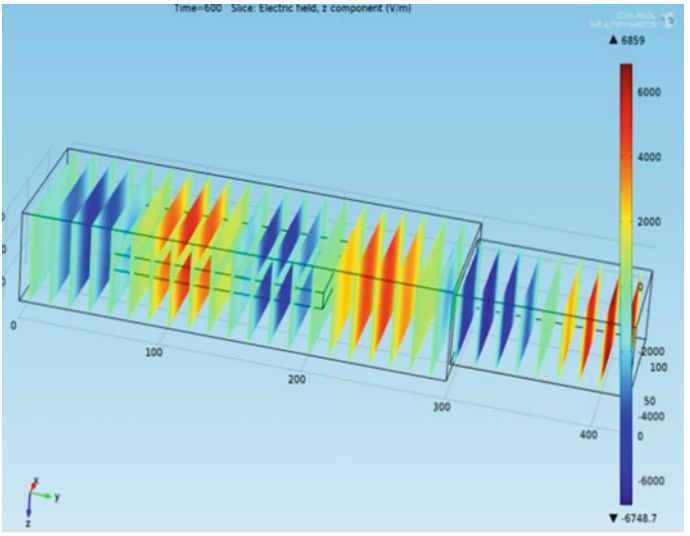


Fig. 29 Distribution of the electric field through sections along the ZX axis throughout the geometry

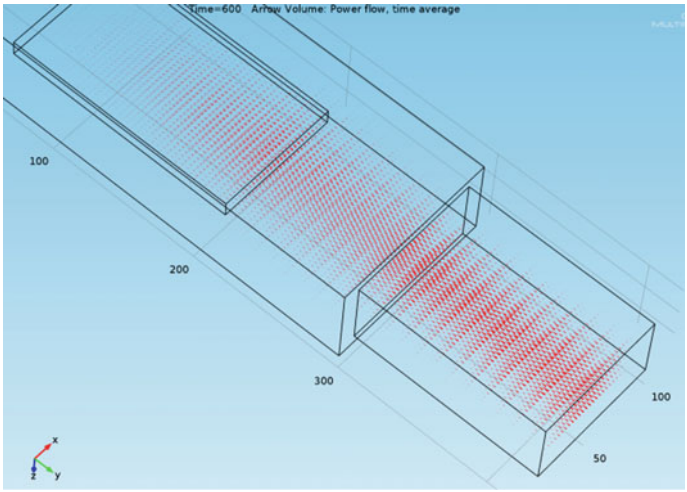


Fig. 30 Power flow through the applicator

3.2.1 Numerical Results of the Simulation of Microwave Field Heating Using the Comsol Multiphysics Software

During the numerical simulations, the following output data were followed:

- maximum temperature [°C]—represents temperature calculated in the dielectric at the end of the processing time [17];
- power absorbed [W]—presents the absorbed power in the material.

The variation in time of the temperature throughout the dielectric volume at different powers and processing time of 600 [s] is shown in Figs. 31 and 32.

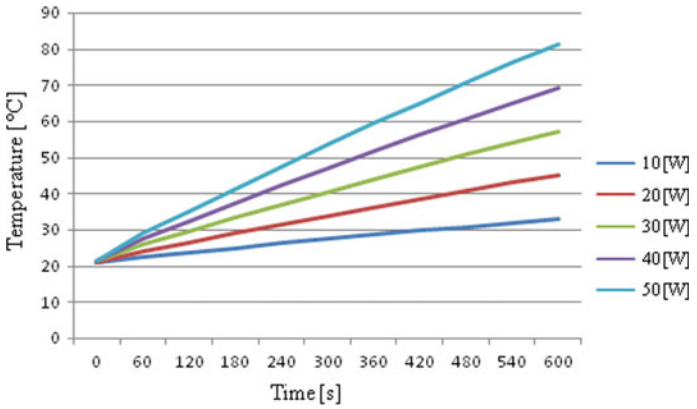


Fig. 31 Temperature variation in the whole volume of the dielectric material in 600 [s]

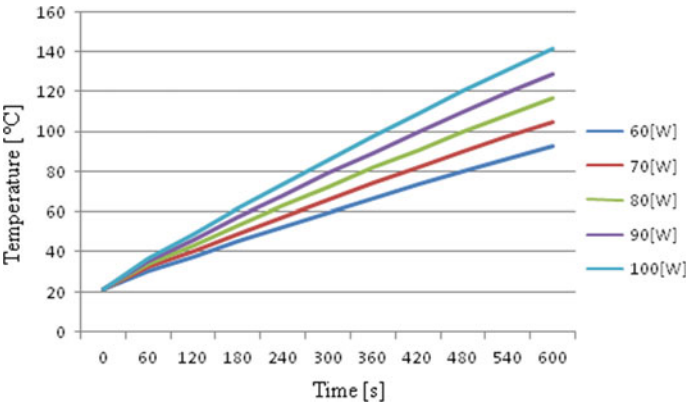


Fig. 32 Temperature variation in the whole volume of the dielectric material in 600 [s]

Figures 33 and 34 present the variation of temperature in dielectric for a processing time of 1200 [s] at power levels P_1 and P_2 ($P_1 < P_2$) [17].

By analyzing the results obtained, it is found that, with the increase of the power, the temperature increases to sufficiently high values that can cause dielectric degradation. For this reason, it is necessary to limit the power down to values at which the temperature does not exceed 80 °C, because above this value the dielectric is destroyed. According to Lambert's law, the power absorbed in the microwave field decreases exponentially in the sample.

The advantage of using the FEM-BEM.3D-RF and Comsol numerical modeling programs determines the field sizes in points of the domain, that are useful for the quantitative evaluation of the heating process.

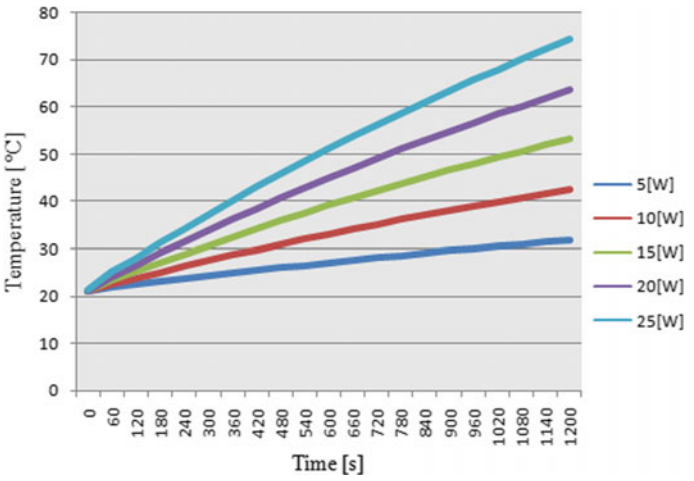


Fig. 33 Temperature variation in dielectric at P_1 power

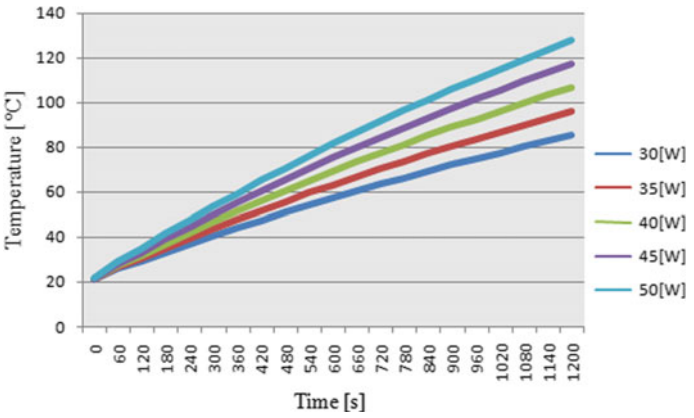


Fig. 34 Temperature variation in dielectric at P_2 power

4 Conclusions

In this chapter, we presented the numerical analysis of the drying process of wood in the radio frequency field, using its own software that combines the analysis of the electric field with the thermal one, with the mass and movement problem of some dielectrics.

Using the facilities of the Comsol Multiphysics commercial software, the numerical analysis of the electric field coupled with the thermal one was done when heating/drying the wood in the microwave field. As it is known, any numerical analysis of processes using IT modeling/simulation techniques provides important and more complete information than analytical calculations, useful in analyzing the process.

Within this study, a hybrid FEM-BEM method was developed to solve the problem of anamagnetic quasi-stationary electromagnetic field in 3D structures, for stationary and moving dielectrics. The main advantage of the method, compared to the FEM method, offered by most commercial programs, is that it separates the FEM component, applied to the dielectric, with an unchanged mesh during movement, from the BEM component, which defines the boundary condition of the FEM problem, but which implies the change of some matrices during the movement of the load. Due to the complexity of the problem, a high importance was devoted to creating techniques for reducing computation time, with many original solutions being proposed.

By radio frequency heating/drying of dielectric materials of large size, with regular and repetitive shapes, it has been confirmed that, especially in the wood industry, this technique always remains a topical solution due to the advantages it offers.

References

1. Leuca T, Molnar CO, Arion MN (eds) (2014) Elemente de Bazele Electrotehnicii. Aplicatii utilizand tehnici informatice. University of Oradea
2. Nelson SO, Datta AK (2001) Dielectric properties of food materials and electric field interactions. In: Datta AK, Anantheswaran RC (ed) Handbook of microwave technology for food applications, 2nd edn. Marcel Dekker, New York, pp 69–107
3. Coman O, Leuca T, Coman S (2012) Numerical modeling of drying wood in high frequency electromagnetic field. J Electr Electron Eng Rom 5(2):37–40
4. Sobol H, Tomiyasu K (2002) Milestones of microwaves. IEEE Trans Microw Theory Tech 50(3):594–611
5. Bandici L, Banu GS, Ficai A, Ficai D (2019) The role of susceptors in the process of obtaining nanopowders using microwaves. In: Paper presented at 15th international conference on engineering of modern electric systems (ICEMES 2019) pp 69–72
6. Gustrau F (eds) (2012) RF and microwave engineering. Fundamentals of wireless communications. Publisher Wiley
7. DKK-Denki Kogyo (2001). www.denkikogyo.co.jp. Accessed 30 August 2019
8. High Frequency Electronics (1988). www.highfrequency.co.nz. Accessed 30 August 2019
9. Nemeth Group (1975)www.nemeth-engineering.com. Accessed 30 August 2019
10. Coman S (eds) (2015) Analiza numerica a campurilor electromagnetic in regim de microunde si termic cuplate. Proiectarea optimala a unui aplicator. University of Oradea
11. Coman O (2014) Algoritmi de optimizare în proiectarea asistată de calculator a instalațiilor de microunde. PhD thesis, University of Oradea
12. Midea Inovation (2012) Worlds first solid state RF microwave owen. http://www.prnewswire.com/news-releases/mide-inovation-worlds-first-solis-state-rf-microwave-owen-159894045.html
13. Von Hippel AR (ed) (1954) Dielectric material and applications. Wiley, New York
14. Bandici L (ed) (2005) Modelarea Numerică a Câmpului Electromagnetic şi Termic Cuplat în Instalaţiile de Încălzire în Câmp de Microunde. Mediamira, Cluj-Napoca
15. Metaxas AC, Meredith RJ (eds) (1983) Industrial microwave heating. Peter Peregrinus, London
16. Gupta M, Eugene WL (eds) (2007) Microwaves and metals. Wiley
17. Laza M, Coman S, Leuca T (2015) Temperature variation in the process of heating oak wood using radio frequency. In: Paper presented at the conference on engineering of modern electric systems (EMES), pp 1–4
18. Kostoreva AA, Kostoreva ZA, Rogovaya LV, Loginov VS (2017) Research of heat and mass transfer processes in conditions of microwave heating of wet wood. In: MATEC web of conferences, p 110
19. Aichholzer A, Schuberth C, Mayer H, Arthaber H (2018) Microwave testing of moist and oven-dry wood to evaluate grain angle, density, moisture content and the dielectric constant of spruce from 8 GHz to 12 GHz. Eur J Wood Wood Prod 76(1):89–103
20. James WL (1975) Dielectric properties of wood and hardboard: variation with temperature, frequency. Moisture Content and Grain Orientation, USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin
21. Nelson SO, Bartley PG (2002) Measuring frequency and temperature dependent permitivities of food material. IEEE Trans Instrum Meas 51(4):589–592
22. Laza MI, Leuca T (eds) (2015) Modelarea si simularea numerica 3D a procesului de incalzire/uscare a dielectricilor in camp electromagnetic de inalta frecventa. University Of Oradea
23. Bandici L, Leuca T, Coman S (2017) The use of microwave field energy in the drying process of wooden sticks. In: Presented at 14th international conference of engineering of modern electric systems (EMES). IEEE. pp 95–98
24. Leuca T, Laza M, Bandici L, Cheregi G, Vasilescu GM, Drosu O (2014) Fem-bem analysis of radio frequency drying of a moving wooden piece. Revue Romaine des Sci Tech 4:361–370
25. Olmi R, Bini M, Ignesti A, Riminesi C (2000) Dielectric properties of wood from 2 to 3 GHz. J Microw Power Electromag Energy 35(3)

26. Oloyede A, Groombridge P (2000) The influence of microwave heating on the mechanical properties of wood. J Mater Process Technol 67–73. Elsevier
27. Aguilar-Reynosa A, Romaní A, Rodríguez-Jasso RM, Aguilar CN, Garrote G, Ruiz HA (2017) Microwave heating processing as alternative of pretreatment in second generation biorefinery. Energy Convers Manag 136:50–65
28. TenWolde A, Simpson W (1999) Physical properties and moisture relations of wood. In: Wood handbook: wood as an engineering material. USDA Forest Service. Products Laboratory, Madison, Wisconsin
29. Torgovnikov G, Vinden P (2010) Microwave wood modification technology and its applications. Forest Prod J 60(60):173–182
30. Nelson SO (2015) Dielectric properties of agricultural materials and their applications. In: Chapter 1—theory and fundamental principles. Elsevier
31. Nelson SO, Trabelsi S (2014) Dielectric properties of agricultural products: fundamental principles, influencing factors, and measurement techniques. Electro-Technologies for Food Processing Series
32. Samuel VG, Samuel LZ (2010) Moisture relation and physical properties of wood. General Technical Report FPL-GTR-190
33. Leuca T, Coman S, Trip ND, Bandici L, Codrean M, Perte M (2019) Neural network modeling of a drying process in radio frequency field. In: Presented at the 15th international conference on engineering of modern electric systems (EMES). University of Oradea, pp 193–196
34. Centrul de Cercetare şi Inginerie Tehnologică în Conversia Energiei Electromagnetice – CCITCEE (2004) https://erris.gov.ro/Center-for-Research-and-Engi. Accessed 07 September 2019
35. Coman S, Coman O, Leuca T, Laza M, Slovac F (2014) The use of experimental design in order to optimize the heating parameters of wood material inside a microwave applicator. In: Experimental results, presented at international symposium on fundamentals of electrical engineering (ISFEE). Bucharest, pp 1–4
36. Coman S, Coman O, Leuca T (2015) Optimal design of the microwave heating process using neural networks and genetic algorithms. In: Presented at 13th international conference on engineering of modern electric systems (EMES). Oradea, pp 1–4
37. Kitchen R (eds) (2001) RF and microwave radiation safety handbook. Oxford, UK

Design and Energy Analysis for Fuel Cell Hybrid Electric Vehicle

Mircea Raceanu, Nicu Bizon, Adriana Marinoiu, and Mihai Varlam

Abstract The environmental issues impose major changes in actual technologies for vehicle manufacturers. Nowadays, further research is focused on the development technologies for the vehicles of the future. Among these technologies, the fuel cell hybrid electric vehicle (FCHEV) has an important role due to the potential to improve significantly the fuel economy. FCHEVs can be more efficient than conventional internal combustion engines being an efficient and promising perspective. The lately research was focused on different configurations of FCHEVs, especially concerning the desired hybridization level involving the specific FC and batteries rules for their interconnection. Proton exchange membrane fuel cells (PEMFCs) is regarded as promising candidates for vehicle applications, mainly due to the mature technology, which can provide electrical power with high efficiency, less noise, compactness, lightness, low operating temperature, and very low emissions compared with conventional internal combustion engines. The electric efficiency usually represents 40–60% while the output power can be changed to meet quickly demanded load. The design of the power source in the FCHEVs is extremely attractive for transport applications. The FCHEV combines the advantage offered by PEMFC with the backup system using the efficient energy management assigned by the Battery. The LiPo rechargeable battery assures a quick transfer of energy during transient responses and continuous power during the absence of reactants. In this chapter, we provide a design and energy efficiency analysis for FCHEV implemented in ICSI ENERGY

M. Raceanu (✉) · A. Marinoiu · M. Varlam
National Center for Hydrogen and Fuel Cell, National Research and Development Institute for Cryogenics and Isotopic Technologies ICSI Rm. Valcea, Valcea, Romania
e-mail: mircea.raceanu@icsi.ro

A. Marinoiu
e-mail: adriana.marinoiu@icsi.ro

M. Varlam
e-mail: mihai.varlam@icsi.ro

N. Bizon
Faculty of Electronics, Communications and Computers, Department of Electronics, Computers and Electrical Engineering, University of Pitesti, Pitesti, Romania
e-mail: nicu.bizon@upit.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_26

Department, ICSI Rm Valcea, Romania. To ensure the required power, an energy management strategy (EMS) has been proposed. The FCHEV performance obtained in simulation using standardized load cycles is validated by taking into account a real experimental speed profile and numerical analysis of the acquired data. This EMS is focused on rule-based fuzzy logic control and state machine control. The developed FCHEV is mainly composed of PEMFC stack, LiPo rechargeable battery, and DC/AC inverter. The LiPo rechargeable battery is the main energy source, while the PEMFC plays the role of the support system. The feeding of the electric motor is assigned by the inverter which can convert the direct current (DC) in alternate current (AC). The PEMFC supplies the stationary/slow variable load, operating close to the maximum efficiency, and the battery supplies the load transients. Moreover, the PEMFC recharges the battery when is necessary, by considering the available extra energy. In order to validate the mentioned strategy, we analyzed the efficiency obtained by using the FCHEV in comparison with the efficiency using an only battery (electric vehicle). The results indicated more than 90% efficiency in the first case in comparison to 75% in the second case, respectively. The reliability of our model was tested and evaluated firstly taking into consideration various results by using of Matlab/Simulink environment. The experimental study was carried out by considering a specific protocol for the extra-urban driving cycle (EUDC). Therefore, this chapter takes into account an energy management strategy in order to analyze the efficiency obtained by using the FCHEV in comparison with efficiency by using only a battery (electric vehicle).

Keywords Fuel cell hybrid electric vehicle · Electric vehicle · Extra urban driving cycle · Energy efficiency · Numerical analysis

Abbreviation and Acronyms

FCHEV	Fuel Cell Hybrid Electric Vehicle
FCHPS	Fuel Cell Hybrid Power System
PEMFC	Proton Exchange Membrane Fuel cell
FCS	Fuel Cell System
EMS	Energy Management Strategy
ESS	Energy Storage System
EUDC	Extra-Urban driving cycle
GES	Global Extremum Seeking
ADVISOR	Advanced Vehicle Simulator
Batt	Battery
UC	Ultracapacitor
PWM	Pulse with Modulation
BMS	Battery Management System
SoC	State of Charge

DC/DC Converter Direct Current
DC/AC Inverter Alternative Current

1 Introduction

A promising solution to provide electricity with zero local emissions in automotive and stationary applications is the fuel cell system (FCS). The most commonly used FCS in hybrid power systems is the proton exchange membrane fuel cell (PEMFC) due to the high-power density, low volume, and low weight compared to other FCs [1, 2]. Meanwhile, a lot of research teams are working on new routes for obtaining hydrogen that uses renewable electricity such as water electrolysis or various processes of biomass gasification to produce clean electricity [3, 4]. In practice, due to sudden changes in the load during the vehicle acceleration phase, improper administration of water management and starvation phenomena with reactant due to the slow response of PEMFC, leading to loss of performance and cutting down of cycle life [5–7]. Consequently, to eliminate these disadvantages of the PEMFC mainly due to the slow dynamics, FCs are connected to other power sources (batteries, ultracapacitors) to meet the fast dynamics of the vehicle's electric motor [8, 9]. Since power is distributed among several sources, it is necessary to establish an energy management strategy (EMS) [10–12].

There are three types of electric vehicles namely: electric vehicles with batteries (BEV), electric vehicles with fuel cells (FCEV) and hybrid electric vehicles with fuel cells (FCHEV), the last uses FC as the main power source and batteries/ultracapacitors as the auxiliary power source. Fuel cell systems and battery packs have proven to be effective when working together to improve vehicle efficiency. The main challenge in developing FCHEV is finding an optimal power of FC/Batt/UC for which the efficiency is maximum, as well as establishing of power management algorithms that have as objectives: reducing hydrogen consumption, protecting FC from sudden loads and increasing the lifetime of FC. In this respect is preferable to operate the FC under the most stable conditions and as close as possible to the maximum efficiency point of the FC for a partial charge, while the battery can operate at a high current to remove the FC's weak points [13, 14]. There are numerous recent studies in the literature for the successful integration of FC into vehicles [11, 15–18]. Unlike pure electric vehicles, the battery system in FCHEV can be reduced in capacity, leading to weight loss and lower prices [19, 20].

PEMFCs are ideal for automotive applications. The most commonly used catalyst in PEMFC is platinum on carbon due to its good catalytic activity, in-creased durability, and corrosion resistance. Platinum is an expensive and rare metal, which is why the PEMFC price is high. However, the high cost of platinum has led researchers to find alternative solutions to reduce the platinum con-tent by using non-metals [21–23], gold nanocatalysts [24–26] and platinum decorated on graphene [26–29]. The specific power of PEMFC used in automotive applications has decreased from 1 to

0.65 kW/kg [30]. Fuel cell systems have an efficiency of up to 60% and are capable of providing a lifetime of 5000 h, performing 240 000 km, as well as large manufacturing volume, which indicates a cost of $30/kW calculated in the last 3 years [31]. A fuel cell is twice as efficient as an internal combustion engine, and depending on the capacity of the hydrogen tank it can have a range of up to 500 km [32, 33].

The energy storage system (ESS) is an important component of an FCHEV. The electric performance of the vehicle depends on the design, storage capacity, and type of storage used by ESS. So, depending on the type of vehicle (EVs, HEVs, and FCHEVs) the storage system capacity differs, as follows. Normally, EVs must have a higher storage capacity (34.5–140 Wh/kg), while the storage capacity for HEV is lower (26.3–77 Wh/kg) [34], and for FCHEV is between 8.06 and 18.45 Wh/kg. The ESS capacity for FCHEV must be carefully chosen to meet the starting at low temperatures, variations in energy demands and energy recovery from the braking of the vehicle. Thus, choosing the size of storage capacity and ESS hybridization is a challenge for FCHEV producers. Generally, the most used are the Li-ion and ultracapacitor batteries that are used either combined or separately. Another challenge of ESS is performance and robustness [35, 36].

Hydrogen has the lowest density and is very difficult to produce and store in a vehicle, due to the complexity of the production facility, which could be by reforming methane gas or by electrolysis of water. Usually, hydrogen is stored in high-pressure tanks, up to 700 bar. Its capacity ensures increased autonomy for an FCHEV. Toyota Mirai has a hydrogen mass of approx. 5 kg in two tanks [32, 33]. The power-on of FCHEV needs a battery system that starts the FC system because FC has a low power density and needs to be helped with a surplus of power from outside in the limit situations. The battery is a very expensive electrochemical device and if it is used improperly its life will be reduced, thus the ultracapacitors are often used for very fast charge variations [37]. The supply of the electric motor with electric power from an FCHEV is carried out by using a power mix provided by FC/Batt/UC and requires a reliable power management system. In Ref. [38], two EMS strategies are proposed in which PEMFC operates under stationary conditions, the battery is protected by depth of discharge, and UC is used in two modes: (1) high-pass filter of the charge/load at which the battery operates in a smooth manner or (2) UC functions as the power source, and the battery operates at constant power. The disadvantage of EMS is that it does not take into account the operation of PEMFC for several power levels. In Refs. [17, 39], the authors designed a hybrid power system composed of FC/Batt/UC, with the possibility of its application in automotive and stationary applications. The EMS system takes into account the slow dynamics of PEMFC that it compensates with ESS. EMS algorithms are implemented in a dSPACE controller and are focused in particular on increasing the lifetime of PEMFC.

In Ref. [40], a comparison between an FCHEV and an internal combustion vehicle is shown. FCHEV is a prototype (FC/Batt/UC). The efficiency of the FC converter is about 96% of the nominal power, and the voltage on the DC bus is regulated very precisely at 400 Vdc. This control is achieved by using PWM techniques that control the three-phase current of the traction motor. In the paper [41], a real-time control strategy for an FCHPS is analyzed, with the objectives of reducing hydrogen

consumption and increasing the electrical efficiency of the system. The strategy is based on a sensitivity analysis of critical parameters, using a Global Extremum Seeking (GES) search algorithm, 100 Hz sinusoidal dither, for which an electric efficiency increases of 7.83% was obtained compared to a classical strategy.

This chapter is organized as follows. Section 1 presents a short introduction of different configurations of fuel cell hybrid electric vehicles with a focus on identifying the advantages of an attractive challenge for future transport applications. Section 2 presents: (a) modelling of PEMFC; (b) modelling of LiPo rechargeable battery; (c) modelling of DC/DC converter and DC/AC inverter. Section 3 outlines the energy management strategy followed by the describing of the system efficiencies. Section 4 involves the experimental tests on a mentioned component which have been modelled in the previous section and their validation. Section 5 presents a detailed analysis of experimental results in respect to electrical vehicle efficiency, all these to demonstrate the strategies ability to ensure the required power by involving some improvements for hydrogen consumption and fuel efficiency. Section 6 presents the conclusions.

2 Description and Modelling of the Components of a Fuel Cell Hybrid Electric Vehicle

2.1 Proton Exchange Membrane Fuel Cell

PEM fuel cells are widely used in automotive applications. PEMFC is constructed of two electrodes (anode and cathode) separated by a Nafion membrane which is fixed by a bipolar plate. Electrodes are coated with a thin layer of the platinum catalyst. PEMFC is an electrochemical device that produces electricity, heat and water. PEMFC can operate continuously as long as it is fed with hydrogen and oxygen through the bipolar plate channels. The theoretical voltage produced by PEMFC is 1.23 V, due to the activation losses, the ohmic losses and the mass losses the voltage will be approximately 1 V. The optimum voltage at which PEMFC operates is in the range of 0.55–0.8 V. The voltage and current are inversely proportional. The reaction of electricity production of a proton exchanger fuel cell is shown in Eq. (1):

$$H_2(g) + \frac{1}{2} O_2(g) \xrightarrow{Pt} H_2O \tag{1}$$

Fuel cells normally operate in the ohmic region and their operation in the mass transport region is avoided. For this reason, ignoring the loss of concentration due to the overpotential of concentration, in a constant state, the terminal voltage of the fuel cell, Vt, is presented in Eq. (2):

$$V_{FC} = E_N - V_{act} - V_{ohmic} - V_{con} \tag{2}$$

where E_N, V_{act}, V_{ohmic} and V_{con} represents, the Nernst voltage and the losses of activation, ohmic and diffusion.

The Nernst voltage is described according to the fuel cell temperature and is shown in Eq. (3):

$$E_n = \left\{ \begin{array}{l} 1.299 + (T - 298)\dfrac{-44.43}{2F} + \dfrac{RT}{2F} \ln\left(P_{H_2} P_{O_2}^{1/2}\right) \; when \; T \leq 100\,^{\circ}\mathrm{C} \\ 1.299 + (T - 298)\dfrac{-44.43}{2F} + \dfrac{RT}{2F} \ln\left(\dfrac{P_{H_2} P_{O_2}^{1/2}}{P_{H_2O}}\right) \; when \; T > 100\,^{\circ}\mathrm{C} \end{array} \right\} \tag{3}$$

where R is the ideal gas constant, T is the operating temperature of the PEMFC (Kelvin), F is the Faraday constant (96,485 A s/mol), P_{H2} is the partial pressure of the reactant (atm.) and P_{O2} is the partial pressure of the oxidant (atm.).

The ohmic overpotential is presented in Eq. (4):

$$V_{ohmic} = R_{FC} \cdot i_{FC} \tag{4}$$

where R_{FC} is the FC resistance, and i_{FC} is the FC current.

Activation losses are the necessary energy for which the chemical reaction begins, these are observed in zone 1 of the polarization curve and are nonlinear. Because the kinetics reaction of the oxygen reduction is much slower than the hydrogen oxidation reaction, it is considered a single reaction. Then the voltage loss due to the activation polarization can be expressed by the Tafel equation:

$$V_{act} = A \cdot \ln\left(\frac{i_{FC}}{i_0}\right) \tag{5}$$

where A is the slope of the Tafel curve, i_0 is the exchange current.

$$V_{conv} = m \cdot e^{(ni_{FC})} + b \cdot \ln\left(\frac{P_{O_2}}{a}\right) \tag{6}$$

Diffusion losses can be described as follows, Eq. (7):

$$V_{con} = m \cdot e^{(n \cdot i_{fc})} + b \cdot \ln\left(\frac{P_{O2}}{a}\right) \tag{7}$$

where a, b, n and m are empirical coefficients that are determined experimentally from the polarization curve.

The rate of hydrogen consumption for the operation under stationary conditions of the PEMFC is according to the flow current and is calculated in Eq. (8):

$$C_{H_2} = \frac{i_{FC} \cdot N}{2F} \tag{8}$$

where N is the number of cells.

The theoretical efficiency of PEMFC is calculated as a ratio between the electrical energy produced and the enthalpy of hydrogen. The theoretical efficiency of PEMFC can be expressed by Eq. (9):

$$\eta = \frac{\Delta G}{\Delta_{\varsigma FC}} \tag{9}$$

where ΔG is the energy input and $\Delta_{\varsigma FC}$ is the useful produced energy.

The PEMFC efficiency is inversely proportional to the temperature and can reach a maximum value of 80% for a temperature of 25 °C [42, 43]. The calculation of the PEMFC efficiency is based on the enthalpy calculation, which is difficult to calculate in the actual operation of the PEMFC. For this reason, in [44] a simplified approach for calculating PEMFC efficiency is proposed. This is approximated according to the measured voltage and a coefficient of fuel use. This is expressed below:

$$\eta_{FC\_v} = \mu_{\mathrm{F}} \frac{V_{FC}}{E_{th}} \tag{10}$$

where μ_{F} is the fuel utilization speed (close to 0.95), E_{th} is the thermoneutral potential and is equal to 1.462 V for HHV and 1.254 V for LLV. The experimental tests were performed on a fuel cell system (FCS) produced by Hydrogenics. FCS is composed from 120 cells and has a nominal electrical power of 30 kW.

The PEMFC stack was conditioned and tested to calculate the power capacity installed in order to vehicle applications. The total test time was about 4 h. The stack was connected to a hydrogen tank, the air was supplied by a blower supplied from the output terminals of the stack. A water/water cooling exchanger has been designed to ensure the optimum operating temperature of the stack. To begin with, the hydrogen line from anode was purged with nitrogen. The load was emulated with a programmable electronic load feed was programmed at a constant current at a rate of 10 A/s. The maximum stopping current was 500 A. The tests were performed up and down, and repeated several times to observe the response of the additional equipment.

Figure 1 shows the voltage/current and power/current characteristic over the entire current range. The experimental tests were performed under steady-state conditions, at a working temperature of about 50–55 °C, the pressure and flow rate were monitored and recorded. From the graph it is observed that the voltage/power decreased/increases with the increase of the current, representing a normal functioning for the PEM fuel cells. The open-circuit voltage was 107 V and decreased to a minimum voltage of 65.9 V (approx. 0.55 V per cell). The voltage/current curve obtained is almost linear. The voltage losses through activation are visible in the domain of very low currents. The ohmic losses are clearly visible and increase with the current. The losses through concentration are very little visible due to the limitation of the equipment and the protection of PEMFC to the starvation phenomenon.

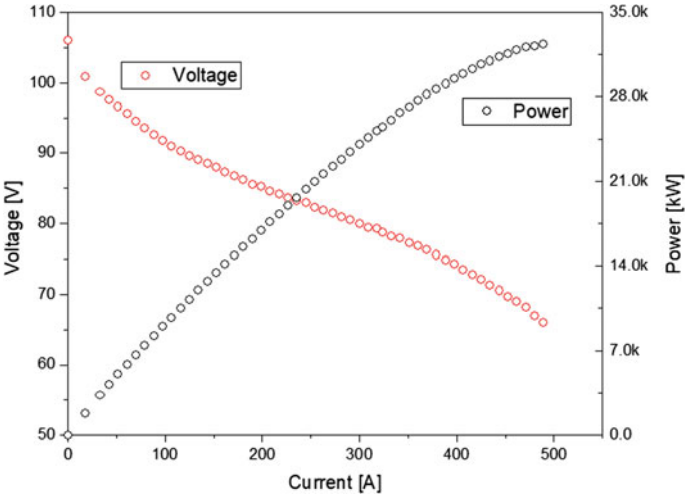


Fig. 1 I/V and I/P characteristics various the stack current

In Fig. 2, the overall efficiency of the fuel cell system was determined experimentally. From the graph it is observed that the efficiency of the system is over 50% in the current range within the range 30–270 A and this decreases slowly with the increase of the current. From the graph it is also observed that at a very low current, below 25 A, the system is extremely inefficient and operation in this menu should be avoided.

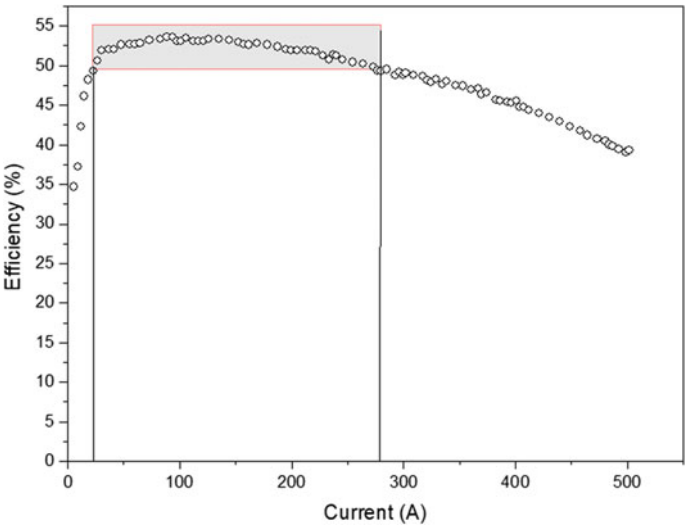


Fig. 2 Overall system efficiency a function of stack current

2.2 Battery

The battery is a storage device to store electricity in the form of chemical energy. Batteries are of two types: (1) primary batteries that are not rechargeable and (2) secondary batteries that can be recharged. Rechargeable batteries are most often used in automotive applications. The main parameters of the batteries are the storage capacity, the stored energy and the charging state. Depending on the type of battery, it works in certain domains of SoC to extend the battery life cycle. Currently, batteries are classified into five groups accessible for automotive applications, such as lead-acid, nickel, zebra, lithium and air-metallic [45].

The most commonly used types of batteries on the market commonly used in electric vehicles are Li-ion and NMC batteries. On the one hand, lithium-based batteries are the lightest, most compact and have a high energy storage capacity. On the other hand, NMC batteries are environmentally friendly, have high reliability and are cheaper [19]. The integration of the batteries in the FCHEV is necessary because it provides the power source to cold start the vehicle, and it ensures the storage of energy produced by the regenerative braking of the traction motor. However, the use of the battery in hybrid power sources needs to be analyzed in detail, as the battery is highly dependent on the discharge depth and the SoC.

The battery management system (BMS) deals with the sustainable use of batteries in power applications. BMS for managing lead batteries is relatively simple. The state of charge is calculated according to the voltage/current polarization curve. BMS has the role of regulating the value of the charge/discharge current supported by the battery and of balancing the voltages in the charging mode.

The "R_{int}" model available in ADVISOR is used to model the battery. The model is simple and has been tested and validated in [46]. The model is built from a voltage source in series with a resistor. The battery voltage and resistance are calculated according to the SoC for the required power, according to [47].

$$P = V_t \cdot I \tag{11}$$

where V_t(V) is the voltage measured at the battery terminals and I(A) is the battery current. V_t is the difference between the open-circuit voltage and the voltage drop on the internal resistance R(Ω), according to Eq. (12):

$$V_t = V_{OC} - I \cdot R \tag{12}$$

where V_{OC} is calculated according to the SoC of the battery pack, and the SoC is experimentally determined by integrating the current extracted from the battery. From Eqs. (11) and (12) the battery current is calculated.

$$I = \frac{V_{OC} - \sqrt{V_{OC}^2 - 4R \cdot P}}{2R} \tag{13}$$

For a more achievable calculation, the term under radical is neglected. Knowing the yield, the voltage measured at the battery terminals can be calculated using Eq. (12). But this approach does not take into account the dynamic behavior of the battery, which is influenced by the charge transfer resistance and capacitive effects between the battery plates. Thevenin model is used to take this behavior into account. Thus, a simple and precise battery model is obtained.

Due to the low budget, the battery system used in the design of the vehicle's propulsion train used lead-acid batteries. The battery pack is built from 32 at 12 V/20 Ah, C20, ST 150 VRLA batteries from ACEDIS, connected in series, with a nominal voltage of 408 V. The battery pack was charged with a constant current of 10 A (0.5 C) until the 441 V (2.3 V per cell) voltage was reached, then the current was lowered to 1 A and charging continued to the voltage at 464 V. In the last stage of charging, a constant voltage of 461 V was maintained over 30 min. In finally, it is considered that the battery pack is fully charged.

The most important parameter of the batteries is the open-circuit voltage, this parameter cannot be measured experimentally in real-time. To determine an empirical relationship between *Voc* and *SoC*, a fully charged battery discharge test was performed at a constant current of 8 A and certain intervals the battery was disconnected from the load and after 10 min of quiet, the voltage at the battery terminals was read, this represents the open-circuit voltage (Fig. 3). The charging/discharge tests were performed under the same conditions and the coulombic efficiency of the battery pack was calculated, as the ratio of the amount of current inserted/drawn from the battery was 92%.

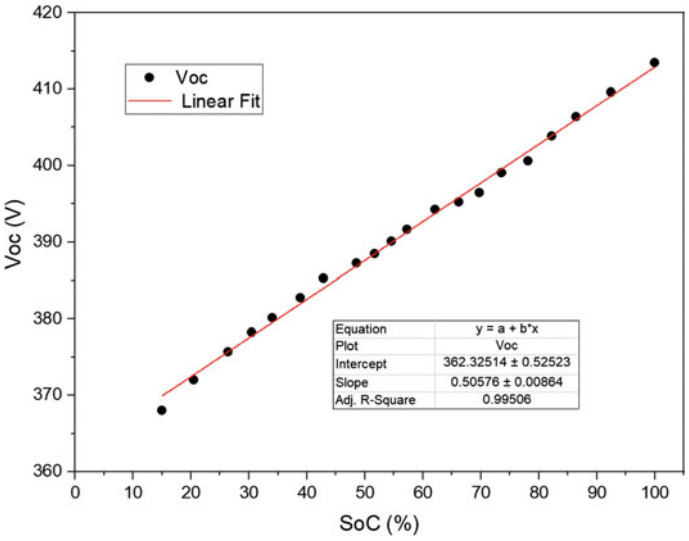


Fig. 3 Battery OCV versus state of charge

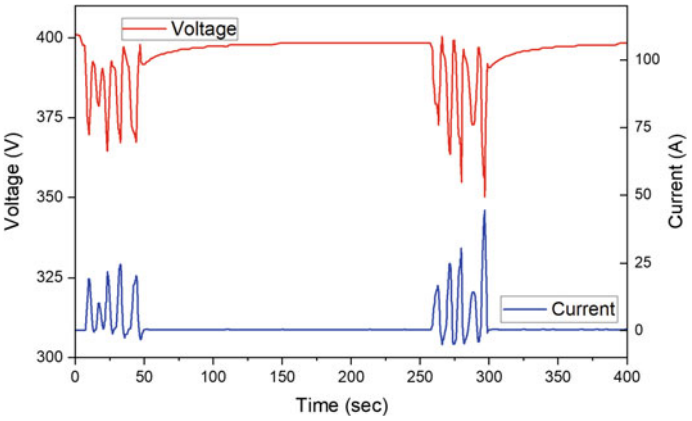


Fig. 4 Voltage response to a variation of current, data recorded at a sampling rate of 1 Hz

The *Voc* voltage decreases linearly with *SoC* over the entire field, as seen in Fig. 3. The relationship between *Voc* voltage and *SoC* was fitted using the least-squares method. This is presented in the following equation:

$$V_{oc} = 0.5SoC + 362.3 \tag{14}$$

The internal resistance of the battery pack varies considerably with increasing current. To determine the value of this resistance performance tests are performed under transient load conditions, the experimental data are presented in Fig. 4.

The test started from a Voc voltage of 401.1 V, the amount of current removed from the battery was 0.45 A/h, which represents 2.25% of the total battery capacity, and at finally the open circuit voltage was 398.4 V. The internal resistance of the battery every second was calculated by Eq. (12). The value of the internal resistance was calculated to 0.062 Ω for a SoC of approximately 75%, which is constant for the SoC range between 25 and 90%, below the value of 25%, the battery resistance increases nonlinearly up to 0.1 Ω. For this reason, experimental and simulation tests are not characterized below this value.

2.3 Ultracapacitor

The ultracapacitor (UC) is an electrochemical capacitor constructed of two layers that have a high energy storage capacity for automotive applications. UCs have a higher power density and a much longer service life than batteries. However, UC has several disadvantages, including very low energy density and a larger volume than batteries. Operating UC to meet power variations in automotive applications results in a large amount of heat produced inside the UC. The UC model must accurately predict the

electrical and thermal components in order to optimize the thermal management system in automotive applications.

UCs are widely used in hybrid electric vehicles powered by PEM fuel cells (FCHEV), where UC provides maximum power during acceleration and stores the energy recovered during braking. FCHEV simulation is a dynamic process that involves technical knowledge from several fields, the models must be accurate and robust.

The proposed UC model is based on the equivalent RC circuit developed in two-stage ladder model. The first stage imitates the internal construction of a single UC network and the performances are tested experimentally in a wide range of frequencies. In step 2 analytical or numerical solutions are found for different loads. Models can be easily combined (series/parallel) to form an ultracapacitor bench. The purpose of the model is to determine the SoC value and the working voltage depending on the open-circuit voltage and the current at which UC runs. The state of charge of UC is estimated by integrating the current over time, as presented in the equations below;

$$SoC = SoC_0 - \int_0^t \eta_c I / Q dt \tag{15}$$

where SoC_0 is the initial value of the state of charge, η_c is the average efficiency (charge/discharge) of the UC, Q is electrical quantity.

The voltage measured at the UC terminals is calculated using the equation below:

$$U_t = U_c - I \cdot R_s \tag{16}$$

where U_t is working voltage, U_c is open circuit voltage, I is demand current, R_s is equivalent series resistance.

The open-circuit voltage can be written according to the initial value of the open-circuit voltage for which the electric potential of the ultracapacitor is decreased.

$$U_c = U_{C0} - \int_0^\tau \frac{I}{C} dt \tag{17}$$

From Eqs. (16) and (17) the working voltage of the ultracapacitor, presented in Eq. (18) is calculated.

$$U_t = U_{C0} - \int_0^\tau \frac{I}{C} dt - I \cdot R_s \tag{18}$$

From Eq. (18) it is observed that the working voltage is linear with the open-circuit voltage and SoC and is expressed explicitly in Eq. (19).

$$U_t = f(SoC) - I \cdot R_s \tag{19}$$

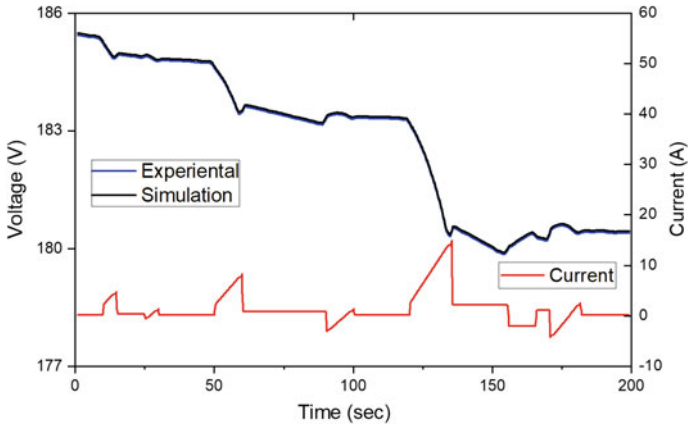


Fig. 5 Current profile for an ECE15 profile and UC voltage response

To validate the UC model, dynamic tests chosen to evaluate the UC performance are performed. The modelled working voltage is compared with the experimental results obtained. The ultracapacitor pack consists of 72 cells connected in series, with a maximum working voltage of 194.4 V and a capacity of 41.7 F. The test bench includes a data acquisition system, an electronic load and a two-way power supply. which emulates the current profile.

Figure 5 shows the comparison between the simulated model and the experimental results for an ECE15 driving cycle. The maximum error between simulation and experiment of the working voltage is of maximum 0.75%.

2.4 DC/DC Converter

Most DC/DC converters that are available on the market for use are suitable for fuel cell applications, usually the control interface is included in the converter. The DC/DC converter is the device that controls the current supplied by the fuel cell. The DC/DC converter is used to convert the fluctuating voltage of the fuel cell to a stable voltage on the DC bus. The converter must have a wide voltage range at the input and be sized for 120% of the rated power of the fuel cell system.

The converters are internally controlled with a current control loop so as to provide the system with security for limiting large currents and to obtain satisfactory working dynamics.

The system is provided with two DC/DC converters, one unidirectional that connects FCS to the DC bus and one bidirectional that controls the current of the battery system, the ultra-capacitive system is connected directly to the DC bus to work very fast at the power transitions.

The control of the two converters is controlled by the references I_{fc} and I_{batt}. In this way, the equations are presented as follows:

$$i_{fc} = i_{fc}^{REF} = \frac{P_{fc}}{V_{fc}} = \frac{P_{fc}^{REF}}{V_{fc}} \tag{20}$$

$$i_{batt} = i_{batt}^{REF} = \frac{P_{batt}}{V_{batt}} = \frac{P_{batt}^{REF}}{V_{batt}} \tag{21}$$

where P_{fc}^{REF} and P_{batt}^{REF} represent the power references of the FC and the battery, respectively.

A current limiter is used to provide power limitation as follows:

$$0 \leq i_{fc}(t) \leq i_{fc\max}(t) \tag{22}$$

$$i_{batt\min}(t) \leq i_{batt}(t) \leq i_{batt\max}(t) \tag{23}$$

3 EMS Control Strategies

EMS algorithms are used to control the operation of FC, Batt and UC, so that the energy demand is met under any vehicle operating conditions. The EMS must establish the reference power of the FC so intact to satisfy the load demand, the FC must operate between the minimum and maximum power and as close to the maximum efficiency point. The required power difference must be provided quickly by the UC and the battery. EMS must ensure voltage stability on the DC bus and reduce voltage fluctuations, thanks to the battery connected directly to the DC bus.

Another EMS objective is to find references for the fuel cell and battery system to provide the vehicle's traction engine power, as defined in Eq. (24).

$$P_{demand} = P_{fc} + P_{batt} + P_{UC} \tag{24}$$

Because battery is directly connected to the DC bus, the battery power cannot be directly controlled. EMS is limited to finding references for FC and UC (System shown in Fig. 6).

The control system inputs that are processed by the fuzzy controller are the following: the power required by the power train (Pload); power of the fuel cell at the output of the unidirectional DC/DC converter (P_FC); vehicle speed (Speed); vehicle acceleration (Acc.); the state of charge of the UC (SoC_UC) and the state of charge of the battery (SoC_Batt).

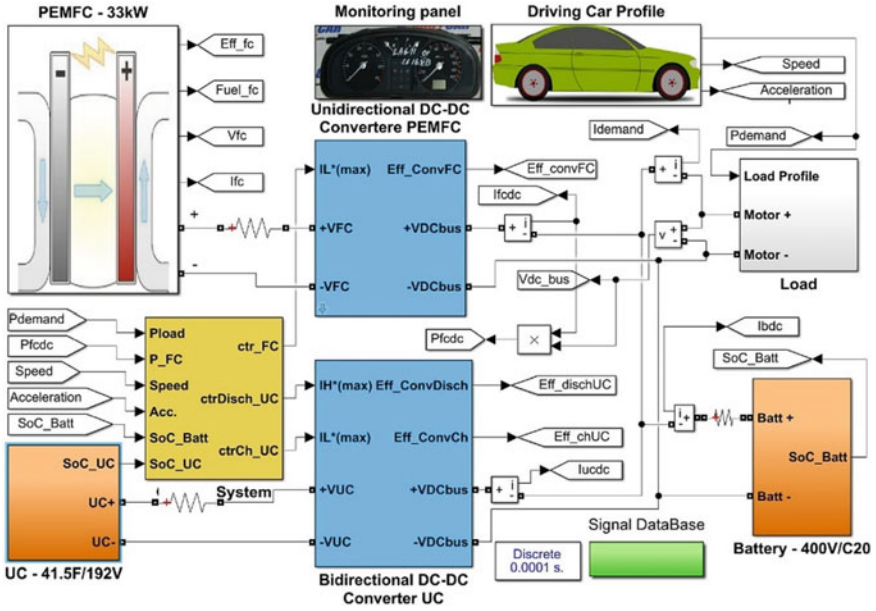


Fig. 6 Fuel cell hybrid electrical vehicle simulation model

The outputs of the control system are the following: the reference current at the FCS output (ctr_FC) and the reference current for discharging and charging the ultracapacitor bank (ctrDisch_UC and ctCh_UC).

Figure 7 shows all the signals that are monitored by four oscilloscopes. The data is saved in the Excel file at a frequency of 10 Hz for analysis and interpretation.

The fuzzy controller system manages the entire power management in FCHEV. As a result, the controller controls the complete operation of the fuel cell (including hydrogen and air flows) and UC SoC. The controller must protect its deep-charge battery system, including overcharging.

Figure 8 shows the FCS (Fuzzy Controller System) for the FCHEV control structure with PI control loop and the limitations imposed for the correct functioning of the components.

FCS is designed from membership functions that are based on IF-THEN rules, which establish certain critical operating situations and conditions that allow the operation of several energy sources at the same time. FCS design requires different processes of fuzzification, rule base and defuzzification. FCS is reliable and capable of achieving maximum efficiency based on reduced fuel consumption.

The shapes of the member functions are triangular in the interior area and trapezoidal on the outside, which are simple to describe due to practical experience in fuel cell testing. In the design it was taken into account that the member functions intersect at a single point, having a degree of membership of 0.5. Figure 9 shows the input functions of the FCS. Figure 10 shows the output functions of the FCS.

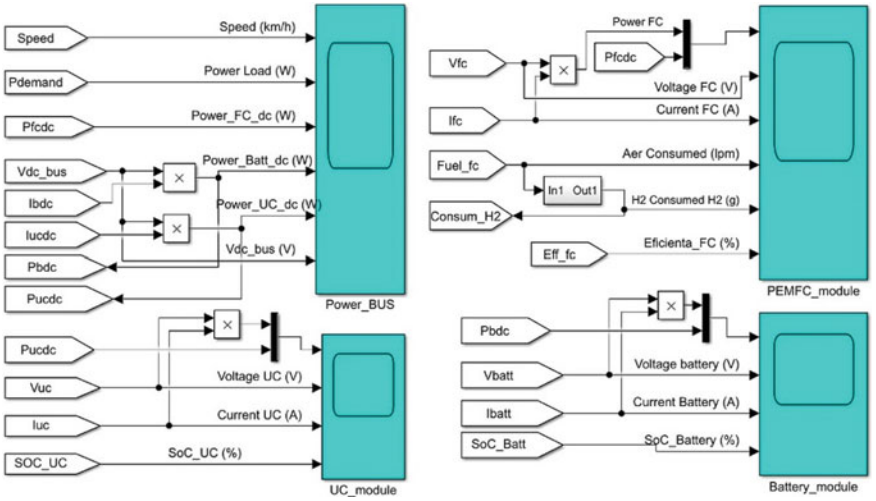


Fig. 7 Monitoring panel

The rules specified in the FLC respect the following order of priority: The fuel cell should operate as consistently as possible in the field; The UC state of charge must be included in the range 50–90%; The minimum power of FC is 0.5 kW and maximum of 30 kW; the battery state of charge of should be over 70% and the voltage on the bus will be about 400 Vdc.

The inference rules based on the Mamdani method are shown in Fig. 11, for the load power of 2.5 kW, the power difference between the load and the fuel cell is − 2.03 kW, the UC state of charge of 86.2, and the acceleration is zero, all these rules suggest that the FC power reference is 7.36 kW.

Figure 12 shows the control surfaces for the power reference of the fuel cell related to the power required by the vehicle, the power availability of the ultracapacitor and the acceleration of the vehicle. It can be observed that, due to the centroid fuzzing method, the control surface is smooth and does not show a sharp increase in the FC power reference, the output varies linearly at 25 kW (maximum). This maximum is obtained when the power of the traction motor exceeds 25 kW and the difference PL-Pfc is greater than 20 kW. The control surfaces are very suggestive.

The simulations are done using the Matlab/Simulink environment. The simulated FCHEV has a total mass of 1500 kg, the equivalent frontal area of the vehicle is 2.05 m^2, and the aerodynamic traction coefficient is 0.28 when considering the air density of 1.2 kg/m^3, and the air speed was considered zero. In the simulation, an extra-urban cycle (EUDC) of 400 s was used, as shown in Fig. 13. The maximum power of the vehicle was about 37 kW for a maximum speed of 120 km/h. The battery is connected to the DC bus took over all the sudden variations of the motor. Due to the control design, it was observed that the voltage at the battery terminals is in the optimum operating range. The UC was controlled so that when the battery voltage

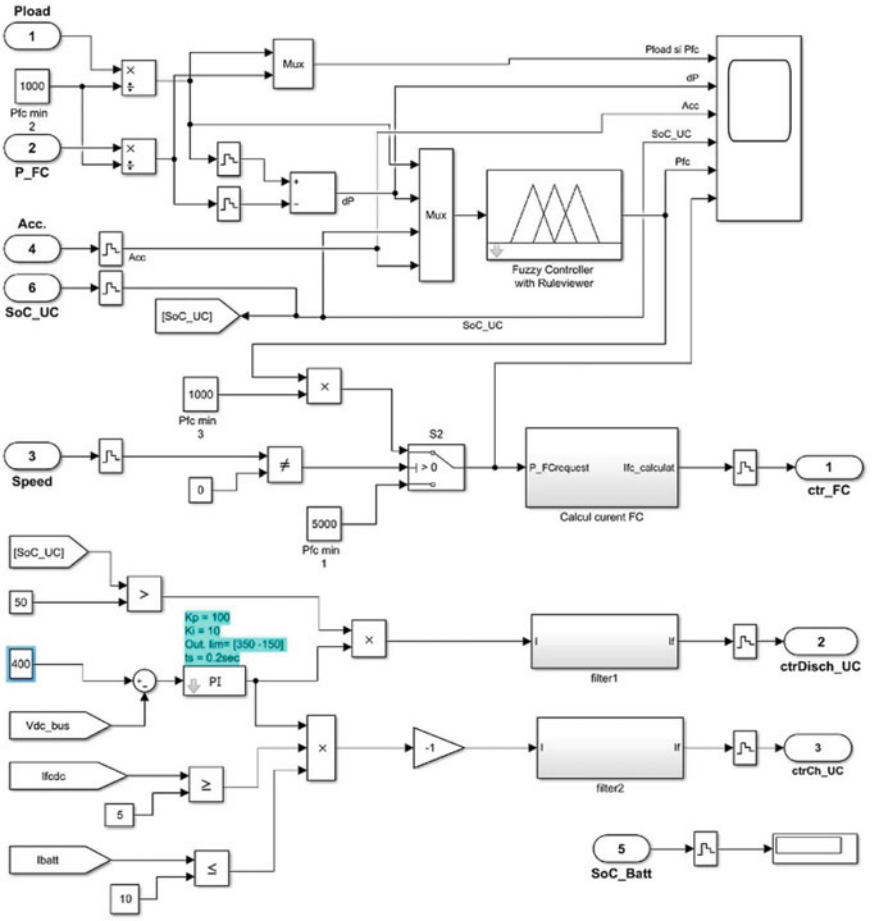
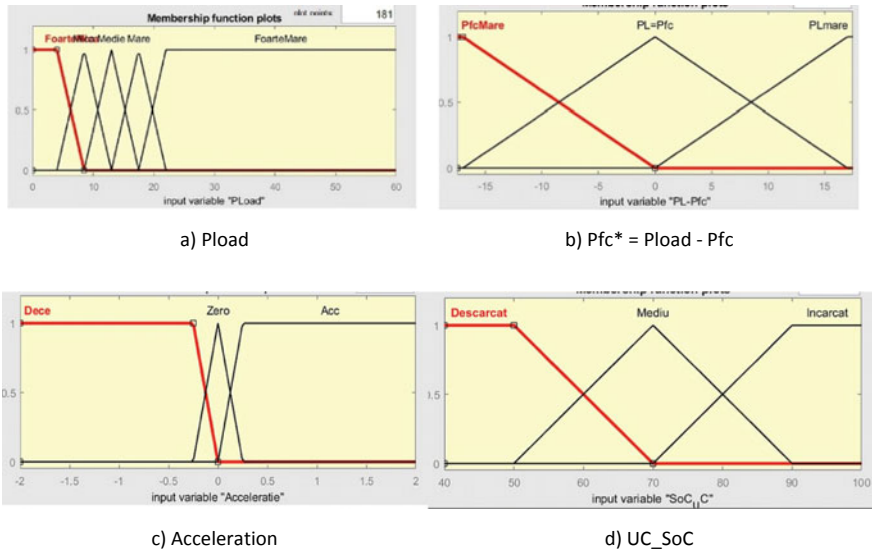


Fig. 8 The fuzzy controller system for FCEV

drops below 400 V it compensates very quickly for the power difference so that the voltage is stable at 400 Vdc.

Figure 14 shows the variation of the fuel cell system operating parameters. The voltage was maintained between the optimal operating parameters. Hydrogen consumption was recorded throughout the driving cycle and was 71 g. The rate of increase of the current was controlled at a maximum of 10 A/s and the efficiency of the fuel cell fluctuated between 50 and 62%. The reference power at the output of the fuel cell has a small ripple and this remains in the considered references.

Figure 15 shows the variation of the voltage, current and state of charge of the UC. The SoC is controlled between 50 and 92% values. UC is controlled to maintain on the DC bus a voltage as constant as 400 Vdc. It also has the role of keeping FC under semi steady sate conditions.



a) Pload b) Pfc* = Pload - Pfc

c) Acceleration d) UC_SoC

Fig. 9 Input membership functions of FCS

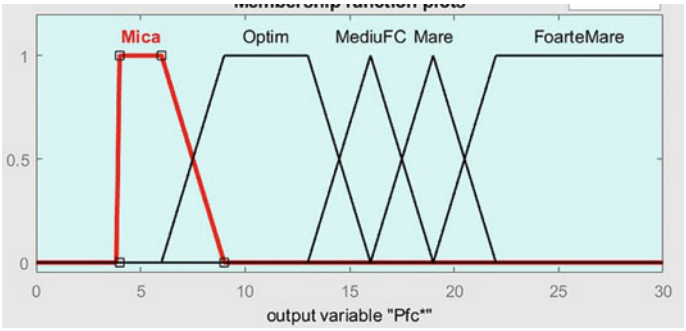


Fig. 10 Output membership functions of FCS

Figure 16 shows all the parameters that vary in the battery system. The battery voltage varies depending on the variation of the battery power, this is most evident in the acceleration or braking phases. SoC battery is kept within the limits and does not very much, because the use of this energy source is limited. In principle, SoC decreases only when the vehicle speed increases, as you have noticed.

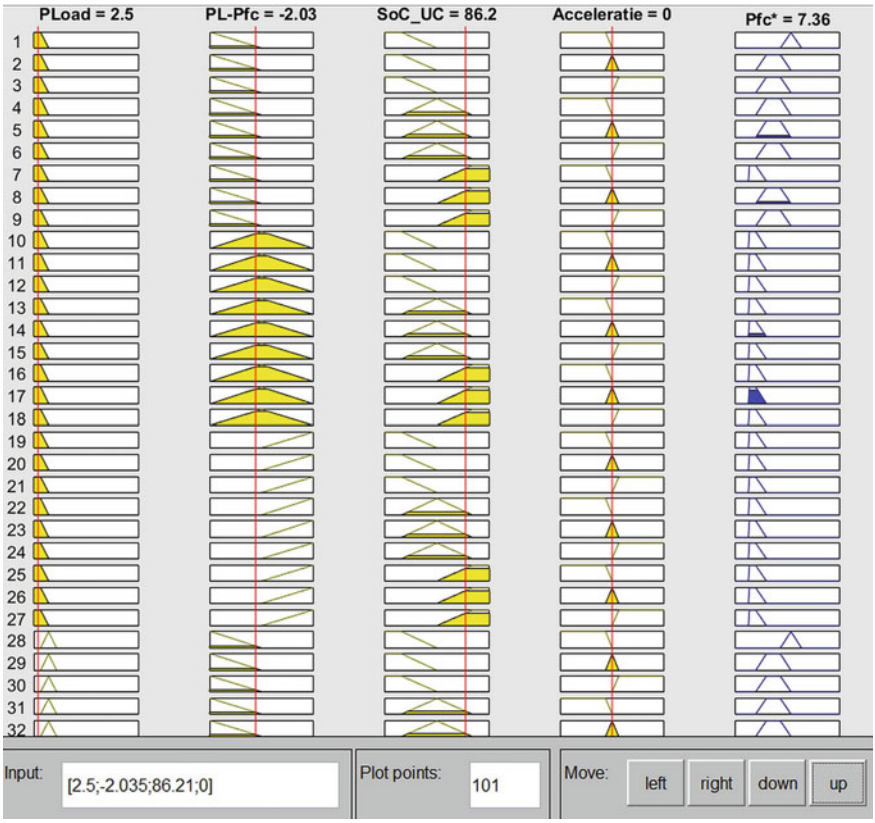


Fig. 11 Rule viewer

4 Experimental Validation

The configuration of the experimental installation for the validation of the EMS algorithms is shown in Fig. 17. To verify the real-time feasibility of the fuzzy control system, several experiment tests were performed. The proposed algorithm is implemented in a NI cRIO-9035 rapid prototyping system. Power sources consist of a Hydrogenics fuel cell system, a battery system and an ultracapacitor bank. FCS and UC bank are connected to the DC bus via two Brusa converters. The battery system is connected directly to the bus. The load conditions are emulated by a programmable electronic load.

EMS based on fuzzy logic was evaluated for a power profile of Fig. 18 which is an extra-urban driving profile (EUDC) and the power evolution of FC, Battery and UC are reported in Fig. 18. The experimental results provided show as the algorithm proposed by EMS allows the complementary operation of all power sources. Figure 19 shows the evolution of the voltages and the load state of the sources.

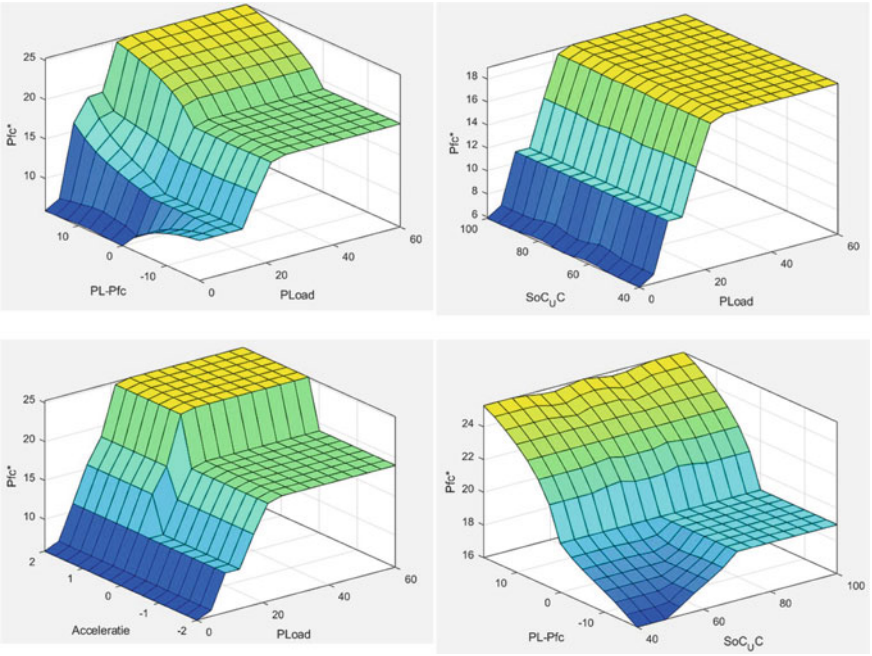


Fig. 12 Control surfaces for the power reference of the fuel cell

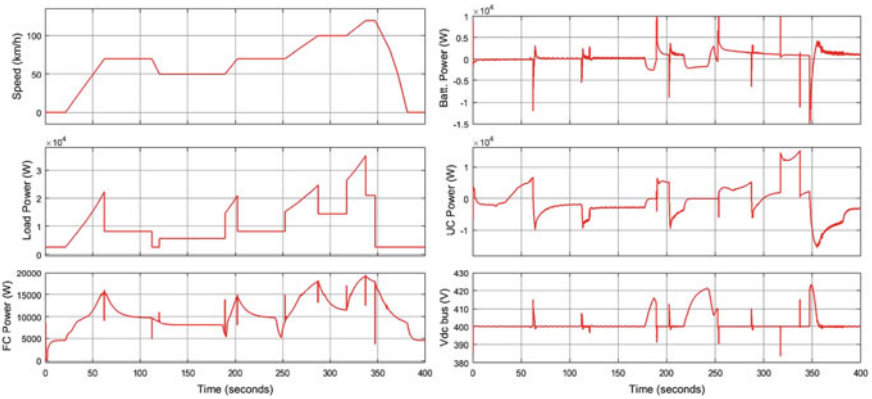


Fig. 13 Balance of power sources, voltage on DC bus and vehicle speed for the EUDC driving cycle

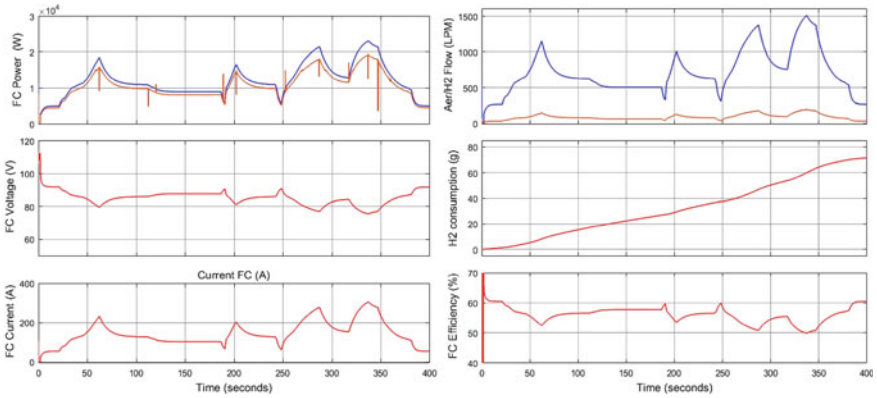


Fig. 14 Variation of fuel cell system parameters for an EUDC driving cycle

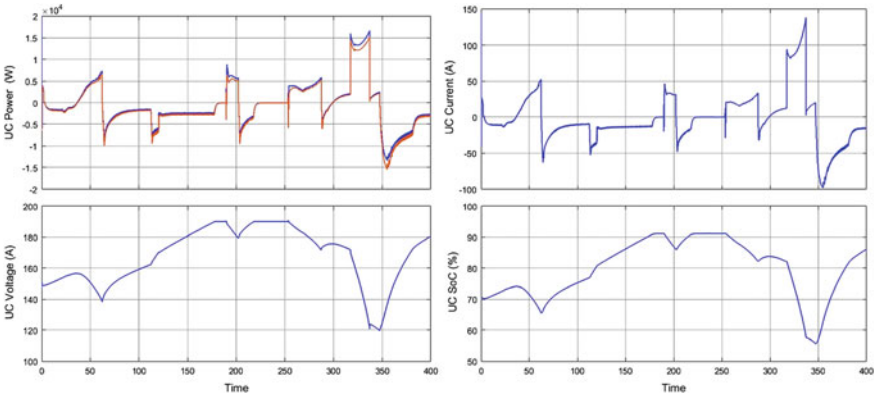


Fig. 15 Variation of ultracapacitor bank parameters for an EUDC driving cycle

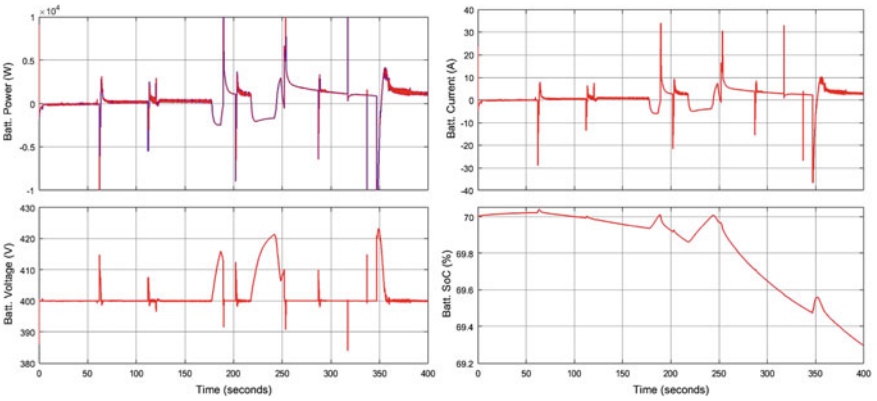


Fig. 16 Variation of battery system parameters for an EUDC driving cycle

Fig. 17 Experimental installation for FCHEV testing

5 Analysis of Experimental Results in Respect to Electrical Vehicle Efficiency

Beyond the advantage that FCHEV has almost zero greenhouse gas emissions, the use of fuel cells in hybrid vehicles needs to be improved in terms of efficiency of hydrogen consumption and increased component life. The electric efficiency of FCHEV was calculated experimentally and was 51% more than the efficiency reported in [48]. From the experimental tests, it was observed that the FC operation was more constant, for example in Fig. 18 it is observed that during the time interval 50–450 FC it operated at a constant power of 4.4 kW although the load power fluctuated according to profile required. The difference in power produced by the FC was injected into the battery and the ultracapacitor. For the high-speed range (120 km/h) the FC power increased with the 180 W/s growth rate controlled by EMS. This increase is imposed by the controller to keep the voltage on the DC bus as stable as possible, thus increasing the battery life.

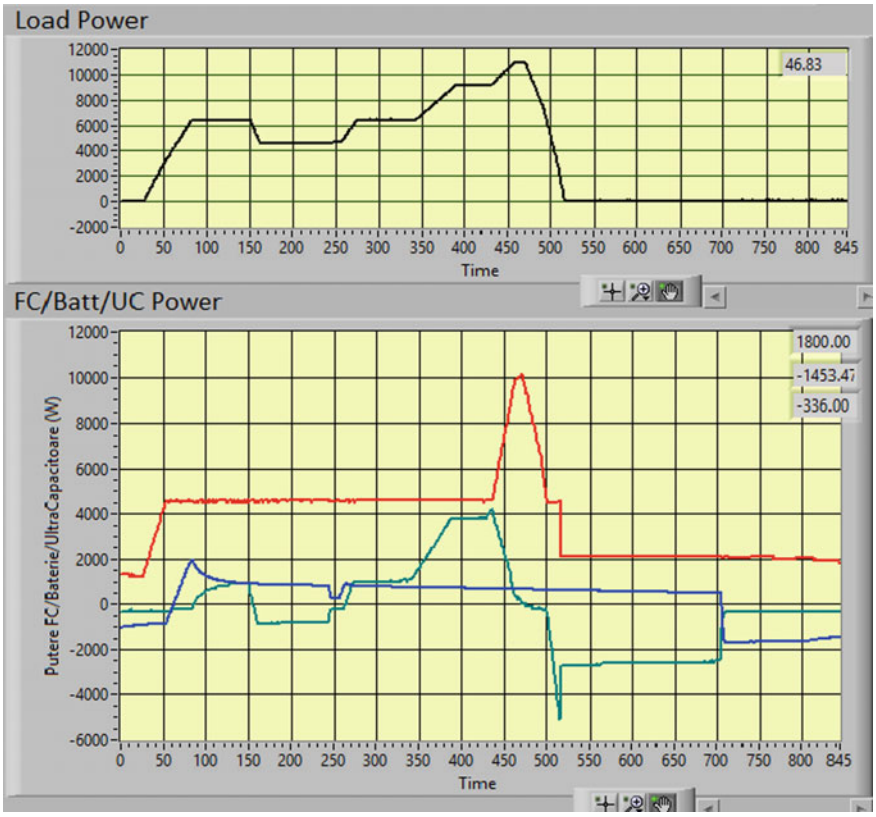


Fig. 18 LabVIEW capture for variation of the FC (red), battery (blue) and ultracapacitor (green) powers for an EUDC driving cycle (real experimental tests)

6 Conclusion

In this chapter, a hybrid electric vehicle powered by fuel cells was analysed, as the main source of power. The auxiliary power source was composed of an ultracapacitor bank and a battery system connected directly to the DC bus. FC and UC were connected to the bus via DC/DC converters. The components were described and found mathematical models as accurate as possible to capture all the power transitions.

The energy management strategy was implemented using a fuzzy controller. The EMS design aimed to find the optimal power references for FC and UC for which they must reduce hydrogen consumption and believe that the electrical efficiency of the system. The electrical efficiency of the system was increased to 51%, and the hydrogen consumption was 1.75 kg per 100 km. As a perspective of this chapter, further research is planned to install and test the FC system on a commercial vehicle.

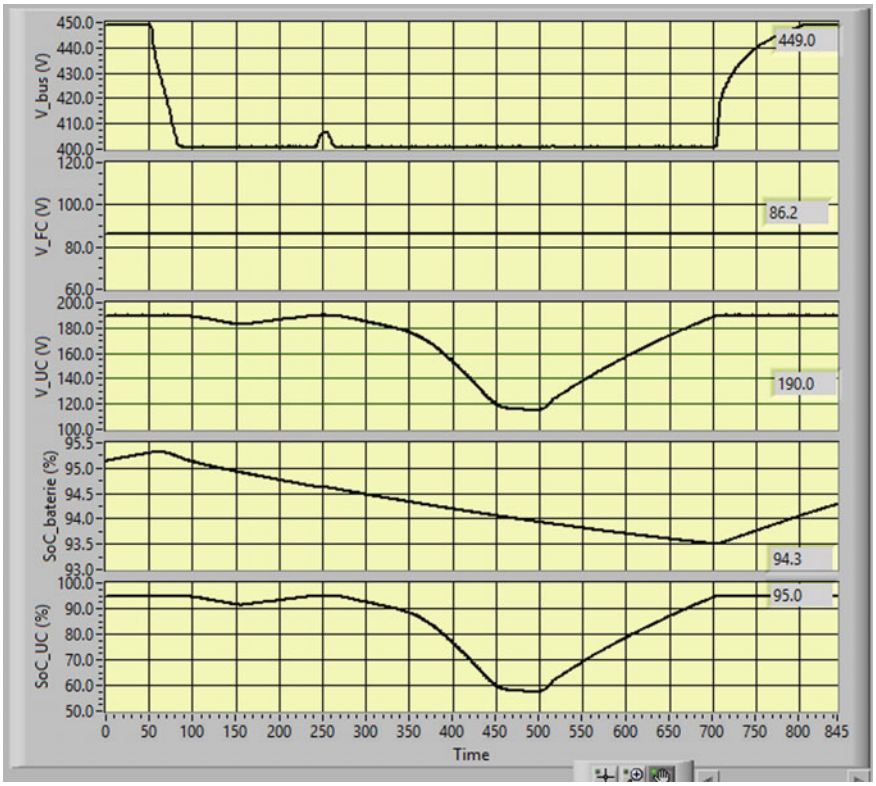


Fig. 19 Experimental results: FC, UC and battery voltage response; battery and UC state of charge response (capture LabVIEW)

Acknowledgements This work has been funded by the National Agency of Scientific Research from Romania by the National Plan of R&D, Project PN 19 11 02 02, PN-III-P-1.2-PCCDI-2017-0194/25 PCCDI and contract 117/2016.

References

1. Sharaf OZ, Orhan MF (2014) An overview of fuel cell technology: fundamentals and applications. Renew Sustain Energy Rev 32:810–853. https://doi.org/10.1016/j.rser.2014.01.012
2. Pei P, Chen H (2014) Main factors affecting the lifetime of proton exchange membrane fuel cells in vehicle applications: a review. Appl Energy 125:60–75. https://doi.org/10.1016/j.apenergy.2014.03.048
3. Olateju B, Monds J, Kumar A (2020) Large scale hydrogen production from wind energy for the upgrading of bitumen from oil sands. Appl Energy 118:48–56. https://doi.org/10.1016/j.apenergy.2013.12.013

4. Feng X, Wang L, Min S (2009) Industrial emergy evaluation for hydrogen production systems from biomass and natural gas. Appl Energy 86:1767–1773. https://doi.org/10.1016/j.apenergy.2008.12.019
5. Raceanu M, Marinoiu A, Culcer M, Varlam M, Bizon N (2014) Preventing reactant starvation of a 5 kW PEM fuel cell stack during sudden load change. In: Proceedings of the 2014 6th international conference on electronics, computers and artificial intelligence, vol 2015, pp 55–60. https://doi.org/10.1109/ecai.2014.7090147
6. Raceanu M, Iliescu M, Culcer M, Marinoiu A, Varlam M, Bizon N (2015) Fuelling mode effect on a PEM fuel cell stack efficiency. Prog Cryog Isot 18:15–24
7. Marinoiu A, Gatto I, Raceanu M, Varlam M, Moise C, Pantazi A et al (2017) Low cost iodine doped graphene for fuel cell electrodes. Int J Hydrogen Energy 42:26877–26888. https://doi.org/10.1016/j.ijhydene.2017.07.036
8. Iliescu M, Raceanu M, Culcer M, Enache A, Varlam M (2017) Fuel cell based powertrain simulations to find the power splitting leading to improved characteristics. Prog Cryog Isot 20:63
9. Aschilean I, Varlam M, Culcer M, Iliescu M, Raceanu M, Enache A et al (2018) Hybrid electric powertrain with fuel cells for a series vehicle. Energies 11:1294. https://doi.org/10.3390/en11051294
10. Bizon N, Oproescu M, Raceanu M (2015) Efficient energy control strategies for a standalone renewable/fuel cell hybrid power source. Energy Convers Manag 90:93–110. https://doi.org/10.1016/j.enconman.2014.11.002
11. Bizon N, Raceanu M (2017) Energy efficiency of PEM fuel cell hybrid power source. In: Bizon N, Mahdavi Tabatabaei N, Blaabjerg F, Kurt E (eds) Energy harvesting and energy efficiency technologies, methods, and their applications, vol 37. Springer International Publishing, Cham, pp 371–391. https://doi.org/10.1007/978-3-319-49875-1_13
12. Ettihir K, Boulon L, Agbossou K (2016) Optimization-based energy management strategy for a fuel cell/battery hybrid power system. Appl Energy 163:142–153. https://doi.org/10.1016/j.apenergy.2015.10.176
13. Yue M, Jemei S, Gouriveau R, Zerhouni N (2019) Review on health-conscious energy management strategies for fuel cell hybrid electric vehicles: degradation models and strategies. Int J Hydrogen Energy 44:6844–6861. https://doi.org/10.1016/j.ijhydene.2019.01.190
14. Wu J, Yuan XZ, Martin JJ, Wang H, Zhang J, Shen J et al (2008) A review of PEM fuel cell durability: degradation mechanisms and mitigation strategies. J Power Sources 184:104–119. https://doi.org/10.1016/j.jpowsour.2008.06.006
15. Sulaiman N, Hannan MA, Mohamed A, Ker PJ, Majlan EH, Wan Daud WR (2018) Optimization of energy management system for fuel-cell hybrid electric vehicles: issues and recommendations. Appl Energy 228:2061–2079. https://doi.org/10.1016/j.apenergy.2018.07.087
16. Sikkabut S, Mungporn P, Ekkaravarodome C, Bizon N, Tricoli P, Nahid-Mobarakeh B et al (2016) Control of high-energy high-power densities storage devices by Li-ion battery and supercapacitor for fuel cell/photovoltaic hybrid power plant for autonomous system applications. IEEE Trans Ind Appl 52:4395–4407. https://doi.org/10.1109/TIA.2016.2581138
17. Thounthong P, Raël S, Davat B et al (2009) Energy management of fuel cell/battery/supercapacitor hybrid power source for vehicle applications. J Power Sources 193:376–385. https://doi.org/10.1016/j.jpowsour.2008.12.120
18. Sulaiman N, Hannan MAA, Mohamed A, Majlan EHH, Wan Daud WRR (2015) A review on energy management system for fuel cell hybrid electric vehicle: issues and challenges, vol 52. Elsevier Ltd. https://doi.org/10.1016/j.rser.2015.07.132
19. Pollet BG, Staffell I, Lei J, Shang JL (2012) Current status of hybrid, battery and fuel cell electric vehicles: from electrochemistry to market prospects. Electrochim Acta 84:235–249. https://doi.org/10.1016/j.electacta.2012.03.172
20. Bubna P, Brunner D Jr JJG, Advani SG, Prasad AK (2010) Analysis, operation and maintenance of a fuel cell/battery series-hybrid bus for urban transit applications 195:3939–3949. https://doi.org/10.1016/j.jpowsour.2009.12.080

21. Marinoiu A, Raceanu M, Carcadea E, Varlam M, Stefanescu I (2017) Low cost iodine intercalated graphene for fuel cells electrodes. Appl Surf Sci 424:93–100. https://doi.org/10.1016/j.apsusc.2017.01.295
22. Marinoiu A, Carcadea E, Raceanu M, Varlam M (2018) Iodine doped graphene for enhanced electrocatalytic oxygen reduction reaction in PEM fuel cell applications. Adv Hydrog Gener Technol (InTech). https://doi.org/10.5772/intechopen.76495
23. Marinoiu A, Raceanu M, Carcadea E, Varlam M (2018) Iodine-doped graphene—catalyst layer in PEM fuel cells. Appl Surf Sci 456:238–245. https://doi.org/10.1016/j.apsusc.2018.06.100
24. Marinoiu A, Raceanu M, Andrulevicius M, Tamuleviciene A, Tamulevicius T, Nica S et al (2019) Low-cost preparation method of well dispersed gold nanoparticles on reduced graphene oxide and electrocatalytic stability in PEM fuel cell. Arab J Chem 13:3585–3600. https://doi.org/10.1016/j.arabjc.2018.12.009
25. Marinoiu A, Jianu C, Cobzaru C, Raceanu M, Capris C, Soare A et al (2017) Facile synthesis of well dispersed au nanoparticles on reduced graphene oxide. Prog Cryog Isot 20:5–14
26. Marinoiu A, Teodorescu C, Carcadea E, Raceanu M, Varlam M, Cobzaru C et al (2015) Graphene-based materials used as the catalyst support for PEMFC applications. Mater Today Proc (Elsevier Ltd) 2:3797–3805. https://doi.org/10.1016/j.matpr.2015.08.013
27. Marinoiu A, Raceanu M, Carcadea E, Pantazi A, Mesterca R, Tutunaru O et al (2018) Noble metal dispersed on reduced graphene oxide and its application in PEM fuel cells. In: Raceanu M (ed) Electrocatalysts for fuel cells and hydrogen evolution—theory to design. IntechOpen, Rijeka, p Ch. 4. https://doi.org/10.5772/intechopen.80941
28. Marinoiu A, Raceanu M, Carcadea E, Varlam M, Soare A, Stefanescu I (2017) Doped graphene as non-metallic catalyst for fuel cells. Medziagotyra 23. https://doi.org/10.5755/j01.ms.23.2.16216
29. Marinoiu A, Cobzaru C, Carcadea E, Raceanu M, Atkinson I, Varlam M (2015) An experimental approach for finding low cost alternative support material in PEM fuel cells. Rev Roum Chim 61:433–440
30. Wilberforce T, El-hassan Z, Khatib FN, Al A, Baroutaji A, Carton JG et al (2017) Sciencedirect developments of electric cars and fuel cell hydrogen electric cars. Int J Hydrogen Energy 42:25695–25734. https://doi.org/10.1016/j.ijhydene.2017.07.054
31. Briguglio N (2011) Fuel cell hybrid electric vehicles. In: Andaloro L (ed). IntechOpen, Rijeka, p Ch. 6. https://doi.org/10.5772/18634
32. Toyota (2016) 2016 Mirai product information. Mirai Prod Inf 2016:1–3
33. Yoshida T, Kojima K (2015) Toyota MIRAI fuel cell vehicle and progress toward a future hydrogen society. Toyota Mot Corp 24:45–49. https://doi.org/10.1149/2.F03152if
34. Amjad S, Neelakrishnan S, Rudramoorthy R (2010) Review of design considerations and technological challenges for successful development and deployment of plug-in hybrid electric vehicles. 14:1104–1110. https://doi.org/10.1016/j.rser.2009.11.001
35. Bizon N (2013) Energy efficiency for the multiport power converters architectures of series and parallel hybrid power source type used in plug-in/V2G fuel cell vehicles. Appl Energy 102:726–734. https://doi.org/10.1016/j.apenergy.2012.08.021
36. Hannan MA, Hoque MM, Mohamed A, Ayob A (2017) Review of energy storage systems for electric vehicle applications: issues and challenges. Renew Sustain Energy Rev 69:771–789. https://doi.org/10.1016/j.rser.2016.11.171
37. Liu C, Li F, Ma L-P, Cheng H-M (2010) Advanced materials for energy storage. Adv Mater 22:E28–E62. https://doi.org/10.1002/adma.200903328
38. Ahmadi S, Bathaee SMT, Hosseinpour AH (2018) Improving fuel economy and performance of a fuel-cell hybrid electric vehicle (fuel-cell, battery, and ultra-capacitor) using optimized energy management strategy. Energy Convers Manag 160:74–84. https://doi.org/10.1016/j.enconman.2018.01.020
39. Garcia P, Torreglosa JP, Fernandez LM, Jurado F (2013) Control strategies for high-power electric vehicles powered by hydrogen fuel cell, battery and supercapacitor. Expert Syst Appl 40:4791–4804. https://doi.org/10.1016/j.eswa.2013.02.028

40. Fathabadi H (2018) Novel fuel cell/battery/supercapacitor hybrid power source for fuel cell hybrid electric vehicles. Energy 143:467–477. https://doi.org/10.1016/j.energy.2017.10.107
41. Bizon N (2019) Sensitivity analysis of the fuel economy strategy based on load-following control of the fuel cell hybrid power system. Energy Convers Manag 199:111946. https://doi.org/10.1016/j.enconman.2019.111946
42. Carcadea E, Varlam M, Marinoiu A, Raceanu M, Ismail MSS, Ingham DBB (2019) Influence of catalyst structure on PEM fuel cell performance—a numerical investigation. Int J Hydrogen Energy 44:12829–12841. https://doi.org/10.1016/j.ijhydene.2018.12.155
43. Carcadea E, Varlam M, Ismail M, Ingham DB, Marinoiu A, Raceanu M et al (2019) PEM fuel cell performance improvement through numerical optimization of the parameters of the porous layers. Int J Hydrogen Energy. https://doi.org/10.1016/j.ijhydene.2019.08.219
44. Larminie AJ and AD (2002) Fuel cell systems explained
45. Ren G, Ma G, Cong N (2015) Review of electrical energy storage system for vehicular applications. Renew Sustain Energy Rev 41:225–236. https://doi.org/10.1016/j.rser.2014.08.003
46. Johnson VH (2002) Battery performance models in ADVISOR. J Power Sources 110:321–329. https://doi.org/10.1016/S0378-7753(02)00194-5
47. Carter R, Cruden A, Hall PJ, Zaher AS (2012) An improved lead-acid battery pack model for use in power simulations of electric vehicles. IEEE Trans Energy Convers 27:21–28. https://doi.org/10.1109/TEC.2011.2170574
48. Williamson SS, Lukic SM, Emadi A (2006) Comprehensive drive train efficiency analysis of hybrid electric and fuel cell vehicles based on motor-controller efficiency modeling. IEEE Trans Power Electron 21:730–740. https://doi.org/10.1109/TPEL.2006.872388

Finite Element Solutions for Magnetic Shielding Power Applications

Dumitru Cazacu, Elena Otilia Virjoghe, Valeriu Manuel Ionescu, and Stefan Castravete

Abstract In this chapter are presented some aspects concerning the finite element analysis of magnetic shielding for power applications. The investigation describes the physical mechanisms of magnetic shielding the magnetic field in a cylindrical shield using magnetic scalar potential and magnetic vector potential. A variational and a Galerkin finite element formulation are described. The mitigation of an OHTL magnetic field inside a shielded building placed near it is evaluated in the case study of this chapter.

Keywords Low frequency magnetic field · Power application · Finite element · Passive magnetic shielding

Symbols and Acronyms

OHTL	Overhead transmission lines
UGTC	Underground transmission cables
MV/LV	Medium voltage/low voltage
CAE	Computer aided engineering
FEM	Finite Element Method
SE	Shielding effectiveness

D. Cazacu (✉) · V. M. Ionescu
University of Pitesti, Pitesti, Romania
e-mail: dumitru.cazacu@upit.ro

V. M. Ionescu
e-mail: valeriu.ionescu@upit.ro

E. O. Virjoghe
University Valahia, Targoviste, Romania
e-mail: elena.virjoghe@valahia.ro

S. Castravete
CAELYNX-Europe, Craiova, Romania
e-mail: scastravete@caelynx.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_27

MSP V_m	Magnetic scalar potential V_m
MVP A	Magnetic potential vector A
δ	Skin depth
$\omega = 2\pi f$	Angular frequency
μ_0	Vacuum magnetic absolute permeability
μ_r	Relative magnetic permeability
σ	Electric conductivity
H	Magnetic field strength
E	Electric field strength
k	Propagation constant
B	Magnetic flux density
J	Current density
SLF	Super low frequency
PDE	Partial Differential Equation
Ni	Shape functions

1 Introduction

The electromagnetic pollution progressed a lot in the last century. The diversity of the electric and electronic systems has evolved in a tremendous way. Those devices operate in different frequency ranges, concentrated on specialized spectral domains.

Some of the most common sources of low frequency magnetic fields are the overhead transmission lines (OHTL), underground transmission cables (UGTC), medium voltage/low voltage (MV/LV) substations and building's electrical distribution systems. Their magnetic fields can generate electromagnetic compatibility problems caused the interference that affects technical features of the electrical and electronic devices and also can represent potential hazards for the human health.

In order to evaluate the possibilities of mitigating the magnetic field of these equipments and to compare the results with the reference values proposed by the international scientific institutions some methods were proposed.

In [1] certain intrinsic and extrinsic methods are described. The first category is referring to modifying the geometrical and electrical parameters of the magnetic field source: layout and compaction, distance management, phase splitting and phase cancellation. The second one comprises passive and active techniques depending on the way in which they attenuate the magnetic field.

The passive solutions refer to the attenuation systems that are located near the source of the magnetic field or in the neighborhood of the protected region.

Passive solutions include conductive and/or ferromagnetic shields and passive loops.

Active solutions use external devices that generates magnetic field that attenuate the incident magnetic field, having the same magnitude, phase and frequency. They are used in so called active loops. These are complex and expensive devices that

monitors and adjust the necessary current in order to obtain the proper counterpart magnetic field [2].

In order to design and evaluate the effectiveness of the attenuation solution generally computer aided engineering (CAE) software tools are used [3, 4].

Analytical solutions exist only for simple geometries and homogenous media.

Those programs use different numerical techniques in order to solve the specific forms of Maxwell equations that describe the operating of the considered systems. The most used methods are the finite difference method, the finite element method, the boundary element method and the transmission line matrix method.

One of the most versatile numerical methods used to simulate electromagnetic fields is the finite element method.

Modeling and simulation of the magnetic shields for power applications by the finite element method is widely described in the literature.

In [5] a precise shielding factor computation, based on the Finite Element Method (FEM) combined with the Jiles-Atherton model is presented, considering random disturbances. In paper [6] the 3D magnetic field of the reactor is computed by the method of edge-node finite element coupling.

In [7, 8] some numerical problems for magnetostatic and time harmonic 3D magnetic shields are presented in [9] different types of magnetic potentials are used for computing 3D magnetostatic shields effectiveness.

The effectiveness of the multilayer magneto static and time harmonic shields is evaluated using finite element method in [10, 11].

In [12] one and two shells magneto static cylindrical shields were considered. Applying the interface conditions at the discontinuity surfaces between the media, the symbolic algebraic systems of equations for those configurations were obtained and solved. Structure of the system's matrices were analyzed and their sparsity patterns were visualized. Analytical and numerical transversal effectiveness of the cylindrical shields were computed for different lengths.

Some numerical and experimental aspects concerning the electromagnetic shielding in microwave range are presented in [13, 14].

Measured values of the electrical field for an open type air substation are compared with numerical results obtained by finite element program Ansys in [15].

One of the new trends in computational electromagnetics is using the multicore machines in order to improve the execution time. Aspects concerning this topic applied for Comsol Multiphysics on a 3D magnetostatic problem are available in [16–17].

In [18] analytic and numerical comparison between the magnetic stored energy in cylindrical and toroidal coils considering the steady state superconducting state is considered.

Numerical shielding solutions concerning the underground power cables, using FEMM finite element software, are described in [19, 20]. In [21] a combination of theoretical analysis and numerical simulations with the finite element method is used to analyze the shielding properties of a passive and active shield developed for a SERF co-magnetometer application.

The progresses in the development of CAE programs that use finite element method are presented in many scientific papers.

In [22] certain recent trends in computational electromagnetics for defense applications are presented. Numerical aspects referring to an a posteriori error estimate of weak Galerkin (WG) finite element methods that can be applied to polygonal meshes and to an ill-posed elliptic Cauchy problems are described in [23, 24].

In this chapter recent trends of using finite element method for the modeling and simulation of passive shielding for low frequency field considering power applications are described.

The structure of the chapter is as follows.

In Sect. 1 is the introduction. Section 2 is entitled Time harmonic magnetic shields. It has two sub sections: Basics magnetic shielding mechanisms and the factors that determine the shielding effectiveness (SE) are described in Sect. 2.1. Solutions of the transversal time harmonic magnetic field in cylindrical shields are obtained in Sects. 2.2.1 and 2.2.2 using the magnetic scalar potential (MSP) and the magnetic potential vector (MVP), respectively. In both cases the shielding effectiveness (SE) was computed and analyzed.

The finite element method is described in Sect. 3, using two formulations.

A variational approach is presented for the magneto static field in Sect. 3.1 and the Galerkin approach is used for the time harmonic magnetic field in Sect. 3.2.

In Sect. 4 a case study that evaluates de magnetic protection of a partially shielded building to the magnetic field produced by an OHTL is presented, using the Ansys finite element software.

The distribution of the magnetic shield inside the building, with and without the shield is described and the maximum values are compared with the reference values indicated by the international commissions.

2 Time Harmonic Magnetic Shields

2.1 Basics of Magnetic Shielding Mechanisms

This type of shields is used to protect some volumes against time variable magnetic fields.

There are shields with closed and open geometries. The first type includes those shields that separate completely the source field and the protected region: the infinite extended plane, the infinite extended cylindrical and the spherical shields [2, 25–26]. For this type the magnetic field occurs in the shielded region by penetration.

For opened geometries (finite length plane shields and finite length cylindrical shields) out of the penetration mechanism appears also the flux leakage.

Two different mechanisms are present in the shielding of low frequency magnetic fields: magnetic flux shunting and magnetic flux attenuation by eddy currents.

The first mechanism is used also for dc magnetic fields [27].

The magnetic flux flows mainly through high permeability magnetic materials, that attract and shunt the flux lines. As a consequence, the magnetic flux lines are concentrated in the shield and avoid the protected region. The shield behaves like a magnetic field concentrator. The shielding effectiveness (SE) depends on the material magnetic permeability, on the geometry (ratio of the thickness over the external diameter multiplied with the magnetic relative permeability), on the shape of the shield and on the thickness of it [25–26, 28, 29].

In the second case the shielding effect is based on the generation of eddy currents in the metallic shells of the enclosure. Eddy currents are induced only in electrically conducting materials. The eddy currents create magnetic fluxes that oppose to the variation of the incident magnetic fluxes. As a consequence, the incident magnetic field is rejected in the neighborhood of the shield [25, 30, 29]. This phenomenon is encountered only in ac magnetic field.

The mechanism is valid no matter what is the value of the magnetic permeability.

Anyway, for materials that have high magnetic permeability the shielding is very efficient. The level of shielding depends on the material permeability and conductivity, on the geometry (ratio of the thickness over the external diameter multiplied with the magnetic relative permeability), on the thickness of the shield and on the frequency of the field source [27].

The eddy current density depends on the frequency of the incident field and on the electric conductivity of the material.

In Fig. 1 the distribution of the magnetic field lines is compared considering two types of magnetic shields: a magneto static cylindrical ferromagnetic shield and a time harmonic magnetic cylindrical shield. The thickness of the shield is 3 mm and the external diameter is 56 mm. The relative magnetic permeability is $\mu_r = 1000$ the electric conductivity $\sigma\ \sigma = 1.45 \cdot 10^6$ S/m and is 50 Hz.

For electromagnetic shields there is an important parameter, called skin depth, which influences the effectiveness of the shield. The expression of it is:

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \mu_r \sigma}} \tag{1}$$

where ω is the angular frequency, μ_r and μ_0 are the relative and absolute magnetic permeability respectively and σ is the electric conductivity of the shield material.

The magnetic field decays in the shield over a distance of few skin depth lengths.

For the shield effectiveness is important the ratio between the thickness and the skin depth g/δ.

If $g \neq \delta$ high values of shield effectiveness could be obtained. If $g = \delta$ the eddy currents have a uniform distribution in the shield thickness and the effectiveness of the shield is less than in the previous case.

For the shields that attenuate the incident field by the mechanism of the eddy current a good effectiveness can be obtained by increasing the maximum dimension, such as the diameter, for a constant thickness.

For the flux shunting mechanism, the increasing of the diameter decreases the SE.

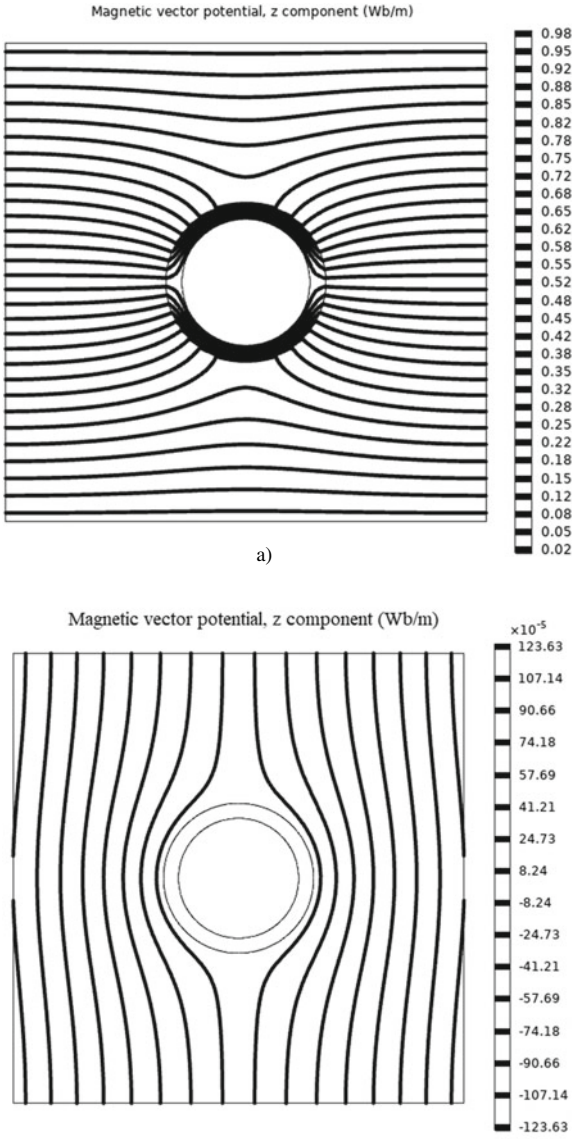


Fig. 1 **a** Magneto static shield as a magnetic field concentrator. **b** Rejection of the magnetic flux lines

In order to increase the SE of the magnetic shields at low frequencies materials with high permeability should be used. But they are expensive and heavy. Mu-Metal and Permalloy are examples of alloys with permeability of up to 100,000, by comparison with ordinary steel that has only few thousands.

A solution to avoid the drawbacks of using those materials is to use a multilayer geometry, composed of combination of ferromagnetic and non ferromagnetic

materials separated by a layer of air. The thickness of the air influences directly the shield effectiveness. Also, two layer shields composed of steel-copper are very efficient against variable magnetic fields. The copper, having high electric conductivity, generates high reflexivity and eddy currents. The steel having high magnetic permeability is proper for low frequencies [21, 31, 32].

One of the new trends for shielding materials is the usage of materials with nonconventional electromagnetic properties, e.g. epsilon-, mu- and index-near-zero metamaterials. They are artificial materials with properties that are not found in nature, e.g. negative permeability. Those materials are manufactured from repetitive structures composed of composite materials such as metals and plastics. In [32] a longitudinal mu-near-zero metamaterial is used to shield quasi-stationary magnetic fields.

2.2 *The Cylindrical Shield in Transversal Time Harmonic Magnetic Field*

Solutions of the magnetic field equations in the cylindrical shield structure can be obtained using magnetic scalar potential (MSP) or magnetic vector potential(MVP). Both approaches will be presented in the next paragraphs.

2.2.1 Solution Using Magnetic Scalar Potential

The cylindrical and planar shields are mainly used to attenuate the SLF (super low frequency) magnetic field.

In this section a cylindrical shield subjected to a transversal time harmonic magnetic field is analyzed.

In Fig. 2 the cylindrical shield has an external diameter $2r_e$, an inner diameter $2r_i$ and the thickness of the shell is $g = r_{ext} - r_{\text{int}}$. The incident transversal magnetic field is $\overline{H_{ext}} = H_{ext}\overline{j}$, oriented in the Oy direction. The shield is considered to be very long and the end effects are neglected.

From the magnetic point of view there are three domains of interest: outside the shield (region 1), inside the shield wall (region 2) and in the interior of the shield (region 3).

(a) **The electromagnetic field equations in domain 1** $r > r_e$

In this domain there are no current densities. The magnetic field strength can be expressed using the MSP [25–26, 33–34]:

$$rot\underline{\overline{H}} = 0 \tag{2}$$

$$\underline{\overline{H}} = -grad\, V_m \tag{3}$$

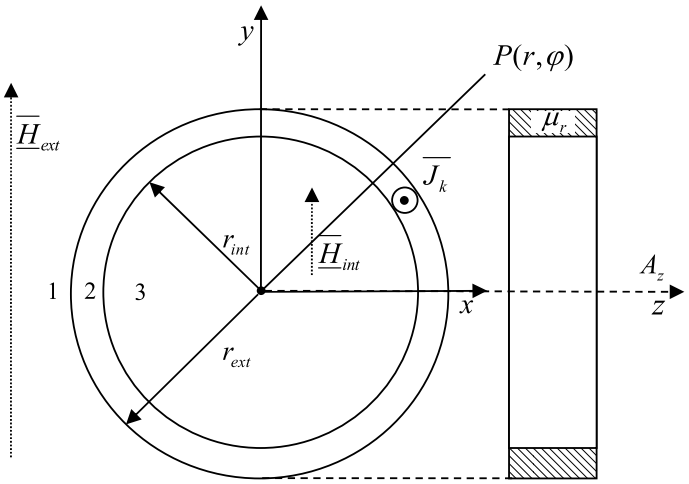


Fig. 2 Cylindrical shield subjected to a transverse time harmonic magnetic field

where V_m is the magnetic scalar potential.

Because $div\ \underline{H} = 0$ it follows that:

$$div\ (\text{grad}\ V_m) = \Delta V_m = 0 \tag{4}$$

In cylindrical coordinates, neglecting the end effects, expression (4) becomes:

$$\Delta V_m = \frac{\partial^2 V_m}{\partial r^2} + \frac{1}{r} \cdot \frac{\partial V_m}{\partial r} + \frac{1}{r^2} \cdot \frac{\partial^2 V_m}{\partial \varphi^2} = 0 \tag{5}$$

In order to solve Eq. 5 the method of separation of variables is used [30, 33–34]:

$$V_m(r, \varphi) = R(r) \cdot \phi(\varphi) \tag{6}$$

The solutions obtained for R and ϕ are:

$$R(r) = C_1 \cdot r^m + C_2 \cdot r^{-m} \tag{7}$$

$$\phi(\varphi) = B_1 \cdot \cos(m\varphi) + B_2 \cdot \sin(m\varphi) \tag{8}$$

and the general expression of the magnetic scalar potential is:

$$V_m(r, \varphi) = \sum_{m=1}^{\infty} \left(C_1 \cdot r^m + C_2 \cdot r^{-m}\right)(B_1 \cos(m\varphi) + B_2 \sin(m\varphi)) \tag{9}$$

After considering the behaviour of the magnetic field far away from the shield, the MSP expression (8) in the domain 1 becomes:

$$V_{m1} = -\left(\underline{H}_{ext} r + C_2 r^{-1}\right) \sin\varphi \tag{10}$$

The magnetic field strength H has the following components:

$$\underline{H_r} = -\frac{\partial V_m}{\partial r} = \left(\underline{H}_{ext} - \frac{C_2}{r^2}\right) \sin\varphi \tag{11}$$

$$\underline{H_\varphi} = -\frac{1}{r} \cdot \frac{\partial V_m}{\partial \varphi} = \left(\underline{H}_e + \frac{C_2}{r^2}\right) \cos\varphi \tag{12}$$

(b) **The electromagnetic field equations in domain 2** $r_{\text{int}} < r < r_{ext}$

In this domain with electrical conductivity the electromagnetic Helmholtz equationdescribes the field [30, 26, 33–35]:

$$\Delta\underline{E} - k^2\underline{E} = 0;\ k^2 = j\omega\mu\sigma \tag{13}$$

where k is the propagation constant for metals.

Because the electric field and the current density are oriented in the 0 z direction we rewrite Eq. 13 in cylindrical coordinates, without vectorial notations:

$$\frac{\partial^2 \underline{E}}{\partial r^2} + \frac{1}{r}\frac{\partial \underline{E}}{\partial r} + \frac{1}{r}\frac{\partial^2 \underline{E}}{\partial \varphi^2} = k^2 \underline{E} \tag{14}$$

Using the separation of the variables the following equations are obtained:

$$\frac{d^2 \underline{E}}{dr^2} + \frac{1}{r}\frac{d\underline{E}}{dr} - \left(k^2 + \frac{n^2}{r^2}\right)\underline{E} = 0 \tag{15}$$

$$\frac{d^2 \underline{E}}{d\varphi^2} + n^2 \underline{E} = 0 \tag{16}$$

Their solutions are:

$$\underline{E}(r) = D_1 I_n(kr) + D_2 K_n(kr) \tag{17}$$

$$\underline{E}(\varphi) = D_3 \cos(n\varphi) + D_4 \sin(n\varphi) \tag{18}$$

We obtain the general solution of (13) as follows:

$$\underline{E}(r, \varphi) = \sum_{n=1}^{\infty} (D_1 I_n(kr) + D_2 K_n(kr))(D_3 \cos(n\varphi) + D_4 \sin(n\varphi)) \tag{19}$$

This solution comprises the modified Bessel functions, of order n, $I_n(kr)$ and $K_n(kr)$.

This solution is proper when the thickness of the shield $g = r_2 - r_1$ is significant by comparation with the radius r_{ext}.

For practical shields, the thickness verifies $g << r_i$.

If the condition $n^2/r^2 << |k^2|$ is fulfilled, then Eq. 15 becomes [30]:

$$\frac{d^2\underline{E}}{dr^2} + \frac{1}{r}\frac{d\underline{E}}{dr} - k^2\underline{E} = 0 \tag{20}$$

and have the solution:

$$\underline{E}(r) = M_1 e^{kr} + M_2 e^{-kr};\ k = \sqrt{j\omega\mu\sigma} \tag{21}$$

The solution of electric field in the shield wall is:

$$\underline{E}(r, \varphi) = \left(M_1 e^{kr} + M_2 e^{-kr}\right)\cos\varphi \tag{22}$$

(c) **The electromagnetic field equations in domain 3**

In this domain the magnetic field is described by similar equations similar with (11) and (12) but without the terms C_2/r^2, because when $r \to 0$ the magnetic field goes to infinite [30, 26, 35, 36].

The expressions of the magnetic field components are:

$$\underline{H}_r = \underline{H}_{ext}\sin\varphi \tag{23}$$

$$\underline{H}_\varphi = \underline{H}_{ext}\cos\varphi \tag{24}$$

The shielding factor

In order to compute the shielding factor the interface conditions among those three domains are used. The shielding factor $\underline{SF}$ is a complex number and is defined by the ratio:

$$\underline{SF} = \underline{H_{\text{int}}}/\underline{H_{ext}} \tag{25}$$

Interface conditions

At the border between the first and the second domains, for $r = r_{ext}$, the normal component of the magnetic flux density and the tangential component of the magnetic field strength are preserved.

As a consequence the following expressions are obtained:

$$\underline{H}_{ext} - \frac{C_2}{r_{ext}^2} = \frac{1}{j\omega\mu_0 r_m}\left(M_1 e^{kr_{ext}} + M_2 e^{-kr_{ext}}\right) \tag{26}$$

$$\underline{H}_{ext} + \frac{C_2}{r_{ext}^2} = \frac{k}{j\omega\mu}\left(M_1 e^{k_{ext} r} - M_2 e^{-kr_{ext}}\right) \tag{27}$$

At the boundary between the second and the third domain, for $r = r_{\text{int}}$, from the same conditions the next expressions follows:

$$\underline{H}_{\text{int}} = \frac{1}{j\omega\mu_0 r_m}\left(M_1 e^{kr_{\text{int}}} + M_2 e^{-kr_{\text{int}}}\right) \tag{28}$$

$$\underline{H}_{\text{int}} = \frac{k}{j\omega\mu}\left(M_1 e^{kr_{\text{int}}} - M_2 e^{-kr_{\text{int}}}\right) \tag{29}$$

From relations (26) to (29) the following unknowns C_2, M_1, M_2, H_{int} are obtained. After the calculation, the expression of the shielding factor is obtained:

$$\underline{F}_e = \frac{\underline{H}_{\text{int}}}{\underline{H}_{ext}} = \frac{1}{\cosh kg + \frac{1}{2}\left(D + \frac{1}{D}\right)\sinh kg} \tag{30}$$

where k is the propagation constant in metals:

$$\underline{k} = \sqrt{j\omega\sigma\mu} = \frac{1+j}{\delta} \tag{31}$$

and D is:

$$D = \frac{kr_m}{\mu_r} \tag{32}$$

and $g = r_e - r_i$ is the thickness of the shield.

Because $g << r_i$ the following approximation have been used: $r_e \approx r_i = r_m$ The SE can be computed using the following expression:

$$SE = \left|\frac{1}{\underline{SF}}\right| = \sqrt{\text{Re}^2\left(1/\underline{F_e}\right) + \text{Im}^2\left(1/\underline{F_e}\right)} \tag{33}$$

The shield effectiveness SE increases with the thickness of the shield wall and with the frequency of the incident field. Ferromagnetic cylindrical shields have higher SE values, at the same thickness, magnetic permeability and frequency than those manufactured from non ferromagnetic materials.

2.2.2 Solution Using Magnetic Potential Vector

The same problem can be solved in terms of the magnetic vector potential A in a cylindrical reference coordinate system. The geometry is symmetrical and the

magnetic vector potential is oriented in the z direction and it is independent of the z coordinate. In domains 1 and 3 of the cylindrical shield, the magnetic vector potential verifies the Laplace equation and in domain 3 the shield verifies the diffusion equation [2, 30, 34]:

$$\nabla^2 A_z = 0, r \geq r_{ext} \tag{34}$$

$$\nabla^2 A_z - k^2 A_z = 0, r_{\text{int}} < r < r_{ext} \tag{35}$$

$$\nabla^2 A_z = 0, r \leq r_{\text{int}} \tag{36}$$

where

$$\nabla^2 A_z = (\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2})A_z \tag{37}$$

The current density is proportional to a $\cos\varphi$ factor so the magnetic field potential A_Z is expressed as a product of a radial function and $\cos\varphi$ [2, 36]. The following expressions are obtained [2, 36]:

$$A_z^1(r, \varphi) = \mu_0 H_{ext} \cos\varphi (r - \frac{c_1}{r}), r \geq r_{ext}, \tag{38}$$

$$A_z^2(r, \varphi) = \mu_0 H_{ext} \cos\varphi [c_2 I_1(kr) + c_3 K_1(kr)], r_{\text{int}} < r < r_{ext}, \tag{39}$$

$$A_z^3(r, \varphi) = \mu_0 H_{ext} c_4 r \cos\varphi, r \leq r_{\text{int}}, \tag{40}$$

where $I_1(\cdot)$ and $K_1(\cdot)$ are the first-order modified Bessel functions of the first and second kind, respectively. The unknown coefficients c_1, c_2, c_3 and c_4 can be determined using the boundary conditionsat the interface of the cylindrical shells, in a similar way as for the MSP. After solving the obtained algebraic linear system of equations, the unknown coefficients are determined. Using them the expressions of the magnetic vector potential in all of the three regions of interest can be obtained.

The shield effectiveness has the following expression [2]:

$$SE = \left| \frac{r_{\text{int}}}{2 r_{ext} \mu_r} \{[\mu_r K_1(k r_{int}) - k r_{int} K_1'(k r_{int})]\Big[\mu_r I_1(k r_{ext}) + k r_{ext} I_1'(k r_{ext})\Big] - [\mu_r I_1(k r_{int}) - k r_{int} I_1'(k r_{int})][\mu_r K_1(k r_{ext}) + k r_{ext} K_1'(k r_{ext})]\} \right| \tag{41}$$

where $I_1'(\cdot)$ and $K_1'(\cdot)$ are the first derivative of the first-order modified Bessel functions of the first and second kind, respectively.

The Eq. 41 has certain simpler forms for different conditions.

Considering the case when the radii of the shield are large compared to the skin depth (i.e. $r_{int}, r_{ext} = \delta$) a much simpler expression is obtained:

$$SE = \left| \begin{array}{l} \frac{\sqrt{r_{int}}}{8\mu_r k r_{ext}\sqrt{r_{ext}}}[k(r_{ext} + 8\mu_r r_{int} + 8\mu_r g)\cosh(kg) + \\ \left(kg + 4k^2 r_{int}^2 + k r_{int} + 4k^2 r_{int} g + 4\mu_r^2\right)\sinh(kg)] \end{array} \right| \tag{42}$$

For the case of a magnetic conducting thin shield, the following expression is obtained:

$$SE; \left|\cosh(\mathrm{kg}) + \frac{1}{2}\left(\frac{\mathrm{kr}_m}{\mu_\mathrm{r}} + \frac{\mu_\mathrm{r}}{\mathrm{kr}_m}\right)\sinh(\mathrm{kg})\right| \tag{43}$$

Considering the low-frequency approximation case, for the case of power applications ($|kg| = 1$) in (43), then the next expression is obtained:

$$SE; |1 + \frac{\mu_r}{2r_m} g + \frac{r_m}{2\mu_r} g k^2| \tag{44}$$

Considering the thick-shield approximation for the magnetic conducting cylindrical shell a different expression is derived, using the following approximations:

$$(|kg| = 1, r_m = \delta \text{ and } r_{ext}; r_{int} = r_0)$$

$$SE; |\frac{r_m k}{4\mu_r} e^{kg}| = \frac{r_m}{2\sqrt{2}\mu_r \delta} e^{g/\delta} \tag{45}$$

3 Finite Element Formulations for Magneto Static and Time Harmonic Magnetic Field

Problems of analysis of macroscopic electromagnetic field accept, out of the differential formulation, an equivalent variational approach.

In order to create the mathematical variational model of the electromagnetic field a variational principle (Lagrangian or Hemiltonian type) must be set. It must allow generating from a stationary condition of a certain functional (generally with a significance of electromagnetic energy) the fundamental equations of the electromagnetic field in material media [36, 37].

Numerical techniques convert the partial differential equations of the electromagnetic field in linear or non linear systems of algebraic equations. The solution of those systems generates an approximate solution of the electromagnetic

field in a discrete number of points of the field domain called nodes. Different approaches convert the PDEs that describe the field into a system of algebraic equations having as unknowns the values of the electric or magnetic potential. There different discretization techniques: finite difference method, finite element method, finite volume method, boundary element method.

The finite element method is recognized as a powerful and versatile method that can be used to a large class of engineering and mathematical problems, including those from electromagnetics.

The operation that named the method consists of discretization (mesh) of the 1D, 2D or 3D field domain into some 1D elements (line segments), 2D elements (triangles or quadrilaterals) and 3D elements (tetrahedrons, hexahedrons (bricks), as seen in Fig. 3a–d.

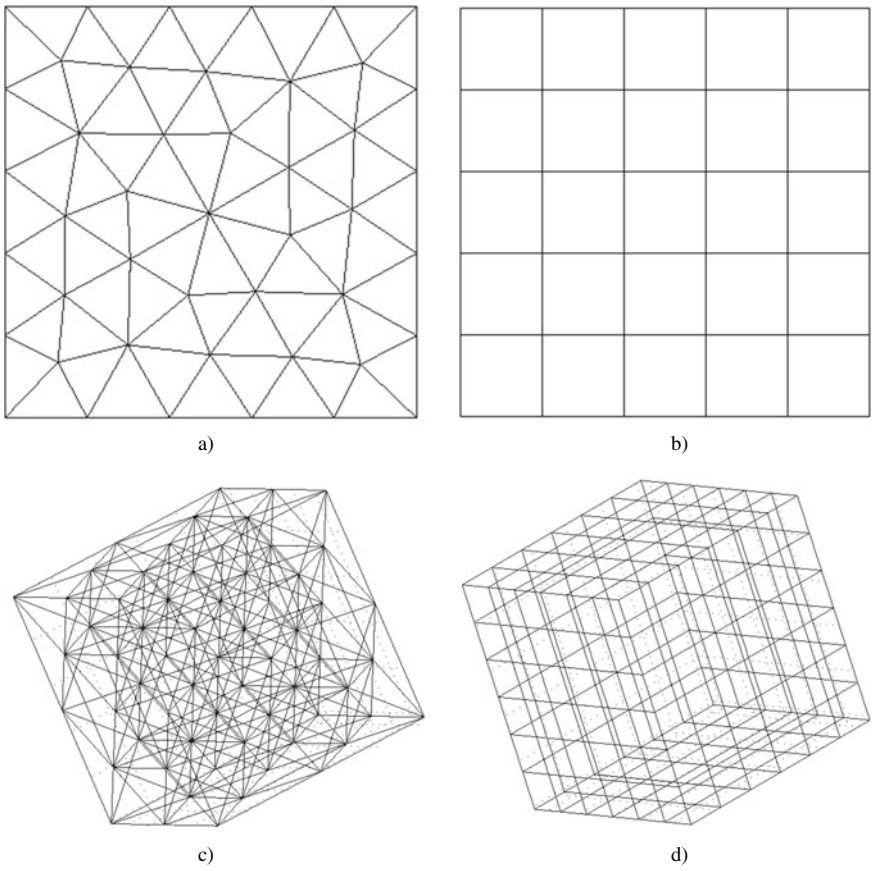


Fig. 3 Finite element discretizations **a** 2D structured mesh with squares **b** 2D unstructured mesh with triangle, **c** 3D non structured mesh with tetrahedrons, **d** 3D structured mesh with hexahedrons

Meshing with triangles and tetrahedron is called unstructured mesh. When quadrilaterals and hexahedrons (bricks) are used we are talking about structured mesh.

Then the unknown electromagnetic potentials at an elemental level are expressed as combinations of their values in the nodes of the mesh and a set of known functions called shape functions [36–38].

There are cases when the energy functional is very difficult to obtain or doesn't exist (generally for non-self adjointpartial differential equations PDEs). In those cases other techniques are used: the Ritz-Rayleigh and the Galerkin method.

They are not related with the minimization of a functional and can be used to solve directly the PDEs with boundary conditions.

The main computational steps that should be performed in order to solve a problem by the finite element method are:

- Discretizing the solution region into small sub domains called finite elements
- Obtaining the governing equation for a certain element (obtaining the element coefficient matrix)
- Assembling of all elements in the solution region (assembling all the elementary matrix into a global matrix, named stiffness matrix)
- Solving the resulting system of equations
- Post processing the results.

3.1 Finite Element Analysis of Magneto Static Field Using a Variational Formulation

We consider a plane-parallelmagnetostatic field, described by a Laplace equation, in 2D domain, having Dirichlet boundary conditionson the frontier Γ_2 [33–34, 36–39], (Fig. 4):

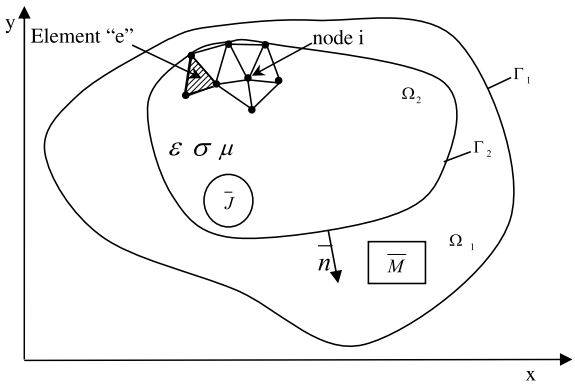


Fig. 4 Electromagnetic field domain

$$\nabla^2 A_z = 0 \tag{46}$$

where A_Z is the z component of the magnetic vector potential A, for this type of field.

Solving Eq. (46) is equivalent with obtaining the solution for a variational problem. It consists in finding the minimum of a certain functional, that represents the magnetic energy from the domain. This means that the distribution of the magnetic potential that satisfies the Laplace equation minimizes the magnetostatic energy from the domain.

The energy functional that corresponds to Eq. 46 has the following expression:

$$F(A) = \int_V \frac{1}{2\mu} \nabla^2 A_z dV \tag{47}$$

This equation presents the energy of the magnetic field from the respective domain. Solving the equation is simpler for plane-parallel or axisymmetricfields.

If we consider a parallel plane magnetic field in magnetostatic and assuming that the depth is in the Oz direction, the perpendicular on the field plane, equal with the unit, the functional. Using Eq. 47 can be written as a surface integral:

$$F(A) = \int_{S\Gamma_2} \frac{1}{2\mu} \nabla^2 \overline{A} dS \tag{47'}$$

Being applied for each element, this transformation decreases the computational costs.

The application of the variational approach will be described as follows, considering a general electromagnetic media, described by parameters ε, σ and μ with magnetization domain $\overline{M}$ and current sources$\overline{J}$ (Fig. 4).

In the first step, the domain Ω_2 is split in smaller sub domains, called finite elements, with triangular shapes (Fig. 4). These elements can have different sizes, depending on the domain configuration, with the mesh density increasing where the field variation rate is higher. Let m be the total number of finite elements and n the total number of resulting nodes. By dividing the domain in a finite element mesh the functional Eq. 47 can be rewritten as the sum of the functionals corresponding to the m finite elements [34, 36]:

$$\Im = \sum_{e=1}^{m} \frac{1}{2\mu_e} \int_{S_e} (\nabla A_{ze})^2 ds \tag{48}$$

Expression (48) indicates that the method can be used to solve the magnetic field problems in heterogeneous domains that occur in practical situations.

Different finite elements can have different permeabilities (μ_{re}), the magnetic medium in a single element being homogenous. Also relation (48) indicates that, for

a linear media, the sum of the magnetostatic field energy from every finite element equals the whole magnetostatic field energy from domain Ω_2.

On the other hand, the use of the Eq. (48) instead of (47) should assume the fulfilling of specific interface conditions for the magnetostatic field at the interface between the adjacent finite elements.

For the magnetostatic field this means the continuity of the magnetic vector potentialA and the preservation of the normal components of the magnetic flux density at the interface between medium 1 and 2:

$$(B_n)|_{\mu_{r1}} = (B_n)|_{\mu_{r2}} \tag{49}$$

Usually polynomial approximations are used (Lagrange, Hermite).

Considering the finite elements small enough it could be considered that the magnetic vector potential (MVP) $A(x, y)$ has a linear variation with x and y, and can be described by a first degree algebraic polynom.

In Fig. 5 a current triangular finite element e, described by the nodes i, j, k with the coordinates $P_i(x_i, y_i)$, $P_j\left(x_j, y_j\right)$ and $P_k(x_k, y_k)$ is presented. For a point P inside the element (e), the MVP $A_z(x, y)$ can be written as follows [33, 40, 36]:

$$A_{ze} = \alpha_1 + \alpha_2 \cdot x + \alpha_3 \cdot y \tag{50}$$

or as a matrix:

$$A_{ze} = \begin{bmatrix} 1 & x & y \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \tag{51}$$

For each element three coefficients should be computed.

Because the aim of the computing is to obtain the values of the magnetic potential in the nodes of the mesh then MVP A_{ze} should be expressed as a function of the magnetic potentials A_{zi}, A_{zj}, A_{zk} from the nodes i, j, k of the finite element "e", in

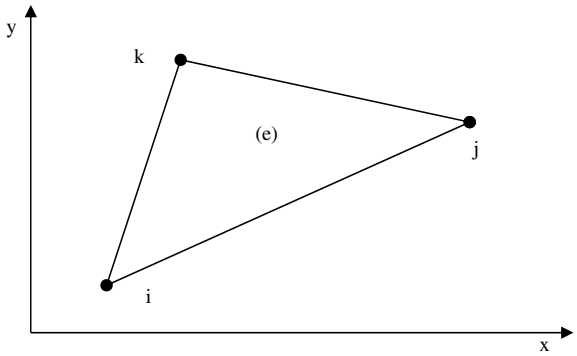


Fig. 5 First order triangular finite element

the following form:

$$A_{ze} = N_{ei} A_{zi} + N_{ej} A_{zj} + N_{ek} A_{zk} \tag{52}$$

or in a matrix form:

$$A_{ze} = \begin{bmatrix} N_{ei} & N_{ej} & N_{ek} \end{bmatrix} \cdot \begin{bmatrix} A_i \\ A_j \\ A_k \end{bmatrix} = [N] \cdot [A] \tag{53}$$

where the coefficients N_{ei}, N_{ej}, N_{ek} are linear functions of x and y, called shape functions. If the point P is in the node i the corresponding shape function has the value 1 and in the other is 0 $\left(N_{ej} = N_{ek} = 0\right)$.

This property can be described as $N_i(x_k, y_k) = \delta_{ik}$, where δ_{ik} is the symbol of Kronecker.

For two adjacent elements the value of the magnetic potential in the common nodes is equal. In order to improve the precision, higher degree polynomials can be used.

For example in Fig. 6, for each side another node has been added in a second order polynomials approach. The magnetic potential has the following expression:

$$A_e = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 x^2 + \alpha_5 xy + \alpha_6 y^2 \tag{54}$$

respectively:

$$A_e = N_{ei} A_i + N_{ej} A_j + N_{ek} A_k + N_{ea} A_a + N_{eb} A_b + N_{ec} A_c \tag{55}$$

For each element six coefficients should be computed. The precision is higher, but the execution time increases.

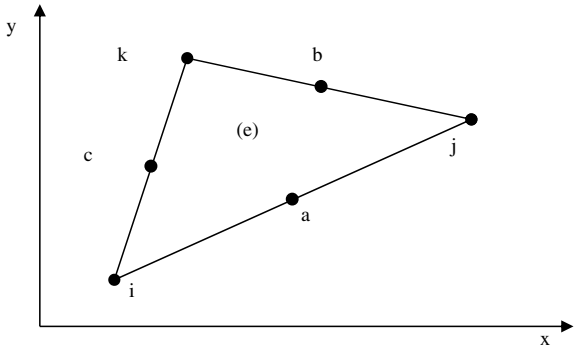


Fig. 6 Second order triangular finite element

In order to increase the accuracy of the finite element solution usually three strategies are used.

The first one is the h-version of the finite element that uses low polynomial degree (i.e. p = 1, 2) and increases the mesh density. With this approach the approximation error decreases algebraically versus the number of the unknowns.

The second technique is to maintain the mesh density and to increase the degree p of the approximation polynomial. This is the p version of the finite element method. For simpler cases, when the solution is smooth, this approach has an exponential convergence versus the number of unknowns. For the practical cases the convergence rate is algebraic. The exponential convergence can be reached again using the hp-version of the finite element method.

The shape functions will be determined for a linear approximation. Writing the Eq. 51 for nodes i, j, k the following matrix expression is obtained:

$$[A] = \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \tag{56}$$

Solving the system in Eq. 56 we obtain the coefficients $\alpha_1, \alpha_2, \alpha_3$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{bmatrix}^{-1} \cdot \begin{bmatrix} A_{zi} \\ A_{zj} \\ A_{zk} \end{bmatrix} \tag{57}$$

The vector of the shape functions is then:

$$[N_e] = \begin{bmatrix} 1 & x & y \end{bmatrix} \cdot \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{bmatrix}^{-1} \tag{58}$$

The determinant of the second matrix from the r.h.s. is equal with $2S_e$ where S_e is the area of the triangle e. The following relations for the shape functions are obtained [33, 40]:

$$\begin{aligned} N_{ei} &= \frac{1}{2S_e}\left(x_j y_k - x_k y_j\right) + \left(y_j - y_k\right)x + \left(x_k - x_j\right)y \\ N_{ej} &= \frac{1}{2S_e}(x_k y_i - x_i y_k) + (y_k - y_i)x + (x_i - x_k)y \\ N_{ek} &= \frac{1}{2S_e}\left(x_i y_j - x_j y_i\right) + \left(y_i - y_j\right)x + \left(x_j - x_i\right)y \end{aligned} \tag{59}$$

After doing the calculation the following expression for the gradientof the magnetic potential is obtained:

$$grad\, A_{ze} = \bar{i}\frac{\partial A_{ze}}{\partial x} + \bar{j}\frac{\partial A_{ze}}{\partial y} = \bar{i}\frac{\partial}{\partial x}\left(N_{ei}A_{zi} + N_{ej}A_{zj} + N_{ek}A_{zk}\right) + \bar{j}\frac{\partial}{\partial y}\left(N_{ei}A_{zi} + N_{ej}A_{zj} + N_{ek}A_{zk}\right) \tag{60}$$

If we take into account Eq. 59 then Eq. 60 becomes:

$$grad\, A_{ze} = \bar{i}\frac{1}{2S_e}\left[\left(y_j - y_k\right)A_{zi} + (y_k - y_i)A_{zj} + \left(y_i - y_j\right)A_{zk}\right] + \bar{j}\frac{1}{2S_e}\left[\left(x_k - x_j\right)A_{zi} + (x_i - x_k)A_{zj} + \left(x_j - x_i\right)A_{zk}\right] \tag{61}$$

Considering for each element that $|grad\, V| = const$ then the functional Eq. 48 becomes [33]:

$$\Im = \sum_{e=1}^{m} \frac{\mu_e S_e}{2} \cdot (\nabla A_{ze})^2 \tag{62}$$

and it be can be rewritten as:

$$\Im = \sum_{e=1}^{m} \frac{\mu_e}{8S_e}\left\{\begin{array}{l}\left[\left(y_j - y_k\right)A_{zi} + (y_k - y_i)A_{zj} + \left(y_i - y_j\right)A_{zk}\right]^2 \\ +\left[\left(x_k - x_j\right)A_{zi} + (x_i - x_k)A_{zj} + \left(x_j - x_i\right)A_{zk}\right]^2\end{array}\right\} \tag{63}$$

The functional Eq. 62 has been transformed into a function that has as variables the magnetic potentials in the n nodes of the field domain.

Minimization of it is obtained imposing in each nodal point i that the derivative in relation to A_{zi} to be 0:

$$\frac{\partial \Im}{\partial A_{zi}} = 0, \mathrm{i} = 1, 2, \ldots, \mathrm{n} \tag{64}$$

and a system with n algebraic equations that has n unknowns, the magnetic potentials in nodes i, is obtained:

$$[C][A] = [b] \tag{65}$$

If the boundary conditions are of Dirichlet type, given the potentials in the nodes of the domain frontier, the total number of variables is less than n, but even in this case it is equal to the number of equations.

In many engineering applications matrix C, the stiffness matrix, is symmetrical and sparse and band structured. The equation system is numerically solved, giving the potential values for the nodes. For 2D problems direct solvers are used and for 3D problems iterative solvers are indicated.

In the postprocessor module, for each finite element we can compute ($grad\ A_{z\lambda}$) as well as the field vector.

The finite element method is a discretization method for field computing, as is the finite differences method, which is suitable to be compared with. While the finite differences method is based on equation approximation with partial derivates by using finite differences equations, the finite element method approximates the potential function for a finite element. Both methods, as well as the boundary element method, lead to an algebraic system of equations where the unknowns are the potentials of the nodes. In the case of boundary elements method the resulting matrix is fully populated.

The finite element method has advantages in certain situations. Thus, it can be used for complex geometrical configurations, for linear and non linear PDEs, for coupled problems (electro-thermal, magneto-structuraland electromagnetic field electric circuits), varying materials and boundary conditions and anisotropicmaterials.

3.2 *Finite Element Galerk in Formulation a Time Harmonic Magnetic Field*

The time magnetic dependent field problems are also referred as eddy current problems. They occur in the cases of the electromagnetic shields.

There are many situations that require solutions of the time harmonic problems, such as: AC electric machines, electromagnetic shields, transformers, magnetic brakes. In the quasi stationary regime, the density of the displacement current is neglected in relation to density of the conduction current $\overline{J_D} = \frac{\partial \overline{D}}{\partial t} = \overline{J} = \sigma \overline{E}$.

The equation of the quasi stationary electromagnetic field for the magnetic vector potentials [35, 38, 41] is the following:

$$\nabla \cdot \frac{1}{\mu} \nabla \overline{A} = \sigma \frac{\partial \overline{A}}{\partial t} + \overline{J_S} \tag{66}$$

where $\overline{J_S}$ is the excitation current and for the linear media:

$$\Delta \overline{A} = \mu \sigma \frac{\partial \overline{A}}{\partial t} + \mu \overline{J_S} \tag{67}$$

For 2D problems the MVP A has only one component A_Z that satisfies the Coulomb gauge and the Eq. (66) becomes:

$$\nabla \cdot \frac{1}{\mu} \nabla \overline{A_z} = \sigma \frac{\partial \overline{A_z}}{\partial t} + \overline{J_{sz}} \tag{68}$$

For the time harmonic steady state regime the complex representation of the MVP A can be obtained using the exponential form:

$$\underline{A} = Ae^{j\omega t} \tag{69}$$

where ω is the angular frequency and j is the imaginary unit.

If we substitute Eq. (69) in Eq. (67), we obtain [40, 34, 36–38]:

$$\nabla \cdot \frac{1}{\mu}\nabla \overline{A} - j\sigma\omega A = \overline{J_S} \tag{70}$$

The solution of Eq. 70 is defined into a finite domain D, subjected to boundary conditions on the frontier Γ. The most encountered boundary conditions require either the normal or the tangential magnetic flux density to be zero respectively.

In terms of MVPA they are equivalent with the following homogenous boundary conditions:

$$\frac{\partial A}{\partial n} = 0 \tag{71}$$

$$\mathrm{A} = 0 \tag{72}$$

We'll apply the weighted residual methodto Eq. 70. In order to satisfy the compatibility condition at the interface at two adjacent finite elements we'll consider shape functions with C_0 continuity. After applying the first Green identity the following relation is obtained [40, 36, 38]:

$$-\int_\Omega \nabla w_i \frac{1}{\mu}\nabla \underline{A} d\Omega + \int_\Gamma w_i \frac{\partial \underline{A}}{\partial n} d\Gamma - \int_\Omega w_i (j\omega\sigma A + J_S) d\Omega = 0 \tag{73}$$

Using the basis functions, the magnetic vector potential can be expressed at elemental level as follows:

$$\underline{A} = \sum N_i \underline{A_i} \tag{74}$$

where for triangular linear elements the shape functionsN_i are:

$$N_i = \frac{a_i + b_i x + c_i y}{2A} \tag{75}$$

where A is the area of an element with the nodes i, j and k.

If we set the shape functions to be the weighted functions$w_i = N_i$, then the weighted residual method becomes the Galerkin method. The discretized form of Eq. 73, at elemental level, becomes:

$$\left[\int_{elem}\frac{1}{\mu}\left(\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x}+\frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y}\right)dxdy-j\int_{elem}\sigma\omega N_iN_jdxdy\right]\cdot[A_i]=\left[\int_{elem}N_iJ_Sdxdy\right] \tag{76}$$

The latter equation can be written in the form of an algebraic system of equations:

$$\left[\underline{k}\right]\cdot\left[\underline{A}\right]=\left[\underline{Q}\right] \tag{77}$$

For the linear triangular finite element elements, the coefficients matrix k_{ij} can be represented as [40]:

$$\underline{k_{ij}}=\underline{p_{ij}}+j\underline{q_{ij}} \tag{78}$$

The real part p_{ij} is described by:

$$p_{\mathrm{ij}}=\frac{1}{4A\mu}\left(\mathrm{b_ib_j}+\mathrm{c_ic_j}\right) \tag{79}$$

and using the Holland- Bell formula [35–37]:

$$\int N_1^aN_2^bN_3^cd_\Omega=2A\frac{a!b!c!}{(a+b+c+2)!} \tag{80}$$

we obtain for q_{ij} the following expression:

$$q_{\mathrm{ij}}=-\sigma\mu\int_{elem}N_iN_jdxdy=\frac{A\sigma\omega}{12}\begin{bmatrix}2&1&1\\1&2&2\\1&1&2\end{bmatrix} \tag{81}$$

where A is the area of the triangular element.

In order to obtain the global matrix, the elemental matrices must be assembled.

The matrix from Eq. 77 is symmetric and the coefficients are complex. The extra time needed to solve an eddy current problem by comparison with magnetostatics case is due to the required time to perform the computations using complex numbers.

The Galerkin method is one of the most versatile methods used for the numerical solution of the PDEs.

4 Case Study: Determination of the Magnetic Field Produced by a High Voltage Electrical Overhead Line Near a Shielded Building Using Finite Element Method

The influence of overhead transmission lines represents a thematic of real interest in the domain of transport of power because it is possible to have an impact on the human body. This application proposes a calculation model of the magnetic field in harmonic regime produced by the high voltage conductors of the transmission systems of electricity using the finite element method. The numerical computation of magnetic field in the vicinity of a high voltage 220 kV electrical overhead line is analyzed. To calculate the magnetic field strength and the magnetic flux density near the overhead power line and the shielded building, the ANSYS Multiphysics software package is used.

The magnetic field strength at ground level depends on the distance to the line and the currents intensity, which flows in the phase conductors. Unlike voltage, the intensity of the electric current can vary very quickly during a day and depending on the season. The intensity of this field depends also on the height and spatial distribution of the conductors. There are other sources (on a small scale) that contribute to the intensity of the magnetic field: currents from the neutral conductor, currents induced in the protective conductors of overhead line, currents in the adjacent ground, in the telecommunications circuits, in pipes or other metallic structures, which it was parallel to the overhead line.

In the literature there are some studies containing an analytical calculation of the magnetic field caused by high voltage power lines. Different techniques can be applied for the calculation of the magnetic field produced by the electrical lines supported by poles of unequal height, unequal distances between poles and or poles whose arrangement is not linear [42]. In another paper, the authors set out to develop a method of calculating the magnetic field generated by conductors arranged in different geometries for single and double circuit electrical lines. For example, in [43], a method of calculating the magnetic field in the vicinity of an electric line called complex double numbers is developed for the following configurations: flat and vertical power lines, lines with polygonal symmetry, delta power line, lines in hexagonal arrangement. The attenuation of the magnetic field created by a double circuit electrical line can also be achieved through an optimization of the arrangement of the phase conductors that feed a railway station and a distribution station [44].

The numerical integration of the Biot-Savart law in differentiated form represents the main analytical method for calculating the density of the magnetic flux [45]. The magnetic field produced by electric power lines is usually calculated numerically with the use of a computer. In [46] specific calculations of the magnetic field for an 110 kV overhead line are presented. In [47] the effect of harmonic components at different electromagnetic frequencies is also taken into account. In [48] historical load databases are used to take into account the relations between magnetic field and electrical load patterns. In [49], 50 the magnetic field distribution is calculated and measured in high voltage substations. To reduce the low-frequency magnetic field in

a building near a high-voltage power line the building is screened with materials that modify its distribution. The shielding factor depends on the permeability, geometry, and thickness of the material [51, 52].

In this application, the magnetic field is calculated using finite elements method (FEM) generated by the conductors of a high voltage electrical line in an area near a building, located in the electric station. The six conductors of a 220 kV transposed high voltage transmission line with double circuit are considered. Consider a system consisting of 6 conductors (double circuit) with 3 conductors on each circuit. The conductors are located at a height of 15 m from earth. Each conductor has a diameter of 30 mm and is located at a distance of 3.6 m and 5 m (central conductors) of the pole of power line. The building is made of concrete with magnetic permeability 1, with a length of 10 m and a height of 6 m, located at a distance of 15 m from the central axis of the conductors. At the top of the building there is a steel screen. Two types of steel with different magnetic permeability of 1000, 4000 and 10,000 respectively were considered. The screen has a length of 11 m and a thickness of 5 mm. Figure 7 shows the physical model consisting of the six conductors and the shielded building and in Fig. 8 is presented a detail that includes the shielded building. The building is surrounded on the outside with a screen with a thickness of 3 mm and in addition to the floor of the building there is a screen of 3 mm. Between the two screens in the floor there is a 4 mm air gap.

The power lines are the conductors of Aluminum Conductor Steel Reinforced (ACSR) type, having the magnetic permeability $\mu_r = 300$ [53, 54]. The conductors are crossed by currents of 375 A, phase shifted by 120° on the three phases. The simulation was done in harmonic mode.

After creating the physical model and define materials, the next step in the preprocessor phase is mesh generation and load applying upon the elements. The finite element mesh of the system with six conductors is shown in Fig. 9. We used a mesh with 314,548 nodes and 628,593 triangular elements. In Fig. 10 is presented the discretization mesh for the shielded building. The boundary conditions and loads to

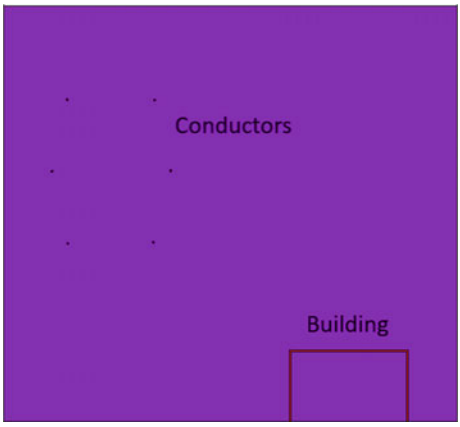


Fig. 7 The physical model subjected to modeling

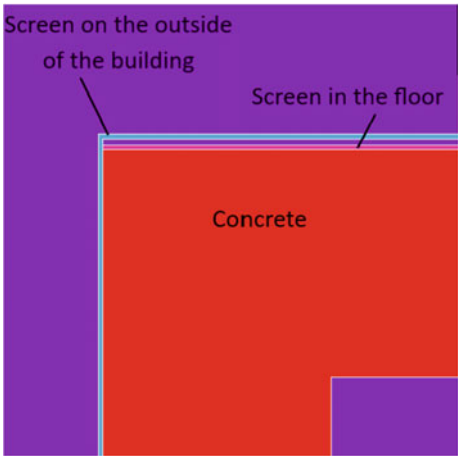


Fig. 8 The physical model comprising the shielded building

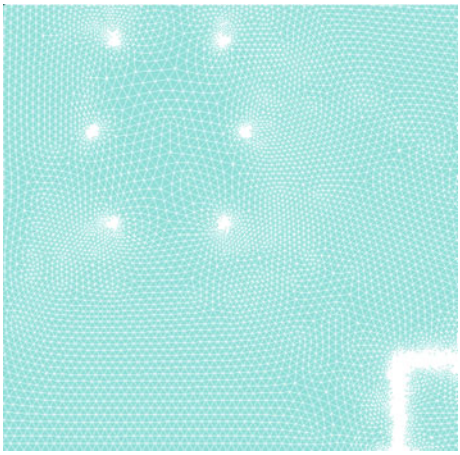


Fig. 9 Finite element mesh

a 2D harmonic magnetic field analysis are applied, both on the plane model (key points, lines, and areas) and on the finite element model (nodes and elements) [55]. The solution of magnetic field problems is commonly obtained using potential functions. Depending on the problem to be solved, one of the two types of potential functions, the MVPor the MSP, is used.

Figure 11 shows the distribution of the magnetic field around the six conductors and shielded building. The maximum value of the magnetic field intensity is obtained around the conductors. In this region, the magnetic field intensity has a maximum of 3403.14 A/m. Inside the building, the intensity of the magnetic field has values between 8.35 and 3.25 A/m for the steel shielded building with magnetic permeability 1000 and between 7.25 and 2.7 A/m for the steel shielded building with the permeability 4000. These values are obtained at a distance of 1.8 m from the ground.

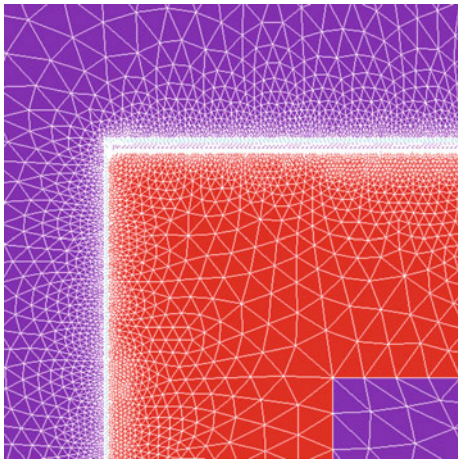


Fig. 10 Finite element mesh (detail building shielded)

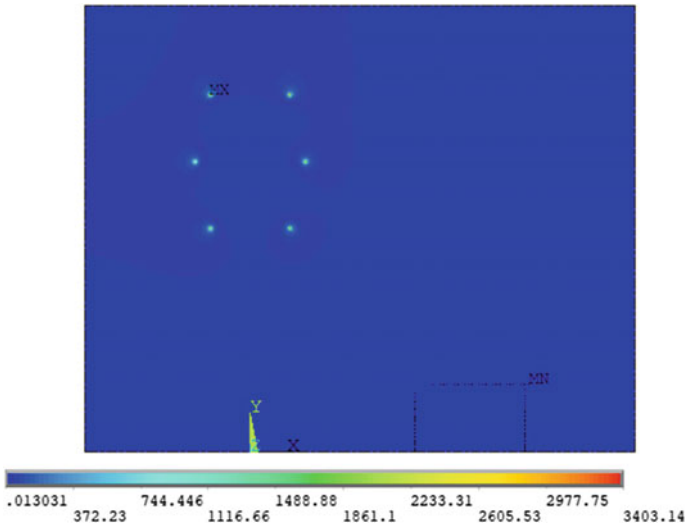


Fig. 11 Distribution of the magnetic field around the conductors and shielded building

Ata distance of 4.8 m from the ground, the maximum values of the magnetic field intensity are 9.37 A/m for the shielded building, with 1000 magnetic permeability steel and 9.06 A/m for the shielded building with 4000 magnetic permeability steel.

The distribution of magnetic flux density all over the surface is shown in Fig. 12 and the distribution of magnetic flux density around the conductor is presented in Fig. 13. The two reference distances were chosen for the evaluation of the magnetic

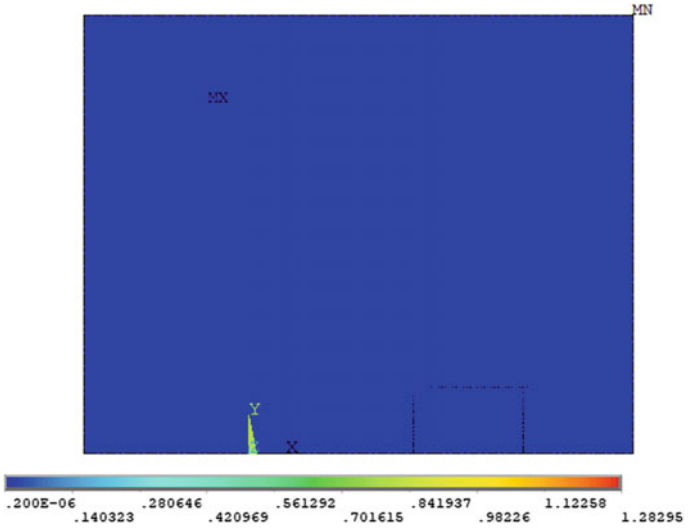


Fig. 12 The distribution of magnetic flux density around the conductors and shielded building

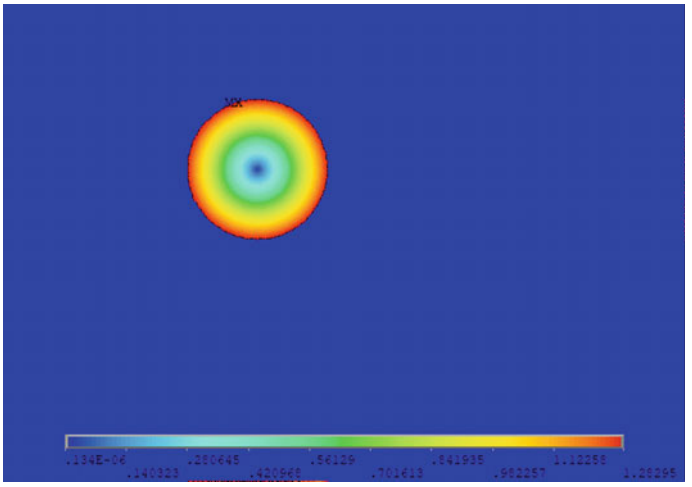


Fig. 13 The distribution of magnetic flux density around the conductor

sizes of 1.8 m and 4.8 m respectively, which means the distance at the head level for the persons on the ground floor and respectively on the floor.

The magnetic induction has values between 9.23 and 2.19 μT for the screen with the magnetic permeability 1000 and 3.64 and 1.09 μT for the screen with the magnetic permeability 4000. These values were obtained on the contour drawn at a distance of 1.8 from the ground.

At a distance of 4.8 m were recorded values between 9.62 and 4.16 μT in the case of the screen with magnetic permeability 1000 and 4.74 and 2.12 μT in the case of the screen with magnetic permeability 4000.

The maximum value of the magnetic flux density is obtained around the conductors. In this region, the magnetic flux density has a maximum of 1.282 T for the conductors traversed by a current of 375 A.

In ANSYS there is a graphical program that displays the resulting fields in the form of contour and density plots. The path for the displayed charts is chosen on a contour consisting of two points placed symmetrical from the shielding building.

Figure 14 and Fig. 15 show the chart of the magnetic flux density inside the building, at 1.8 m and 4.8 m above the ground, respectively. Screening is more efficient when using a material with higher magnetic permeability. In Figs. 16 and 17 represented the values of the magnetic induction in the case of shielding with ferromagnetic steel with magnetic permeability 10,000. It was considered the case

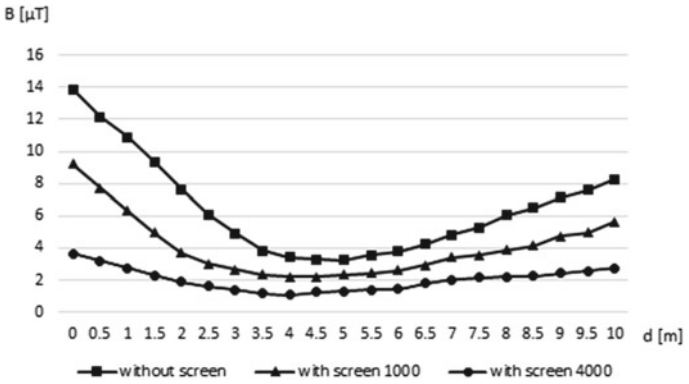


Fig. 14 Chart of magnetic flux density inside the building, at a distance of 1.8 m

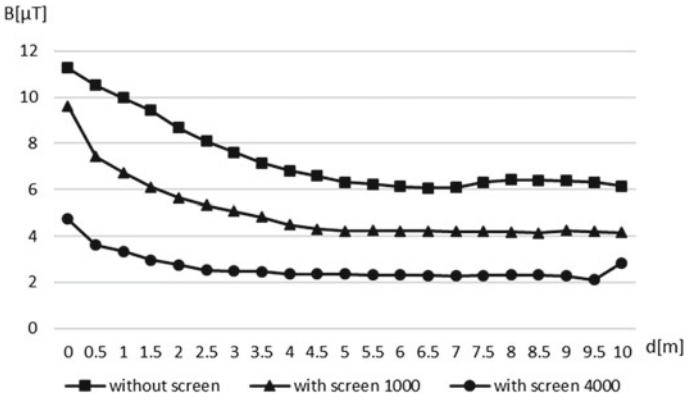


Fig. 15 Chart of magnetic flux density inside the building, at a distance of 4.8 m

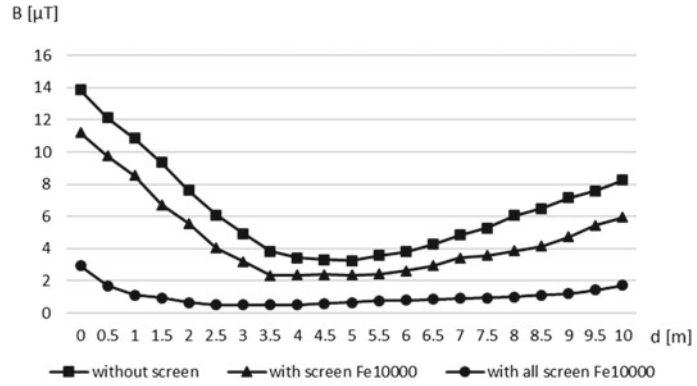


Fig. 16 Chart of magnetic flux density inside the building, at a distance of 1.8 m

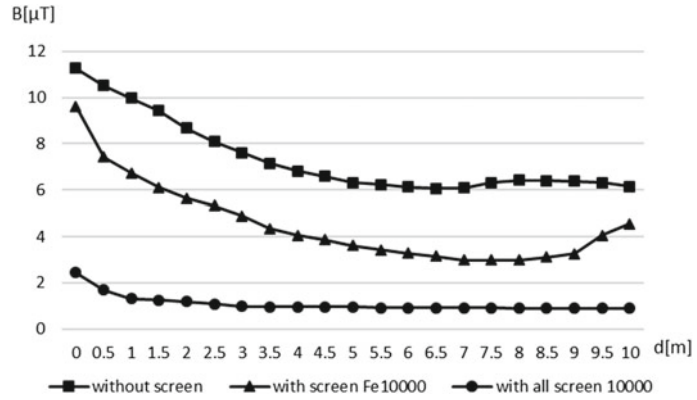


Fig. 17 Chart of magnetic flux density inside the building, at a distance of 4.8 m

of the screen only in the floor of the building (with screen Fe 10,000) but also the complete screen outside the building and in the floor (with all screen Fe 10,000), for those two distances. It is found that the screen only of the floor is not efficient, the values of the magnetic field in this case being comparable with those existing in the absence of the screen. The obtained values are within the limits imposed by ICNIRP standards [56].

5 Conclusions

In this paper, we presented the calculation of the magnetic field produced by a high voltage electrical overhead line in a nearby building located in the electrical station. Screening is efficient when using multi-layer screens made of materials with

the highest magnetic permeability. In the case of complete shielding with 10,000 ferromagnetic steel with magnetic permeability, magnetic induction values below 2 μT were obtained.

Different materials and multi-layer passive shields will be considered in the future. Also, active shields will be considered.

The presented method can be useful in the design of a transmission power lines for insulation distance estimation. Finite element numerical simulation approach can be used to predict the magnetic field generated by high voltage overhead power lines. It can also be used for evaluation of the shielding techniques used for SLF magnetic flux mitigation.

References

1. Bravo JC, del Pino LJ, Cruz P (2019) A survey on optimization techniques applied to magnetic field mitigation in power systems. Energies 12:1332. https://doi.org/10.3390/en12071332
2. Celozzi S, Araneo R, Lovat G (2008) Electromagnetic shielding. Wiley, Inc. https://doi.org/10.1002/9780470268483
3. Fireteanu V, Popa M, Tudorache T (2004) Modele Numerice in Studiulsi Conceptia Dispozitivelor Electrotehnice (publication in Romanian). MatrixRom, Bucharest Romania
4. Salinas E (2001). Mitigation of power-frequency magnetic fields: with applications to substations and other parts of the electric network. PhD thesis, Chalmers University of Technology Göteborg, Sweden, https://publications.lib.chalmers.se/records/fulltext/523/523.pdf. Accessed 10 Jan 2019
5. Zhao Y, Sun Z, Pan D, Lin S, Jin Y, Li L (2019) A new approach to calculate the shielding factor of magnetic shields comprising nonlinear ferromagnetic materials under arbitrary disturbances. Energies 12:2048. https://doi.org/10.3390/en12112048
6. Yang J, Zhang W, Zou L, Wang Y, Sun Y, Feng Y (2019) Research on distribution and shielding of spatial magnetic field of a DC air core smoothing reactor. Energies 12:937. https://doi.org/10.3390/en12050937
7. Cazacu D, Enescu F, Castravete S (2013) Numerical problems in 3D FEM simulation of harmonic magnetic fields. In: 2013 international conference on electronics, computers and artificial intelligence, ECAI 2013, pp 1–4. https://doi.org/10.1109/ECAI.2013.6636193
8. Cazacu D (2013) Numerical problems in 3D magnetostatic FEM analysis. In: 15th WSEAS international conference on automatic control, modelling and simulation ACMOS '13. Brasov, Romania, pp 385–390
9. Cazacu D, Castravete S (2012) Efficiency evaluation via finite element method of 3D magnetostatic shields using different types of magnetic potentials, pp 25–30. https://doi.org/10.1109/SIITME.2012.6384340
10. Cazacu D, Stoica C (2015) Teaching computational aspects on the efficiency of multi layer spherical magneto static shields. In: Proceedings of the 2014 6th international conference on electronics, computers and artificial intelligence, ECAI 2014, pp 59–64. https://doi.org/10.1109/ECAI.2014.7090199
11. Cazacu D (2012) Factors that influence the magnetostatic shields effectiveness. Bull Electr Eng Faculty, Valahia Univ Tirgoviste 1:6–11
12. Cazacu D, Ionescu V, Ionita S, Virjoghe E (2016) On the multishell cylindrical magnetostatic shields: analytical and numerical approach, pp 1–4. https://doi.org/10.1109/ECAI.2016.7861092
13. Cazacu D, Stanescu C (2008) Finite element models of a transmission line for a plane electromagnetic wave in a shielding material. In: The 1st WSEAS international conference on finite differences, finite elements, finite volumes, boundary elements. Malta, pp 117–121

14. Stanescu C, Cazacu D (2008). Experimental and numerical aspects concerning the electromagnetic shielding materials. In: Proceedings of the 4th international conference on technical and physical problems of power engineering. Pitesti, Romania, pp 150–156
15. Virjoghe EO, Băncuţă I, Husu AG, Cazacu D, Florescu V (2019) Measurement and numerical modelling of electric field in open type air substation. J Sci Arts 19(1):249–259
16. Ionescu V, Cazacu D (2016) Analysis of COMSOL multiphysics parallel performance on a multi-core system, pp 1–5. https://doi.org/10.1109/ICATE.2016.7754605
17. Ionescu VM, Cazacu D (2015) Virtualization impact on comsol processor detection. Univ Pitesti Scienti Bull, Series Electron Comput Sci 15(2):25–32
18. Cazacu D, Jubleanu R, Bizon N, Monea C (2019) Comparative numerical analysis of the stored magnetic energy in cylindrical and toroidal superconducting magnetic coils. In: Proceedings of the 2019 11th international conference on electronics, computers and artificial intelligence (ECAI). Pitesti, Romania. https://doi.org/10.1109/ECAI46879.2019.9042162
19. Êvo M, de Paula H, Lopes I, Mesquita R, Souza D (2017) Study of the influence of underground power line shielding techniques on its power capability. J Control Automat Electr Syst 28:1–11. https://doi.org/10.1007/s40313-017-0319-x
20. Ates K, Carlak F, Ozen S (2016) Magnetic field exposures due to underground power cables: a simulation study. https://doi.org/10.11159/eee16.133
21. Li J, Quan W, Han B, Liu F, Xing L, Liu G (2018) Multilayer cylindrical magnetic shield for SERF atomic co-magnetometer application. IEEE Sens J 1–1. https://doi.org/10.1109/JSEN.2018.2890083
22. Sankaran K (2019) Recent trends in computational electromagnetics for defence applications. Defence Sci J 69:65–73. https://doi.org/10.14429/dsj.69.13275
23. Mu L (2019) Weak Galerkin based a posteriori error estimates for second order elliptic interface problems on polygonal meshes. J Comput Appl Mathemat 361. https://doi.org/10.1016/j.cam.2019.04.026
24. Wang C, Wang J (2019) Primal-dual weak Galerkin finite element methods for elliptic Cauchy problems. Comput Mathemat Appl 79(3):746–763. https://doi.org/10.1016/j.camwa.2019.07.031
25. Hortopan G, Vlase O, Nitu S (1990) Ecranarea Electromagntica in Tethnica Curentilor Intensi (publication in Romanian). Ed Tehnica, Bucharest
26. Kou Y, Jin K, Zheng X (2011) Numerical simulation on the shielding efficiency of magnetic shielding enclosures in the ITER applications. Comput Mater Continua 22:129–146
27. Yashchuk V, Lee S, Paperno E (2013) Magnetic shielding. In: Budker D, Kimball DJ (eds) Optical magnetometry. Cambridge University Press, Cambridge, pp 225–248. https://doi.org/10.1017/CBO9780511846380.013
28. Whelan B, Kolling S, Oborn BM, Keall P (2018) Passive magnetic shielding in MRI-Linac systems. Phys Med Biol 63(7):075008
29. Hoburg JF (1995) Principles of quasistatic magnetic shielding with cylindrical and spherical shields. IEEE Trans Electromag Compat 37(4):574–579. https://doi.org/10.1109/15.477342
30. Hortopan G (2005) Principii si Tehnici de Compatibilitate Electromagnetica (publication in Romanian). Editura Tehnica, Bucharest
31. Losito O, Dimiccoli V, Barletta D (2011) Low frequency shielding improvement by multilayer design. In: Proceedings of EMC Europe 2011 York—10th international symposium on electromagnetic compatibility, pp 640–643
32. Lipworth G, Ensworth J, Seetharam K, Lee J, Schmalenberg P, Nomura T, Reynolds M, Smith D, Urzhumov Y (2015) Quasi-static magnetic field shielding using longitudinal Mu-near-zero metamaterials. Scienti Rep 5:12764. https://doi.org/10.1038/srep12764
33. Sora C (1982) Bazele Electrotehnicii (publication in Romanian). Editura Didacticasi Pedagogica, Bucharest
34. Kaiser KL (2006) Electromagnetic shielding. Taylor & Francis
35. Mocanu CI (1981) Teoria Campului Electromagnetic (publication in Romanian). Editura Didacticasi Pedagogica, Bucharest
36. Polyak D (2007) Advanced modelling in computational electromagnetic compatibility. Wiley

37. Mindru G, Radulescu M (1986) Analiza Numerica a Campului Electromagnetic (publication in Romanian). Dacia Publishing House, Bucharest
38. Jin JM (2010) Theory and computation of electromagnetic fields. Wiley
39. Humphries S Jr (1997) Finite element methods for electromagnetics. CRC Press
40. Binns KJ, Lawrenson PJ, Trowbridge CV (1992) The analytical and numerical solution of electric and magnetic fields. Wiley
41. Kuczmann M (2009) Potential formulations in magnetic applying the finite element method. https://maxwell.sze.hu/docs/C4.pdf. Accessed 22 Feb 2019
42. El Dein A (2009) Magnetic-field calculation under EHV transmission lines for more realistic cases. IEEE Trans Power Delivery 24:2214–2222. https://doi.org/10.1109/TPWRD.2009.2028794
43. Filippopoulos G, Tsanakas D (2005) Analytical calculation of the magnetic field produced by electric power lines. IEEE Trans Power Delivery 20:1474–1482. https://doi.org/10.1109/TPWRD.2004.839184
44. Molnar-Matei F, Andea P, Pană A, Teslovan R (2012) Double circuit 110 kV overhead line magnetic field analysis, pp 780–783. https://doi.org/10.1109/MELCON.2012.6196546
45. Andreuccetti D, Zoppetti N, Conti R, Fanelli N, Giorgi A, Rendina R (2003) Magnetic fields from overhead power lines: advanced prediction techniques for environmental impact assessment and support to design, vol 2, 7 pp. https://doi.org/10.1109/PTC.2003.1304684
46. Lunca E, Istrate M, Salceanu A, Tibuliac S (2012) Computation of the magnetic field exposure from 110 kV overhead power lines. In: EPE 2012—proceedings of the 2012 international conference and exposition on electrical and power engineering, pp 628–631. https://doi.org/10.1109/ICEPE.2012.6463803
47. Okrainskaya I, Sidorov AI, Gladyshev SP (2012) Electromagnetic environment under over head power transmission lines 110–500 kV, pp 796–801. https://doi.org/10.1109/SPEEDAM.2012.6264574
48. Mazzanti G (2010) Evaluation of continuous exposure to magnetic field from AC overhead transmission lines via historical load databases: common procedures and innovative heuristic formulas. IEEE Trans Power Delivery 25:238–247. https://doi.org/10.1109/TPWRD.2009.2035390
49. Munteanu C, Pop, LT, Tapa V, Hangea C, Gutiu T, Lup S (2012) Study of the magnetic field distribution inside very high voltage substations, pp 660–663. https://doi.org/10.1109/ICEPE.2012.6463571
50. Aponte G, Cadavid H, Perez R, Escobar A, Mora A, Bolanos H (2003) Electric and magnetic fields measured in Colombian lines and substations, vol 2, pp 1195−1198. https://doi.org/10.1109/ICSMC2.2003.1429132
51. Beiu C, Golovanov N, Cornel T, Buica G (2017) Low frequency electromagnetic shielding solutions. https://doi.org/10.1109/SIELMEN.2017.8123301
52. Livesey K, Camley R, Celinski Z, Maat S (2017) Magnetic shielding of 3-phase current by a composite material at low frequencies. AIP Adv 7:056328. https://doi.org/10.1063/1.4978702
53. Andrei H, Fluerașu C, Vîrjoghe EO, Fluerașu C, Enescu D, Popovici D, Husu AG, Andrei PC, Predușcă G, Diaconu E (2011) Numerical methods, modelling and simulation applied in electrical engineering. Ed. Electra, Bucharest
54. Tupsie S, Isaramongkolrak A, Pao-la-or P (2010) Analysis of electromagnetic field effects using FEM for transmission lines transposition. Int J Electr Comput Eng 5(4):713–717
55. Salari J, Mpalantinos A, Silva JI (2009) Comparative analysis of 2- and 3-D methods for computing electric and magnetic fields generated by overhead transmission lines. IEEE Trans Power Delivery 24:338–344
56. Lin J, Saunders R, Schulmeister K, Söderberg P, Stuck BE, Swerdlow A, Taki M, Veyret B, Ziegelberger G, Repacholi M, Matthes R, Ahlbom A, Jokela K, Roy C (2010) ICNIRP guidelines for limiting exposure to time-varying electric and magnetic fields (1 Hz to 100 kHz) Health Phys 99. https://doi.org/10.1097/HP.0b013e3181f06c86

Regression Analysis-Based Load Modelling for Electric Distribution Networks

Gheorghe Grigoras and Bogdan Constantin Neagu

Abstract The decision making in the electric distribution systems is based on data collected from consumers and the various measurement points located in the network (transformer substations, supply points, branch points, etc.). The information obtained from a 100% integration of the smart metering to consumers comes to fill the data acquired through the Supervisory Control and Data Acquisition (SCADA) system, so that the Distribution Network Operator (DNO) can accurately estimate the state of the supervised system. However, the implementation of smart metering is in various implementation stages in different countries of the world, so that today it can be stated that there is no complete integration. The same aspect should be emphasized in the case of the SCADA system at the level of distribution networks, which is not 100% integrated in the low/medium voltage electric substations. In these conditions, the DNO should apply the mathematical tools that take into account the similarities between the consumers' behaviour and respectively the structure of the load supplied from the electric substations. The regression analysis-based approaches for the load modelling from the nodes of electricity distribution networks were treated in the chapter. The approaches refer to the estimation of the required powers in the supply points with a mixt structure of the load (i.e. residential, commercial, and industrial) at the hour when the maximum value of the load is recorded and the demand of residential consumers which represent the highest percentage from the load structure fed from the electric substations. The proposed approaches were tested in real operation conditions of the distribution networks from Romania.

Keywords Regression · Correlation · Peak load · Distribution networks · Residential consumers

G. Grigoras (✉) · B. C. Neagu
Electrical Engineering Faculty, Power System Department, Gheorghe Asachi Technical University of Iasi, Iasi, Romania
e-mail: ggrigor@tuiasi.ro

B. C. Neagu
e-mail: bogdan.neagu@tuiasi.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_28

Abbreviation and Acronyms

SCADA	Supervisory Control and Data Acquisition
DNO	Distribution Network Operator
LV	Low Voltage
MV	Medium Voltage
HV	High Voltage
LR	Linear Regression
PR	Parabolic Regression
HR	Hyperbolic Regression
ER	Exponential Regression

1 Introduction

The loads from the nodes of electricity distribution systems (represented by the Medium Voltage/Low Voltage (MV/LV) electric substations) vary in time and have particular characteristics in each consumption point. Therefore, to solve the problems regarding the optimal network planning and operation, the demand management and the correct billing of consumers, the Distribution Network Operators (DNOs) need to know the dynamic behaviour of the loads in their networks [1–3]. On the other hand, the load variations are influenced by several factors, such as consumer type, time factor, climatic factors, other electrical loads correlated with the analysed load, historical values, and consumption profile [4–6].

The modelling of electricity consumptions is made using the records from the databases which describe the evolution of individual and aggregated loads. These data are recorded and processed systematically using appropriate methods. The following input information is frequently used in an analysis: the daily maximum value of load, the hourly power consumption, the daily/weekly electricity amount. For a better accuracy in the modelling process, a large database should be used, including the electricity consumptions for a long-time interval and, if possible, the evolution of demographic, climatic and economic activity indices for the geographical area and time interval of interest [5, 6].

Also, there are some restrictions which the Decision Makers must consider them in their analyses [7, 8]:

- The power flows must satisfy the fundamental laws of electrotechnics (Kirchhoff laws);
- The balance between the obtained loads in the estimation process and the measured values.
- The load does not depend by the structure of network.

The randomly selected working sample from the database must be subjected to a detailed analysis to identity the outliers, then following the correlation process to find

the relationships between the variables represented by the power/energy consumption and the climatic and weather factors [9, 10].

In the chapter, various approaches for the load modelling from the nodes of electric distribution networks, based on the correlation and regression analysis will be proposed. The support of the proposed approaches is represented by the processing process of the load profiles belonging to the MV/LV electric substations or LV consumers recorded with the help of smart meters using the statistical tools. The structure of chapter is divided in two parts: a short review about the correlation and regression analysis is made in the first part, and in the second part the regression analysis based-approaches are presented regarding the estimation of the powers in the MV/LV electric substations (at the hour when the maximum value of the load from the system is recorded) and the demands of the residential consumers.

2 Correlation and Regression Analysis

To understand the operation of electric distribution systems, it is necessary to be studied the relationships between the state variables that characterize them (voltages, currents, powers, etc.). For these variables, the relationships can be analysed using the regression and correlation methods.

The regression methods allow the measurement and study of the relation between two or more variables, as well as the discovery of the connection laws between these. A mathematical expression can be obtained with the aim to estimate the values of one independent variable according to the values of other variables [11, 12].

Correlation analysis measures the intensity of the relationship between one or more variables. Depending on the regression model, the correlation can be treated as a single or multiple correlation [13, 14].

The following issues must be solved in a study which is based on the regression and correlation analysis [12]:

- Identify the existence of the relationship between variables. Solution: A logical analysis of the possibility of a relationship between the variables can be applied.
- Establishing the meaning and form of the relationship. Solution: Regression analysis methods can be used.
- Determining the intensity degree of relationship. Correlation analysis methods can be used.

2.1 Correlation Methods

2.1.1 Interdependent Parallel Statistical Series-Based Method

The analysis of statistical relationships takes into account the estimation of a regression model and measuring the intensity of the relationship between variables. The

Table 1 Cross-correlation matrix

x	y	
	$y_1 \ldots y_j \ldots y_p$	$\sum_j n_{ij} = n_i*$
x_1 … x_i … x_k	$n_{11} \ldots n_{1j} \ldots n_{1p}$ … $n_{i1} \ldots n_{ij} \ldots n_{ip}$ … $n_{k1} \ldots n_{kj} \ldots n_{kp}$	n_{1*} … n_{i*} … n_{i*}
$\sum_i n_{ij} = n_{*j}$	$n_{*1} \ldots . n_{*j} \ldots . n_{*p}$	$n** = \sum_i \sum_j n_{ij}$

analysis of the statistical relationship compares the terms of two interdependent parallel series x (independent variable) and y (dependent variable). For example, when two time series are compared, their elements are chronologically sorted, such that the existence and direction of the relationship can be easy identified. Thus, if both variables have a variation in the same direction, there is a direct relationship. If the variation is different, an inverse correlation is obtained. If the two time series vary independently, or one varies and the other remains constant, there is no relationship [8].

The method can be used for the time series with few variables, when there is a relationship between the pairs of variables (x_i, y_i, $i = 1, \ldots, N$).

2.1.2 Cross-Correlation Matrix Based Method

The principle of method is based on the grouping the elements of a data set using simultaneously both correlated variables (x and y). Equal intervals and an identical number of groups for both variables are recommended to be used. Thus, in the matrix, the existence, direction and intensity of the relationship can be appreciated using the distribution model of frequencies $n_{ij,}$ as it can be seen in Table 1.

If the frequencies n_{ij} are scattered relatively uniformly inside the matrix, there is no relationship between the variables considered. But, if they are concentrated around the diagonals, a stronger correlation can be identified between the variables x and y.

2.1.3 Graphical Method

The method involves the graphical representation of the pairs of values corresponding to the variables in a coordinate system, such that the existence, meaning, form and intensity of the correlation can be easy identified. The graph corresponds to the case where a relationship is defined in concordance with interdependent statistical parallel series-based method.

2.1.4 Analytical Methods

The analytical models allow determination of the mathematical relations and the numerical measurement of the intensity between variables. The regression models aim to represent the distribution type of correlated variables. The regression curves indicate the correspondence between the pairs (x_i, y_i). The following steps should be performed to establish and analyse a regression model:

- Building the correlation graph.
- Establishing the theoretical regression model of the relationship (based on the correlation graph adjustment) and identification of the equation corresponding to the chosen regression model.
- Determining the coefficients of the regression equation (with the least squares method) and interpreting the regression according to their sign and value.

2.1.5 Regression Models with Two Variables

The relationship between two variables x and y can be expressed by a regression equation:

$$y_x = f(x) + e \tag{1}$$

where $f\ (x)$ represents a function which is dependent on the variable x, and e is the approximation error.

If the size of the database will grow, the approximation error e will decrease. Thus, a higher number of observations can lead at a stronger relationship. Function $f(x)$ can have different models depending by the data scatter.

Linear regression (LR) model

The LR model is most used in the practice. The relationship can be expressed using the following equation:

$$y_x = a + bx + e \tag{2}$$

The Eq. (2) can be plotted using a line. The variable e represents a random error given by:

$$e = y_i - y_{x_i}\ ; \quad i = 1, \ldots, N \tag{3}$$

where a and b are unknown coefficients, their values being determined using the least squares method.

The coefficient b from the expression (2) can have different signs which characterize the direction of the relationship between variables: "+", positive relationship; "null", no relationship, and "−", negative relationship.

The value of coefficient b shows the dependence degree between variables, namely how much the variable y increases or decreases when the variable x increases or decreases with one unit.

Parabolic regression (PR) model

In order to express this model, the second degree polynomial is usually used:

$$y_x = a + bx + cx^2 + e \tag{4}$$

where coefficients a, b and c are determined using the least squares method.

Hyperbolic regression (HR) model

$$y_x = a + \frac{b}{x} + e \tag{5}$$

Exponential regression (ER) model

In order to express this model, the following equation is used:

$$y_x = ab^x + e \tag{6}$$

For each sample, rel. (6) can be linearized by logarithm:

$$\log y_x = \log a + x \log b \tag{7}$$

2.2 Intensity of the Relationship Between Two Variables

The intensity of the relationship, if there is between two variables (x, y), indicates a concentration degree of or scattering of the values y around the regression model y_x. The intensity of the relationship can be measured based on the correlation coefficient and the correlation ratio.

2.2.1 Correlation Coefficient

The correlation coefficient is used to appreciate the intensity of relationship between the analysed variables. The calculation of this coefficient can be made using the relation:

$$\rho(x, y) = \frac{C(x, y)}{\sigma_x \cdot \sigma_y} = \frac{\sum_i (x_i - x_m)(y_i - y_m)}{n \cdot \sigma_x \cdot \sigma_y}, \quad i = 1, \ldots, N \tag{8}$$

where: $C(x, y)$—the covariance between analysed variables; x_m, y_m—the mean values of the variable; N—number of pairs of values; σ_x and σ_y—the standard deviation of variables x and y.

Between the regression coefficient b from relation (2) and the correlation coefficient, $\rho\ (x, y)$, there is the following relationship:

$$\rho = b \cdot \frac{\sigma_x}{\sigma_y} \tag{9}$$

The analysis of the relation (9) highlights that the sign of the correlation coefficient is identically with the sign of the regression coefficient, because σ_x and σ_y are positive or equal with zero. The value of the correlation coefficient is in the range $[-1, 1]$. These two extreme values represent a perfect linear relationship between the two variables ("positive" or "negative"). The missing of a relationship between the two variables can be recorded if $\rho = 0$.

2.2.2 Correlation Ratio

The correlation ratio η is defined by the relation:

$$\eta = \sqrt{\frac{\sigma_{y_x}^2}{\sigma_y^2}} \tag{10}$$

where

$$\sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n};\ \sigma_{y_x}^2 = \frac{\sum \left(y_{x_i} - \bar{y}\right)^2}{n} \tag{11}$$

The correlation ratio have the values into the range [0, 1]. The value 1 indicates the existence of a relationship, namely the variation of the variable y depends only on by the variation of variable x.

3 Case Studies in the Electric Distribution Networks

3.1 Power Correlation Problem

The quality and efficiency of complex problem-solving process regarding the optimal operation and planning of the electric distribution networks are largely determined by the accuracy of the load estimation methods. The estimation of the power demand and the electricity consumption is made starting from the historical data on the evolution of consumption, which is recorded systematically, processed by appropriate

methods. The main factors which can be taken into account are: daily peak load, hourly electricity consumption, and daily or weekly electricity [15, 16]. In order to have the most accurate estimation, a large-size database should be used including the hourly electricity consumptions for a sufficiently long period (minimum 1 year), the evolution of demographic and climatic factors, and economic indexes in the analysed areas [4–6]. These information must be subjected to a pre-processing stage to eliminate systematic, gross, and random errors, and then if it possible to find a relationship between variables represented by the electricity consumption and the climatic and weather factors [7, 8, 14, 17].

The practice applications have concluded that the success of an estimation method is based on the achievement of some appropriate conditions, such as: an accurate selection of estimation period, the applied method, the confidence of the initial data, the flexibility, and taking into account the climatic and weather factors. In the load estimation process (including the peak load), there are more mathematical methods developed in the literature. The most of the proposed approaches use the dependence between the maximum value of the load (peak load) and the annually/monthly/daily electricity consumption [7, 8].

Today, the most Distribution Network Operators (DNOs) from the European countries are in full process of implementing the smart metering system in the MV/LV electric substations and at the end consumers. The problem is that this process is slow and there are enough electric substations for which DNOs do not have yet information on their loading and the peak load to estimate the operation regime of the electric network. In this case, the loads, generally, and the peak load, particularly, can be estimated based on correlation studies, as will be shown in the following [6–8, 18].

If a simple linear regression model is used for the relationship between the mean values of the variables P and Q, then the following relations can be accepted (see Fig. 1):

$$Q = \rho_{PQ} \cdot \frac{\sigma_Q}{\sigma_P} \cdot P + k_{PQ} \tag{12}$$

$$\rho_{PQ} = \frac{C_{PQ}}{\sigma_P \cdot \sigma_Q} \tag{13}$$

$$C_{PQ} = \overline{P \cdot Q} - \overline{P} \cdot \overline{Q} \tag{14}$$

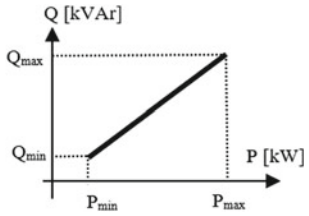


Fig. 1 The correlation between P and Q (direct variation)

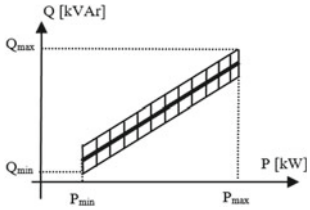


Fig. 2 The correlation between P and Q (opposite variation)

$$\sigma_P^2 = \overline{P^2} - \overline{P}^2 \tag{15}$$

$$\sigma_Q^2 = \overline{Q^2} - \overline{Q}^2 \tag{16}$$

where: P—the active power [kW], Q—the reactive power [kVAr], ρ_{QP}—the correlation coefficient between P and Q; C_{PQ}—covariance between P and Q; σ_P, σ_Q—the standard deviation of P and Q.

The overline indicates the mean value, and the coefficient k_{QP} is determined for each particular case, based on the correlation studies [7].

But, there are cases where the powers P and Q have an opposite variation. In these cases, a "variation belt" should introduced (see Fig. 2).

3.2 Peak Load Estimation Using Power Correlation

3.2.1 Solution Description

The estimation of the loads from the MV/LV electric substations at the hour when the maximum value (peak load) in the electric distribution system was recorded, will be made in this paragraph using a power correlation-based method.

In the initial step, a statistical analysis of the load profiles regarding to the active power from a database belonging a DNO in the MV/LV electric substations without the installed smart meters is performed. Different time frames can be used in this analyse, depending on the technical and load characteristics of the network. The length of the time frames (L_h with $h = 7$ or 24) could be chosen from the following: *L_{24} frame*, *L_7 frames* ($h_{PL} \pm 3\ h$), ($h_{PL} - 4\ h$; $h_{PL} + 2\ h$) and ($h_{PL} - 5\ h$; $h_{PL} + 1\ h$), where h_{PL} is the hour when the maximum value of load (peak load) from the system was recorded.

Using the LR model, the steps of the estimation method are the following:

1. Consideration of a main variable in relation to which the correlation analysis will be performed. The main variable can be chosen as the HV/MV electric substation because the hourly powers P and Q are recorded all along using smart meters.
2. Determining the peak load and the hour when is recorded for the reference electric substation.

3. Calculation of the correlation coefficients between the profiles of the powers P and Q, recorded in each MV/LV electric substation, and the profile for the power P, recorded in the HV/MV electric substation chosen as reference. Also, the standard deviation of the powers P and Q recorded in the MV/LV electric substation will be calculated.
4. Determination of the values for the coefficients $b_{P_r}^{P_i}$, $b_{P_r}^{Q_i}$, a^{P_i}, and a^{Q_1} with the relations:

$$b_{P_r}^{P_i} = \rho_{P_r P_i} \cdot \frac{\sigma_{P_i}}{\sigma_{P_r}}; \quad i = 1, \ldots, N \tag{17}$$

$$b_{P_r}^{Q_i} = \rho_{P_r Q_i} \cdot \frac{\sigma_{Q_i}}{\sigma_{P_r}}; \quad i = 1, \ldots, N \tag{18}$$

$$a^{P_i} = \sum_{j=1}^{h} (P_{ij} - b_{P_r}^{P_i} \cdot P_{rj})/L_h; \quad i = 1, \ldots, N \tag{19}$$

$$a^{Q_i} = \sum_{j=1}^{h} (Q_{ij} - b_{P_r}^{Q_i} \cdot P_{rj})/L_h; \quad i = 1, \ldots, N \tag{20}$$

where: P_r—the active power corresponding to the HV/MV electric substation chosen as reference; P_i, Q_i—the active and reactive powers from the MV/LV electric substation i; N—the number of MV/LV electric substations from the analysed network; L_h—the length of time frame ($h = 7$ or 24).

5. Estimation of the powers P and Q from the MV/LV electric substations at the hour when the maximum value of load in the system was recorded can be made using the following LR models:

$$P_i = b_{P_r}^{P_i} \cdot P_{r\max} + a^{P_i} \tag{21}$$

$$Q_i = b_{P_r}^{Q_i} \cdot P_{r\max} + a^{Q_i} \tag{22}$$

where: $P_{r\,max}$—the peak load corresponding to the reference; P_i, Q_i—the estimated powers in the MV/LV electric substation $i = 1, \ldots, N$.

3.2.2 Testing the Solution

This paragraph presents testing the proposed method based on database belonging an electric MV distribution system (20 kV) with 34 MV/LV electric substations. The

Table 2 The estimated active powers in the MV/LV electric substations

No.	P_m (kW)	Frame L_{24}		Frame L_7 (h_{PL} − 4 h; h_{PL}+ 2 h)		Frame L_7 (h_{PL} − 3 h; h_{PL}+ 3 h)		Frame L_7 (h_{PL} − 5 h; h_{PL}+ 1 h)	
		P_e (kW)	Er_P (%)	P_e (kW)	Er_P (%)	P_e (kW)	Er_P (%)	P_e (kW)	Er_P (%)
1	240.5	239.87	−0.205	241.24	0.31	242.91	1.00	239.18	−0.54
2	216.3	224.75	3.90	220.51	1.95	221.72	2.50	214.72	−0.73
3	311.1	346.66	11.43	321.91	3.47	325.90	4.75	315.86	1.53
4	436	397.62	−8.80	442.13	1.40	448.77	2.92	430.6	−1.23
5	410.7	381.8	−7.03	414.33	0.88	404.67	−1.46	422.27	2.81
6	420.3	422.38	0.49	411.41	−2.11	415.96	−1.03	407.03	−3.15
7	600.6	614.11	2.24	604.11	0.58	597.28	−0.55	610.54	1.65
8	617.7	609.06	−1.39	611.49	−1.00	605.77	−1.93	622.15	0.72
9	561.2	560.48	−0.12	561.27	0.01	553.07	−1.44	567.94	1.20
10	208.7	213.18	2.15	205.07	−1.73	211.44	1.31	200.76	−3.8
11	617	607.32	−1.56	612.51	−0.72	598.1	−3.06	626.26	1.50
12	588.1	562.06	−4.42	588.08	−0.00	593.15	0.86	586.35	−0.29
13	404.9	408.13	0.79	389.69	−3.75	378.81	−6.44	406.06	0.28
14	357.2	397.83	11.37	360.54	0.93	362.83	1.57	359.61	0.67
15	360.4	384.7	6.74	356.33	−1.12	354.75	−1.56	363.06	0.73

(continued)

peak load in this system is recorded at the hour 15. Following the steps of method, the LR models for different time frames were used in the analysis. The values of the active powers at the hour when the load peak was recorded in the analysed system, for the time frames L_{24} and L_7, are presented in Table 2.

The RL models obtained for all considered time frames in the case of a MV/LV electric substation (no. 28) from the analysed system are represented in Figs. 3, 4, 5 and 6 to observe the estimation accuracy for some time frame.

The errors were calculated with the relation:

$$Er_p = \frac{P_e - P_m}{P_m} 100 \quad [\%] \tag{23}$$

where: P_e—estimated active power; P_m—measured active power.

It can be observed that the errors are smaller in the case of the time frame *7 h* ($h_{PL} - 5\ h$; $h_{PL} + 1\ h$) than in the others frames, the average error being 1.48%.

Table 2 (continued)

No.	P_m (kW)	Frame L_{24}		Frame L_7 (h_{PL} − 4 h; h_{PL}+ 2 h)		Frame L_7 (h_{PL} − 3 h; h_{PL}+ 3 h)		Frame L_7 (h_{PL} − 5 h; h_{PL}+ 1 h)	
		P_e (kW)	Er_P (%)	P_e (kW)	Er_P (%)	P_e (kW)	Er_P (%)	P_e (kW)	Er_P (%)
16	365.6	371.42	1.59	376.21	2.90	388.43	6.24	356.33	−2.53
17	545.7	596.06	9.22	577.8	5.88	596.63	9.33	543.04	−0.48
18	254.9	257.03	0.83	258.62	1.46	257.09	0.86	260.77	2.30
19	191.3	213.61	11.66	196.52	2.73	202.63	5.92	191.78	0.25
20	168.1	196.46	16.87	175.47	4.38	183.60	9.22	162.46	−3.35
21	667.5	622.26	−6.77	633.51	−5.09	618.86	−7.28	662.08	−0.81
22	421.4	409.54	−2.81	423.75	0.55	414.28	−1.68	428.87	1.77
23	440.4	388.9	−11.69	419.91	−4.65	411.62	−6.53	436.36	−0.91
24	637	642.21	0.81	620.27	−2.62	633.43	−0.56	619.42	−2.75
25	452.3	444.02	−1.82	453.28	0.21	459.63	1.62	448.10	−0.92
26	623.7	534.84	−14.24	622.70	−0.16	615.64	−1.29	616.91	−1.08
27	402.2	392.49	−2.41	403.55	0.33	394.5	−1.91	408.35	1.53
28	671.8	677.4	0.83	682.12	1.53	686.71	2.21	679.72	1.18
29	634.1	643.68	1.51	638.69	0.72	640.00	0.93	628.87	−0.82
30	594.2	569.71	−4.12	606.74	2.11	613.92	3.32	594.83	0.10
31	388.1	437.63	12.76	392.07	1.02	386.38	-0.44	399.55	2.95
32	329.8	341.49	3.54	333.66	1.17	332.31	0.76	330.87	0.32
33	571	568.94	−0.36	551.65	−3.38	563.85	−1.25	549.30	−3.80
34	635.4	667.39	5.03	637.9	0.39	630.39	−0.78	655.06	3.09

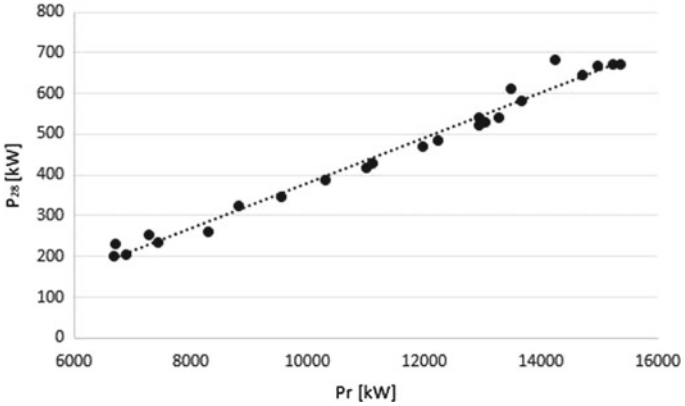


Fig. 3 Linear regression model $P_{28} = 0.0556 \cdot P_r - 175.5$ (Time Frame L_{24})

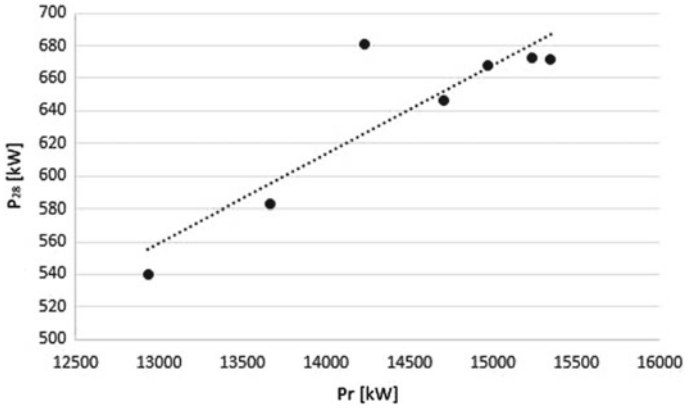


Fig. 4 Linear regression model $P_{28} = 0.0543 \cdot P_r - 146.1$ (Time Frame L_7 ($h_{PL} \pm 3\ h$)

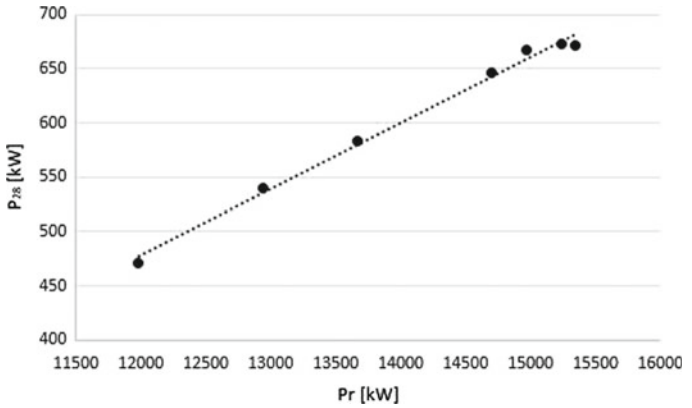


Fig. 5 Linear regression model $P_{28} = 0.0609 \cdot P_r - 251.7$ (Time Frame L_7 ($h_{PL} - 4\ h$; $h_{PL} + 2\ h$))

3.3 Residential Load Estimation Using a Regression—Correlation-Based Method

3.3.1 Solution Description

Load estimation has seen in the latest decades an increase in importance, complexity and need of accuracy. Before 1970, the electricity demand was relatively predictable, and a good forecast required simple mathematical models, limited to trend extrapolation. Also, the "7% rule" was used, which stated the doubling of electricity demand in each 10 years [19]. The load estimation studies are influenced by more factors, which can be grouped as follows [1]:

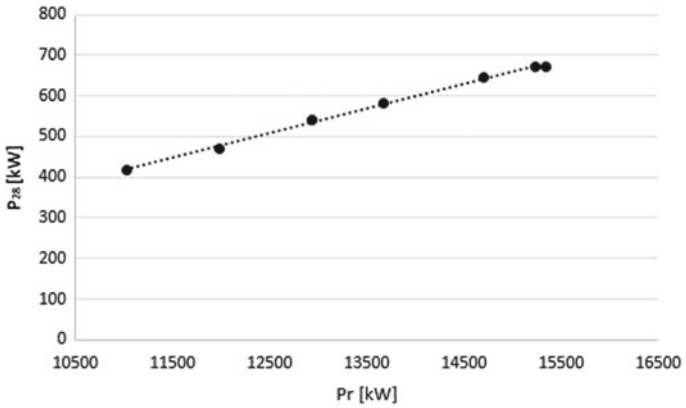


Fig. 6 Linear regression model $P_{28} = 0.0602 \cdot P_r - 244.03$ (Time Frame L_7 ($h_{PL} - 5\ h$; $h_{PL} + 1\ h$))

- Economic: for long and medium time period. These factors aren't responsible for hourly load variations and aren't considered in short term forecasts.
- Temporal—seasons, daily and weekly cycles, holidays, daylight intervals.
- Weather: temperature, humidity, wind speed and direction, clouds, rain.
- Casual: holidays, worker strikes, public events.

Practical studies have shown that the demand variation in time or according to other considered parameters has four main components [17]: season $S(t)$; cyclic $C(t)$; trend $T(t)$, and random $R(t)$. The demand can be written as the sum of the four factors, using the following equation:

$$W(t) = S(t) + C(t) + T(t) + R(t) \tag{24}$$

The mathematical function used in the estimation process is determined by successive steps, taking into account the consumption history and a qualitative and quantitative analysis of the technical and economic factors which influence in time over the consumer demand. In order to obtain a model for the demand of a consumer group or a geographical area requires the testing of several approximation approaches. For electrical load estimation, the optimal approximation functions are obtained using specialized software tools, which choose the best variant among a wide range of options.

The accuracy of the selected estimation model is assessed by computing indices which give the spread of the initial data (earlier demand values) with regard to the considered trend. Usually, a low spread indicates a good approximation which can be expressed by the quality indices (I_k). The mathematical expressions of these indices are given below:

- the mean absolute values of deviations:

$$I_1 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \hat{y}_i - y_i \right| \tag{25}$$

- the mean absolute percentage values of deviations:

$$I_2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \cdot 100 \tag{26}$$

- the mean absolute deviation:

$$I_3 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \hat{y}_i - \bar{y} \right| \tag{27}$$

- the dispersion:

$$I_4 = \sigma^2 = \frac{1}{n - m - 1} \cdot \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2. \tag{28}$$

but the value is different with the total variance of y:

$$\sigma_t^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - \bar{y}_i)^2 \tag{29}$$

- the mean square deviation of the selection

$$I_5 = \sigma \tag{30}$$

- the variation coefficient:

$$I_6 = v = \frac{\sigma}{x} \tag{31}$$

- the correlation coefficient from (8) in a particular form:

$$I_7 = \rho = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\pm\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{32}$$

- the particular form of (10) of correlation ratio will be:

$$I_8 = \eta = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{33}$$

where $\hat{y}_i$—the estimated value, y_i—the real demand, $\bar{y}$—the mean value of the historical consumption, m—the degree of the polynomial used for trend approximation.

In order to compute the trend, as recommended in the literature, continuous functions were used, which can be represented as continuous growth curves and limited growth curves. Their coefficient was determined using time series regression, with normal and modified methods using the sum squared error criterion. This approach is frequently used for residential load estimation.

The load estimation in the MV/LV electric substations is more difficult, because of the lack of historical demand data from consumers. Moreover, load estimation at the level of each DNO is possible with much better accuracy, using load data recorded through the continuous monitoring in the HV/MV electric substations and applying the global estimation methods [17].

Thus, for a year j from the estimation interval P_{m+j}, the load estimation can be obtained based on a mathematical model which uses historical load data:

$$P_{n+j} = \frac{\sum_{k=0}^{m-1} \sum_{i=1}^{n-j-k} \frac{P_{i+j+k}}{P_i}}{\sum_{k=0}^{m-1} (n - j - k)} \tag{34}$$

where: n—the previous years for which recordings exist; m—the previous years used as forecast base; j—forecast year; k—base year.

Previous studies have shown that ambient temperature has a significant influence on demand [4]. The load estimation with the temperature (computed for several consecutive years) can be:

$$P_{pr} = \frac{P_r}{1 + \frac{a}{b}\Delta\theta} \tag{35}$$

where P_r—the real load, measured in a given year; $\Delta\theta$—the difference between the real and average temperature recorded for several years, over a given time interval; a—regression coefficient with the temperature θ; b—the average load ratio for years j and $j-1$.

Accounting for the (load-temperature) correlation, which differs monthly, and sometimes is greater at the night hours than at the day hours, if temperature forecasts are known for the next year, then the load estimation for the next year can be computed:

$$P_{(n+1),\theta} = P_{n,\theta_n}(b + a \cdot \Delta\theta) \tag{36}$$

where $\Delta\theta$—the difference between the next year temperature forecast and the multi-year temperature; P_{n,θ_n}—the load from the last year; a—regression coefficient with the temperature θ; b—the average load ratio for years j and $j-1$.

Using statistical methods [2, 3, 20] the peak load level growth for individual residential consumers can be computed with:

$$S_{\max} = \overline{S}_{\max} + \lambda \cdot \sigma \tag{37}$$

where $\overline{S}_{\max}$—the mean value of the peak load for the residential consumer:

$$\overline{S}_{\max} = \sum_{i=1}^{n} S_{\max_i} \tag{38}$$

σ—mean square deviation, computed as a particular form:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n}\left(S_{\max_i} - \overline{S}_{\max}\right)^2} \tag{39}$$

n—number of residential consumers with the available measurements;

λ—rated deviation of the normal distribution.

For the estimation of the monthly load, the profile of the warm season (December—month 12) and the profile of the cold season (June, month 6) can be used in any month l:

$$P_{t,l} = \frac{P_{t,12} + P_{t,6}}{2} + \frac{P_{t,12} - P_{t,6}}{2} \cos\frac{\pi \cdot l}{2} \tag{40}$$

where: $P_{t,l}$—the active power at hour $t = 1, \ldots, 24$, in month l; $P_{t,6}$, $P_{i,12}$—the active power at hour $t = 1, \ldots, 24$, in month 6 (June) and month 12 (December).

If the yearly load growth is considered, (40) can be rewritten as:

$$P_{t,l} = \frac{\alpha \cdot P_{t,12} + \frac{1+\alpha}{2} \cdot P_{t,6}}{2} + \frac{\alpha \cdot P_{t,12} - \frac{1+\alpha}{2} \cdot P_{t,6}}{2} \cdot \cos \frac{\pi \cdot l}{2} \tag{41}$$

where α is the yearly load growth coefficient.

The estimation model or function is chosen according to the least squares' criterion, which seeks the minimization of the sum S of the squared differences between the computed and the real energy consumption values, written as:

$$S = \sum_{k=1}^{n} d_k^2 = \sum_{k=1}^{n} [y_k - f(x_k, a_0, a_1, \ldots, a_n)]^2 \tag{42}$$

If the obtained values have different variances, then the measured values were obtained with measurement devices having different precision classes (42) can be rewritten as:

$$S = \sum_{k=1}^{n} d_k^2 = \sum_{k=1}^{n} \left\{ [y_k - f(x_k, a_0, a_1, \ldots, a_n)]^2 \cdot \omega_k \right\} \tag{43}$$

where ω_k are weights inversely proportional with the variance of the measured values, respectively:

$$\omega_1 = \frac{1}{\sigma_1^2}; \quad \omega_2 = \frac{1}{\sigma_2^2}; \quad \ldots \quad \omega_n = \frac{1}{\sigma_n^2} \tag{44}$$

The values $a_0, a_1, \ldots, a_n$, are obtained by minimizing S $(a_0, a_1, \ldots, a_n)$:

$$\frac{\partial S}{\partial a_0} = 0; \quad \frac{\partial S}{\partial a_1} = 0; \quad \ldots \quad \frac{\partial S}{\partial a_n} = 0 \tag{45}$$

By solving (45), the best regression coefficients are determined for a function family $y = f(x)$. The direct extrapolation procedure used for determination the best regression coefficients for the load estimation is illustrated in the following for the logistic and power functions. The logistic function used for the estimation of the trend term in time series has the following expression:

$$y = \frac{a}{1 + b \cdot e^{-c \cdot x}} \tag{46}$$

where a is the limit value of y in time, and can be frequently assessed with non-statistical means.

In order to find a, b and c in (46), a possible approach is to empirically choose three values (y_1, y_2, y_3) which correspond to the (x_1, x_2, x_3) equidistant points illustrated in

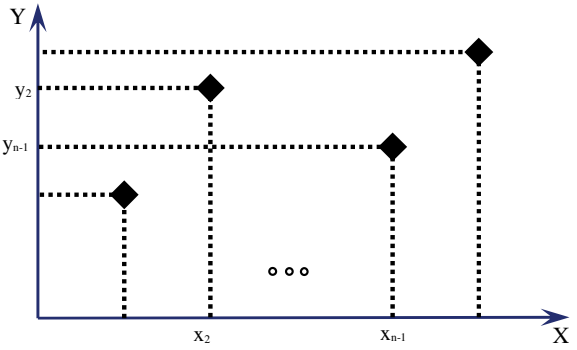


Fig. 7 Representing Y values using equidistant X values

Fig. 7. For simplifying the computation effort, the following notations can be used:

$$x_1 = 0; \quad x_2 = \theta; \quad x_3 = 2\theta \tag{47}$$

Thus, the logistic function (47) can be written:

$$\frac{a - y}{y} = b \cdot e^{-c \cdot x} \tag{48}$$

If $x = x_1 = 0$, then b can be computed with:

$$b = \frac{a - y_1}{y_1} \tag{49}$$

Using the natural logarithm transformation, (48) becomes

$$\ln b - c \cdot x = \ln\left(\frac{a - y_1}{y_1}\right) \tag{50}$$

Similarly, if $x = x_2 = \theta$ and $x = x_3 = 2\,\theta$,

$$\ln b - c\theta = \ln\left(\frac{a - y_2}{y_2}\right); \ \ln b - 2c\theta = \ln\left(\frac{a - y_3}{y_3}\right) \tag{51}$$

By using (50), multiplying the first equation by (-2) and adding it with the second equation from (51), we obtain

$$\frac{a - y_1}{y_1} = \left(\frac{a - y_2}{y_2}\right)^2 \cdot \frac{y_3}{a - y_3} \tag{52}$$

Using (52), a can be written as:

$$a = \frac{2y_1 \cdot y_2 \cdot y_3 - y_2^2(y_1 + y_3)}{y_1 \cdot y_3 - y_2^2} \tag{53}$$

Once a from the logistic function (46) is computed using (53), b can be determined with (49), and c with (50), follows using:

$$c\theta = \ln b - \ln\left(\frac{a - y_2}{y_2}\right) = \frac{a - y_1}{y_1} - \ln\left(\frac{a - y_2}{y_2}\right) \tag{54}$$

or

$$c\theta = \ln \frac{a - y_1 y_2}{a - y_2 y_1}; \quad c = \frac{1}{\theta} \cdot 2.3026 \cdot \log \frac{y_2(a - y_1)}{y_1(a - y_2)} \tag{55}$$

Knowing a, b and c, the logistic function can be computed for any each value of the variable x.

As presented in the literature [21–23], the logistic function can be used for yearly estimations only for longer intervals (8–10 years), especially for consumer categories with similar appliances and demand profiles. As for the use of the power function in load extrapolation, its initial expression is

$$y = a \cdot x^b \tag{56}$$

By using the transformation of natural logarithm, we get

$$y = \ln(A) + B \ln(x) \tag{57}$$

and by substituting $Y = \ln y$; $a = \ln A$; $X = \ln x$; $B = b$, a linear function is obtained:

$$Y = b \cdot X + a \tag{58}$$

The best regression curve fulfils the least mean square criterion:

$$S = \sum_{k=1}^{l} (Y_k - b \cdot X_k - a)^2 \rightarrow \min \tag{59}$$

To find the minimum value of S, it's the first order derivatives in report with a and b must be set to zero ($\partial S/\partial a = 0$; $\partial S/\partial b = 0$), which gives the following equations system:

$$\begin{cases} b \cdot \sum\limits_{i=1}^{l} \ln x_i + m \cdot a = \sum\limits_{i=1}^{l} \ln y_i \\ b \cdot \sum\limits_{i=1}^{l} (\ln x_i)^2 + m \cdot a \cdot \sum\limits_{i=1}^{l} \ln x_i = \sum\limits_{i=1}^{l} \ln x_i \cdot \ln y_i \end{cases} \tag{60}$$

By solving the linear equations system (60), the power function coefficients are obtained:

$$a = \frac{\sum_{i=1}^{l} \ln y_i \cdot \sum_{i=1}^{l} (\ln x_i)^2 - \sum_{i=1}^{l} \ln x_i \cdot \sum_{i=1}^{l} \ln x_i \cdot \ln y_i}{m \cdot \sum_{i=1}^{l} (\ln x_i)^2 - \left(\sum_{i=1}^{l} \ln x_i\right)^2} \tag{61}$$

$$b = \frac{m \cdot \sum_{i=1}^{l} \ln x_i \cdot \ln y_i - \sum_{i=1}^{l} \ln x_i \cdot \sum_{i=1}^{l} \ln y_i}{m \cdot \sum_{i=1}^{l} (\ln x_i)^2 - \left(\sum_{i=1}^{l} \ln x_i\right)^2} \tag{62}$$

The Romanian standards recommends the use a power function for residential load estimation:

$$P(t) = A \cdot t^b = P(t) \cdot t^b \tag{63}$$

If it is considered the 2000–2030 interval, the signification of terms from (63) is the following: *P(t)*—the estimated load for year *t*; *t*—a year from the range [2000, 2020], ($t = 1$ for year 2000); $A = P(t)$—the demand in the first year (2000), used as base value; *b*—regression coefficient, based on historical data, whose value differs for each consumer category.

The estimation functions for the demand evolution in urban areas, considered as power required by MV/LV electric substations, maximum and minimum value, are given in [24] for the 2000–2035 interval. It should also be noted that for the estimation of the demand for the apartments found in crowded areas or in individual buildings more than 4 levels, the following supplemental values should be added: for staircase lighting—0.2 kW/store (4/6 apartments); elevators—10 kW/drive; fire hose enclosure lighting:—2 kW/entrance.

The choice between the maximum and the minimum value should be made in the design stage, taking into account the geographical area, the economic environment, consumer density etc. [5, 6].

3.3.2 Testing the Solution

Using the capabilities of the Smart Meters, which can record consumption values, data was recorded for seven consecutive years (2012–2018) on the LV side of four MV/LV electric substations located in an electric distribution network belonging of a DNO from Romania. The monitored substations supply 390 apartments with 2 and 3 room apartments.

A first category (Group I) contains 205 apartments which use natural gas for cooking and receive hot water and heating from the central thermal power plant. The

second category (Group II) contains 185 apartments which use natural gas for cooking and individual thermal plants for hot water and heating. Table 3 and Fig. 8 show the electricity demand evolution measured in the four monitored MV/LV substations, as measured by the smart meters.

Initially, in order to identify the most representative mathematical model for the load estimation, as described in the previous sections, continuous growth functions (linear, parabolic, polynomial, exponential) limited growth (power, logarithmic, modified exponential, logistic) and modified combinations functions were used.

In the second stage, the regression coefficients were determined for each function and apartment category, using the time as interest variable and the minimum least square criterion. The results confirmed that the power function has the smallest sum of squared estimate of errors (SSE), confirming the validity of the estimation function type recommended in the standards. However, the regression coefficients differ slightly:

$$\text{Group I}: \; W(t) = 0.420 \cdot t^{0.201} \tag{64}$$

Table 3 The demand evolution for each apartment category, [kW/ap]

Year	Group I	Group II
2012	0.420	0.467
2013	0.482	0.543
2014	0.523	0.594
2015	0.554	0.632
2016	0.58	0.664
2017	0.602	0.691
2018	0.637	0.736

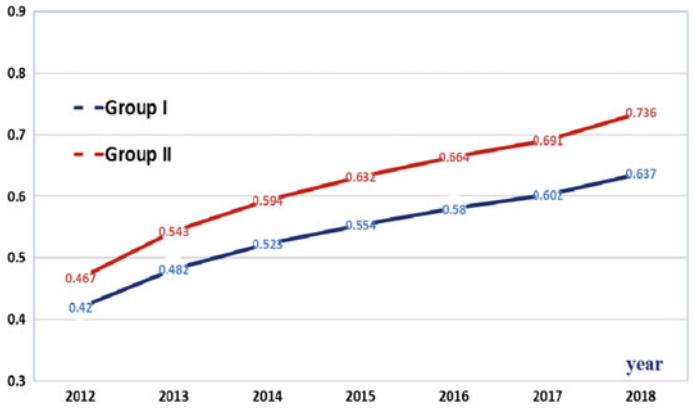


Fig. 8 The demand evolution measured at the LV side of electric substation, for the considered apartment categories and years 2012–2018

Table 4 The demand evolution on the LV side of electric substations, for the considered apartment categories using new and the recommended coefficients

Year	Calculated values		The extreme values of range from the Romanian normative	
	Group I	Group II	Minimum	Maximum
2012	0.420	0.467	0.450	0.520
2013	0.482	0.543	0.500	0.580
2014	0.523	0.594	0.540	0.630
2015	0.554	0.632	0.570	0.670
2016	0.580	0.664	0.600	0.710
2017	0.602	0.691	0.630	0.739
2018	0.637	0.736	0.680	0.800
2019	0.667	0.773	0.727	0.851
2020	0.692	0.804	0.768	0.899
2021	0.703	0.818	0.790	0.920
2022	0.713	0.832	0.804	0.942
2023	0.733	0.857	0.840	0.980
2024	0.750	0.879	0.870	1.020
2025	0.767	0.899	0.899	1.053
2026	0.781	0.918	0.927	1.085
2027	0.788	0.927	0.940	1.100
2028	0.804	0.947	0.969	1.133
2029	0.816	0.963	0.992	1.160
2030	0.829	0.979	1.016	1.187

$$\text{Group II}: \; W(t) = 0.467 \cdot t^{0.219} \tag{65}$$

It should be noted that in the Romanian standard, the power function coefficients have different values according to the number of rooms in the apartment and the heating/cooking type, as described earlier.

For the apartment types used in the study case, two different coefficient sets are provided:

$$\text{Minimal } W(t) = 0.305 \cdot t^{0.35} \tag{66}$$

$$\text{Maximal } W(t) = 0.357 \cdot t^{0.35} \tag{67}$$

For a comparative analysis of the coefficients associated the estimation function obtained in the study case, relations (64) and (65), and given in the standards, relations

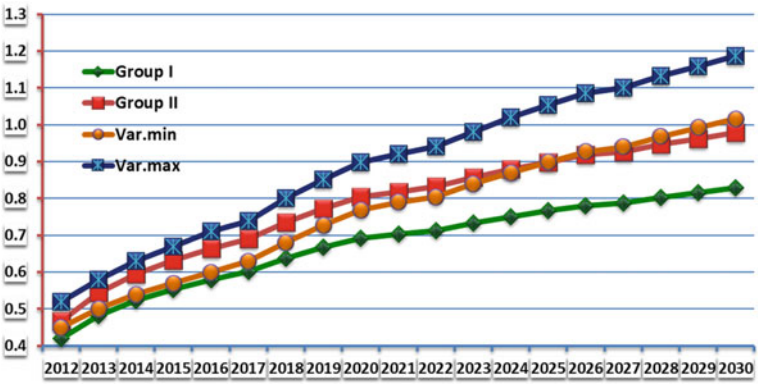


Fig. 9 The demand evolution for the categories of considered apartment

(66) and (67), Table 4 and Fig. 9 present the demand evolution on the LV side of electric substations, between 2012 and 2030 (estimated).

The following conclusions can be highlighted:

- The coefficients of the power function computed in the study case have different values comparative with those from the normative.
- The new estimated values are inside the range given in [24] for the apartments using individual thermal substations.

4 Conclusions

The estimation of the loads in different parts of the distribution system represents a main function of the DNOs. The electricity cannot be efficiently stored on a large scale (relative to the produced amount), which means that for the DNOs, the estimation of the loads is an indispensable factor in the distribution process. The regression models are some of the most commonly used statistical techniques. For the estimation of electricity/power consumption such approaches are used to model the relationship between consumption and other factors such as weather, type of day, nature of consumption, etc. Usually RL model is used using in most cases the temperature. The advantages of this model are related to the relatively simple implementation, the easy understanding of the relationship between the input and output variables and the easy estimation of the performance of the forecasting model. However, due to the complex dependence between electricity consumption and influence factors, inherent problems arise in identifying the correct model.

To solve this problem, the regression analysis-based approaches for the load modelling from the nodes of electricity distribution networks were treated in the chapter. The approaches refer to estimation of the required powers in the supply

points with a mixt structure of the load (i.e. residential, commercial, and industrial) at the hour when the maximum value of the load is recorded and the demand of residential consumers which represent the highest percentage from the load structure fed from the LV/MV electric substations. The proposed approaches were tested in real operation conditions of MV distribution networks from Romania. Thus, the estimation of the loads from the MV/LV electric substations of a test network, at the hour when the maximum value (peak load) was recorded, using the proposed method based on the power correlation, led at an average error for the time frame 7 h (hPL $-$ 5 h; hPL $+$ 1 h) below 1.48% than in the others frames L_{24} frame or L_7 frames ($h_{PL} \pm 3\ h$), ($h_{PL} - 4\ h$; $h_{PL} + 2\ h$).

Regarding the estimation of the demand in the case of residential consumers, the comparative analysis of the coefficients associated the estimation function and to those given in the Romanian standard highlighted that the estimated values are lower than the minimum recommended values for the apartments which use natural gas for cooking and individual thermal plants for hot water and heating. This behaviour could be the result of a modified behaviour of the customers or due to the used database which belonging of a characteristic electric substation from the analysed area, while in used data from the Romanian standard are collected from the whole country.

References

1. Phuangpornpitak N, Prommee W (2016) A study of load demand forecasting model in electric power system operating and planning. Greater Mekong Subreg Acad Res Netw Int J 10:19–24
2. Neagu B, Georgescu G (2010) Load and energy forecast on a proximate, medium and long horizon in public electricity repartition and distribution systems. Buletinul Institutului Politehnic din Iaşi LVI (LX) 3:71–82
3. Li K, Wang B, Wang Z, Wang F, Mi Z, Zhen Z (2017) A baseline load estimation approach for residential customer based on load pattern clustering. Energy Procedia 142:2042–2049
4. Avdakovic S, Ademovic A, Nuhanovic A (2013) Correlation between air temperature and electricity demand by linear regression and wavelet coherence approach: UK, Slovakia and Bosnia and Herzegovina case study. Arch Electr Eng 62(4):521–532
5. Vu D, Muttaqi KM, Agalgaonkar AP (2014) Assessing the influence of climatic variables on electricity demand. In: IEEE power and energy society general meeting, Washington, USA, 2014, pp 1–5
6. Mfonobong Umoren A, Okpura N, Markson I (2017) Rural electrification peak load demand forecast model based on end user demographic data. Math Softw Eng 3(1):87–98
7. Grigoraș G, Cârțină G (2013) The fuzzy correlation approach in operation of electrical distribution systems. Int J Comput Math Electr Electron Eng 32(3):1044–1066
8. Grigoraş G, Scarlatache F, Neagu BC (2010) Clustering in power systems. Applications. Lambert Academic Publishing, Germany
9. Chen C, Zhou JN (2014) Application of regression analysis in power system load forecasting. Adv Mater Res 960(961):1516–1522
10. Zheng R, Zhijian JG, Hongqiao Peng J, Zhu Y (2019) Regression analysis of time series for forecasting the electricity consumption of small consumers in case of an hourly pricing system. Int Trans Electr Energy Syst e12100033
11. Gelman A, Hill J (2007) Data analysis using regression and multi-level/hierarchical models. Cambridge University Press, New York

12. Jaba E (2000) Statistics (in Romanian). Economic Publishing, Bucharest, Romania
13. Nagumo T, Ito H, Sano T (2017) Load current forecasting using statistical analysis. In: 24th international conference & exhibition on electricity distribution (CIRED), Glasgow, Scotland
14. Supapo KRM, Santiago RVM, Pacis MC (2017) Electric load demand forecasting for Aborlan-Narra-Quezon distribution grid in Palawan using multiple linear regression. In: IEEE 9th international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM), Manila, Philippines
15. Bai J, Jiang M, Liu L, Sun Y, Wang Y, Zhang J (2019) Correlation analysis and prediction of power network loss based on mutual information and artificial neural network. In: IOP conference series: earth and environmental science [Online]. https://iopscience.iop.org/article/. https://doi.org/10.1088/1755-1315/227/3/032023/pdf
16. Rahman H, Selvarasan I, Begum J (2018) A short-term forecasting of total energy consumption for India-A black box based approach. Energies 11, paper 2442 [Online]. https://www.mdpi.com/1996-1073/11/12/3442
17. Almeshaiei E, Soltan H (2011) A methodology for electric power load forecasting. Alexandria Eng J 50(2):137–144
18. Azad MK, Uddin S, Takruri M (2018) Support vector regression based electricity peak load forecasting. In: IEEE 11th international symposium on mechatronics and its applications (ISMA), Sharjah, United Arab Emirates
19. Abdulkareem A, Okoroafor EJ, Awelewa A, Adekitan A (2019) Pseudo-inverse matrix model for estimating long-term annual peak electricity demand: the Covenant University's experience. Int J Energy Econ Policy 9(4):103–109
20. Friedrich L, Armstrong P, Afshari A (2014) Mid-term forecasting of urban electricity load to isolate air-conditioning impact. Energy Build 80:72–80
21. Rabie AH, Saleh AI, Abo-Al-Ez KM (2015) A new strategy of load forecasting technique for smart grids. Int J Modern Trends Eng Res 2(12):332–341
22. Muller CJ, MacLehose RF (2014) Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. Int J Epidemiol 43(3):962–970
23. Gajowniczek K, Ząbkowski T (2017) Two-stage electricity demand modeling using machine learning algorithms. Energies 10(10):1547
24. PE 132/2003, Normative of design of the public distribution networks, S.C. ELECTRICA S.A., Bucharest (2003)

Finite Element Analysis of Electromagnetic Fields Emitted by Overhead High-Voltage Power Lines

Eduard Lunca, Bogdan Constantin Neagu, and Silviu Vornicu

Abstract The overhead high-voltage power lines (OHVPLs) are considered significant sources of extremely low frequency (ELF) electric and magnetic fields (EMFs), whose potential health effects became during the past decades a matter of scientific debate and public concern all over the world. In this chapter, a simple and yet effective finite element (FE) approach is proposed to compute and analyze—from the perspective of public exposure—both electric and magnetic fields emitted by typical configurations of OHVPLs belonging to the Romanian power grid. First, a 2D ANSYS Maxwell model is developed for the specific instance of a 110 kV double-circuit OHVPL and validated against two software tools based on quasi-static analytical methods, PowerELT and PowerMAG. Next, it will be used to investigate exposure to ELF-EMFs emitted by a selection of OHVPLs with nominal voltages of 110 kV, 220 kV and 400 kV, taking into consideration influencing factors such as loading, phasing and ground clearance. Compliance with the exposure guidelines specified by the International Commission on Non-Ionizing Radiation Protection (ICNIRP) for general public is assessed for each particular case. As a result, all calculated magnetic fields are below the ICNIRP limit of 100 μT, while the electric fields exceed the ICNIRP limit of 5000 V/m only in limited areas beneath the 400 kV OHVPLs. The calculated field levels are in line with those reported in the scientific literature for similar OHVPLs.

E. Lunca (✉) · S. Vornicu
Department of Electrical Measurements and Materials, Gheorghe Asachi Technical University of Iasi, Iasi, Romania
e-mail: elunca@tuiasi.ro

S. Vornicu
e-mail: silviusieca@gmail.com

B. C. Neagu
Power System Department, Electrical Engineering Faculty, Gheorghe Asachi Technical University of Iasi, Iasi, Romania
e-mail: bogdan.neagu@tuiasi.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_29

Keywords Overhead high-voltage power line · Electric field · Magnetic field · Public exposure · 2D ANSYS maxwell model

Abbreviations

A. Acronyms

2D	Two-Dimensional
ACSR	Aluminum Conductor Steel-Reinforced
ELF	Extremely Low Frequency
EMF	Electric and Magnetic Fields
EU	European Union
FE	Finite Element
FEM	Finite Element Method
IARC	International Agency for Research on Cancer
ICNIRP	International Commission on Non-Ionizing Radiation Protection
IEEE	Institute of Electrical and Electronics Engineers
OHVPL	Overhead High-Voltage Power Line
RMS	Root Mean Square
SW	Shield Wire
T	Transposed
U	Untransposed
WHO	World Health Organization

B. Symbols/Parameters

B	Magnetic flux density
E	Electric field strength
i	Conductor number
I_i	Phase current of conductor i
U_i	Phase voltage of conductor i
d_i	Lateral distance from centerline to conductor i
h_i	Height of conductor i
h_g	Line-to-ground clearance
h	Calculation height above ground
R_{eq}	Bundle conductor equivalent radius

1 General

The electricity has many benefits in our daily life. But generating, transmitting, distributing and using electricity can expose people to ELF-EMFs, which interact with the human body by mainly inducing electric currents in it. During the past

decades, a lot of research has been devoted to investigation of possible health effects of exposure to ELF-EMFs, including childhood and adult cancers, reproductive dysfunctions, cardiovascular and developmental disorders, immunological modifications, neurological effects, etc. Particularly, a (poor) statistical link between childhood leukemia and prolonged exposure to residential ELF magnetic fields higher than 0.3–0.4 μT has been reported by a number of epidemiological studies [1, 2]. In 2002, based on these findings, the International Agency for Research on Cancer (IARC)—an intergovernmental agency activating within the World Health Organization (WHO)—has concluded that the ELF magnetic fields are "possibly carcinogenic to humans" (Group 2B carcinogens, designating agents for which the evidence in humans is limited and the evidence in animals is "less than sufficient"). As for ELF electric fields, IARC has concluded that they are "unclassifiable as to carcinogenicity in humans" (Group 3 carcinogens) [3].

Aiming at preventing the established health effects associated with short-term exposure to high intensity ELF-EMFs, principally induced currents, ICNIRP and IEEE (the Institute of Electrical and Electronics Engineers) have formulated exposure guidelines in 1998 [4] and 2002 [5], respectively. According to the scientific information currently available, long-term exposure to ELF field levels not exceeding the limits prescribed by these guidelines is considered safe for the purpose of protecting human health. There is no established evidence that exposure to ELF-EMFs emitted by power lines, substations, transformers or other electrical equipment, regardless of the proximity, can cause any known health effects. But there is a continuous debate as to what might be adequate precautionary approaches at these lower field levels. Furthermore, the general public often expresses concern about ELF-EMFs, especially in relation with setting up new overhead high-voltage power lines or living in their vicinity [6–9].

The OHVPLs are considered significant sources of both electric and magnetic fields. Both fields are strongest directly under the OHVPL and sharply reduce with distance from it. Of course, in addition to distance, there are many other factors influencing the ELF-EMFs originating from OHVPLs, including voltage, current, phasing, ground clearance, observation height above the ground, balance within circuit, balance between circuits, conductor bundle, existence of parallel lines, ground resistivity (conductivity), etc. Moreover, the electric fields are greatly attenuated by buildings, walls, fences, trees and other obstacles in the neighborhood, but the magnetic fields pass through most materials and cannot be attenuated as easily as the electric fields [10–12].

To determine ELF electric and magnetic field levels emitted by OHVPLs and to assess compliance with relevant exposure limits, both measurements and computations can be performed [13–19]. Computations are often preferable to measurements because they can be conducted for any desired exposure scenario rather than being confined to the particular conditions at the time of taking measurements. Analytical and numerical methods can be used for computations, usually employing two-dimensional (2D) models because of their simplicity [20–33]. Very often, the numerical calculations (simulations) exploit the finite element method (FEM), which is

recognized for its ability to generate accurate 2D electric and magnetic field distributions in the transverse section of the OHVPLs and of other power–frequency systems [19, 29–33].

In this chapter, a simple and yet effective FEM approach is proposed to compute and analyze—from the perspective of public exposure—ELF electric and magnetic fields produced by typical configurations of OHVPLs belonging to the Romanian power grid. Computations are performed with ANSYS Maxwell 2D electromagnetic simulation software, mainly in the form of RMS electric field strength and RMS magnetic flux density lateral profiles, at the standard height $h = 1$ m above the ground level. It is worthwhile to remark that Romania, as a member of the European Union (EU), has implemented exposure limits derived from the *Council Recommendation of 12 July 1999 on the limitation of exposure of the general public to electromagnetic fields (0 Hz–300 GHz)* [34], which is based on the guidelines issued by ICNIRP in 1998. For power–frequency electric and magnetic fields, these limits are 5000 V/m and 100 μT, respectively.

From this point, the chapter is organized as follows. First, a 2D ANSYS Maxwell model for computing ELF electric and magnetic fields around OHVPLs will be developed and validated against simulation software based on analytical methods. Next, it will be used to investigate exposure to ELF-EMFs generated by a selection of OHVPLs with nominal voltages of 110, 220 and 400 kV. As already mentioned, special attention will be given to the field distribution at 1 m height above the ground, taking into consideration influencing factors such as loading, phasing and ground clearance. Compliance with the ICNIRP exposure limits for general public will be assessed for each particular case.

2 2D ANSYS Maxwell Model for Computing ELF Electric and Magnetic Fields Around OHVPLs

ANSYS Maxwell 2D is a high-performance low frequency electromagnetic field simulation software that uses the finite element method for solving electric, magnetostatic, eddy current and transient problems. Therefore, it may serve as an appropriate tool for computing exposure to ELF-EMFs originating from OHVPLs, but such investigations are rather rare and mostly focused only on the magnetic field exposure [17, 35, 36]. As an extension of a recent study by the authors [11], this section presents the development and validation of a 2D Maxwell model for computing both electric and magnetic fields emitted by various OHVPLs. The model is implemented for a common configuration of 110 kV double-circuit line used for primary power distribution, but it can easily be applied to any other OHVPL. The model validation is achieved against previously developed software based on quasi-static analytical methods.

2.1 Model Development

The 110 kV double-circuit OHVPL selected for model implementation—often found in the proximity of urban settings—has geometry dictated by suspension towers of Sn 110.252 type, as illustrated in Fig. 1. The phases of the two circuits are realized with Aluminum Conductor Steel-Reinforced (ACSR) cables of Sect. 240/40 mm^2 (21.7 mm exterior diameter), while the shield wire (SW) is represented by a 160/95 mm^2 ACSR conductor (20.75 mm exterior diameter). The OHVPL is considered to operate at a load of 500 A (close to the maximum rated current), with the phases of the two circuits perfectly balanced. In addition, because the field level largely depends on the relative phasing between the two circuits, we assumed both untransposed (ABC/A'B'C') and directly transposed (ABC/C'B'A') phase arrangements, which

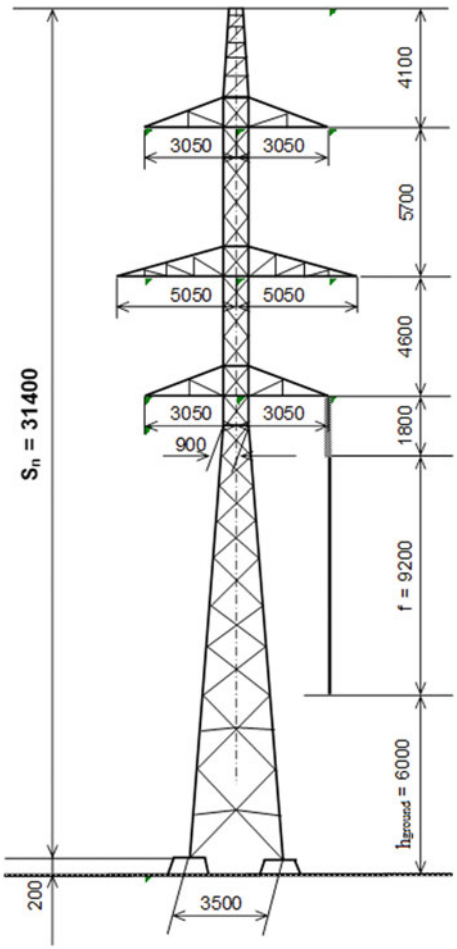


Fig. 1 Suspension tower of Sn 110.252 type (dimensions are given in mm)

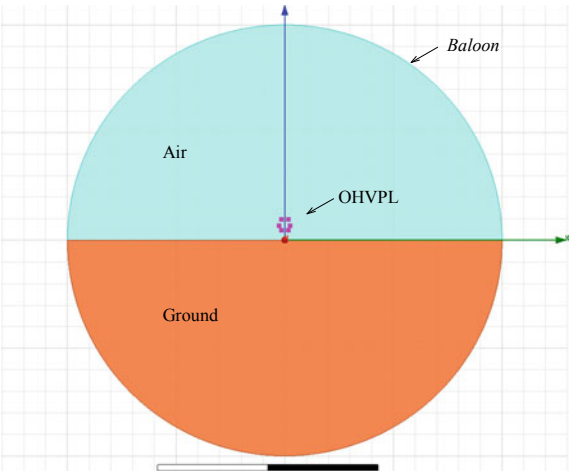


Fig. 2 Global FEM geometric model of the 110 kV double-circuit OHVPL

clearly determine the maximum and minimum exposure to the sides of the line [37, 38], where people are most likely to live or spend time.

a. *FEM geometric model*

The global FEM geometric model is presented in Fig. 2, where the considered 110 kV double-circuit OHVPL is placed above a ground with the electrical conductivity of 0.01 S/m, the relative electric permittivity of 10 and the relative magnetic permeability of 1. The ground clearance, namely 9 m, corresponds to an "average height" of the OHVPL above the ground, calculated as $h_{avg} = h_{max}$—(2/3)·f [39], where $h_{max} = 15.2$ m represents the maximum height of the conductors (at tower) and $f = 9.2$ m is the conductors sag. The active conductors are modeled as presented in Fig. 3, as simple aluminum cylinders with the electrical conductivity of $3.8 \cdot 10^7$ S/m, the relative electric permittivity of 1 and the relative magnetic permeability of 1, while the influence of the SW on the electric and magnetic field distribution is ignored (the SW is not included in simulation). The applied boundary conditions are of Balloon type, which models the region outside the defined space as extending to infinity. The radius of the bounded region is taken $R = 200$ m, sufficiently large to determine the behavior of the two fields well outside the power line corridor, even for OHVPLs with higher nominal voltages. All simulations conducted in this study assume a total number of mesh elements of 1,223,286, but it can be lowered for more rapid and yet satisfactory analyzes.

b. *Magnetic field calculation*

The magnetic field distribution around the OHVPL is obtained using the *eddy current solver*, which allows calculating magnetic fields that oscillate with a frequency (in this case, 50 Hz). However, because the magnetic field distributions generated with this solver are reported in terms of instantaneous magnetic flux density values over a 20 ms period, further post-processing is necessary to generate RMS

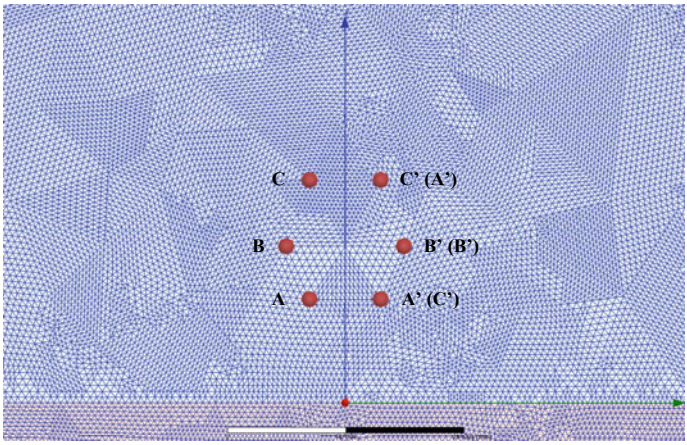


Fig. 3 Discretized OHVPL region

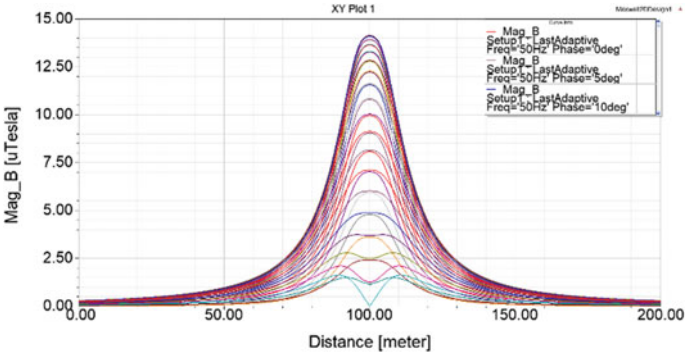


Fig. 4 Instantaneous magnetic flux density profiles at $h = 1$ m, for U phasing

magnetic flux density (lateral) profiles at the height $h = 1$ m, as often used for assessing exposure to ELF-EMFs from overhead power lines. Consequently, multiple instantaneous magnetic flux density profiles have been imported into Microsoft Excel, where they have been processed in a point-by-point fashion, according to the formula [11]:

$$B_{RMS}(i) = \sqrt{\frac{1}{N}\sum_{n=1}^{N} B_n^2(i)}, \tag{1}$$

where $B_1(i), \ldots, B_N(i)$ represent the instantaneous values of the magnetic flux density corresponding to the point i of the profile and $N = 73$ is the total number of values.

Figure 4 presents instantaneous magnetic flux density profiles obtained for untransposed (U) phasing, while Fig. 5 shows similar profiles obtained for transposed (T) phasing. The correspondent RMS magnetic flux density profiles—computed with Eq. (1)—are comparatively presented in Fig. 6. Starting at a certain distance from the centerline, any other phase arrangement will generate an RMS magnetic flux density profile between these two limit plots.

Figures 7 and 8 illustrate the magnetic field distribution around the OHVPL at the time instants corresponding to the maximum field profiles in Figs. 4 and 5, respectively. As evident, an extra degree of cancellation between the magnetic fields generated by the two circuits can be observed for transposed phasing. The distance from the centerline is 30 m in both distributions.

c. *Electric field calculation*

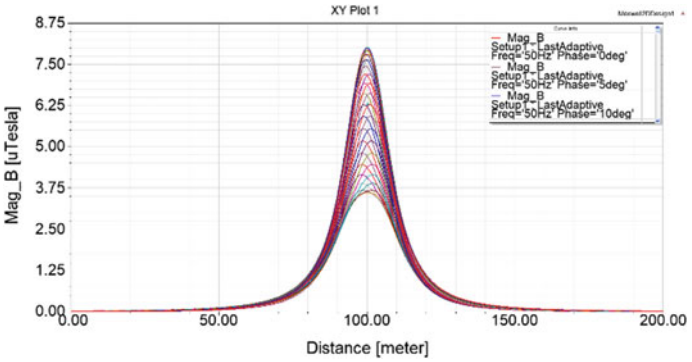


Fig. 5 Instantaneous magnetic flux density profiles at $h = 1$ m, for T phasing

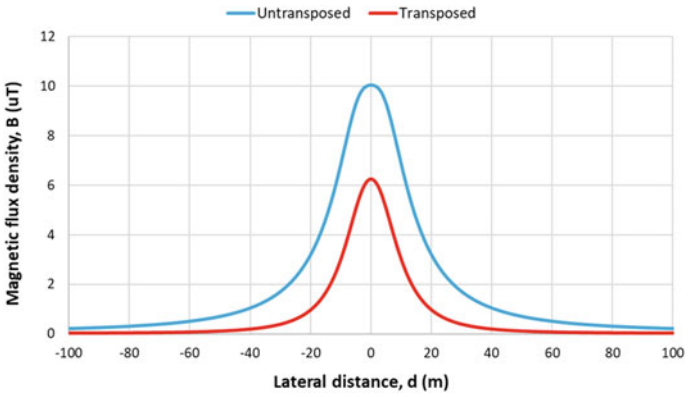


Fig. 6 Comparison between RMS magnetic flux density profiles obtained for untransposed and transposed phasing

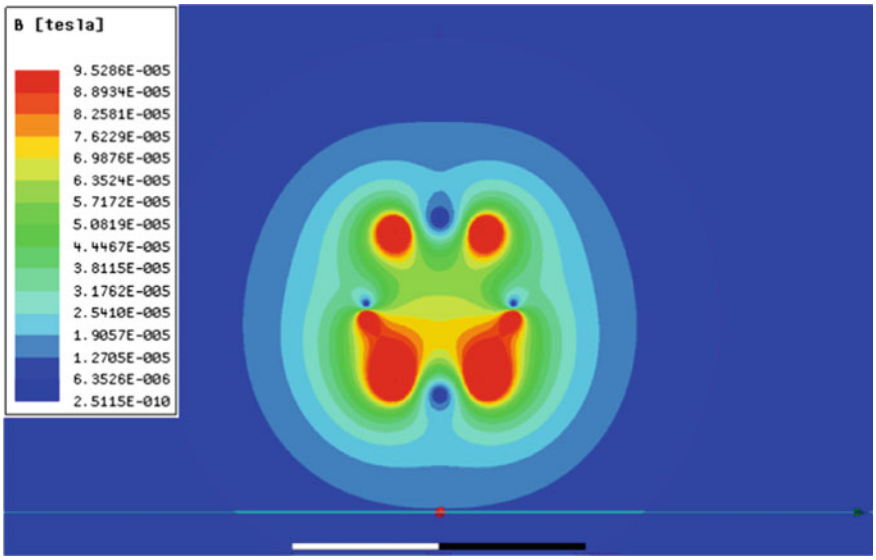


Fig. 7 The magnetic field distribution around the 110 kV double-circuit OHVPL for U phasing ($t = 10.83$ ms)

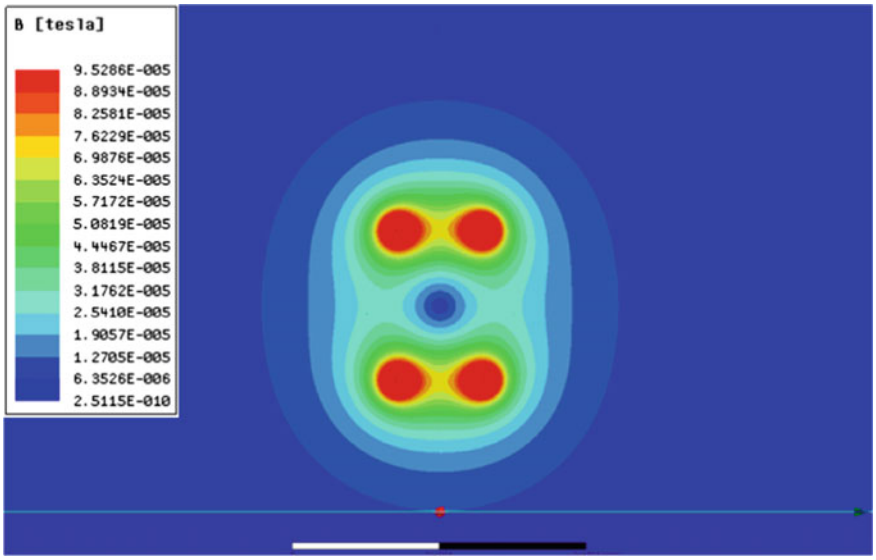


Fig. 8 The magnetic field distribution around the 110 kV double-circuit OHVPL for T phasing ($t = 11.66$ ms)

The electric field distribution around the OHVPL is determined using the *AC conduction solver*, which allows calculating sinusoidally-varying electric fields (here, varying at 50 Hz). And this time, to generate RMS electric field strength lateral profiles at the height $h = 1$ m, multiple profiles of instantaneous electric field strength have been imported into Microsoft Excel, where they have been processed in the same way, by applying the formula [11]:

$$E_{RMS}(i) = \sqrt{\frac{1}{N}\sum_{n=1}^{N} E_n^2(i)} \tag{2}$$

where $E_1(i), \ldots, E_N(i)$ represent the instantaneous values of the electric field strength corresponding to the point i of the profile and $N = 73$ is the total number of values.

As in the case of magnetic field, Fig. 9 gives instantaneous electric field strength profiles obtained for U phasing, while Fig. 10 shows profiles obtained for T phasing. The two associated RMS electric field strength profiles—computed with Eq. (2)—are compared in Fig. 11.

Figures 12 and 13 illustrate the momentary distribution of the electric field around the OHVPL corresponding to the maximum field profiles in Figs. 9 and 10, respectively. Once again, an extra degree of cancellation between the electric fields generated by the two circuits can be observed for transposed phasing. As in the case of magnetic field, the distance from the centerline is 30 m in both distributions.

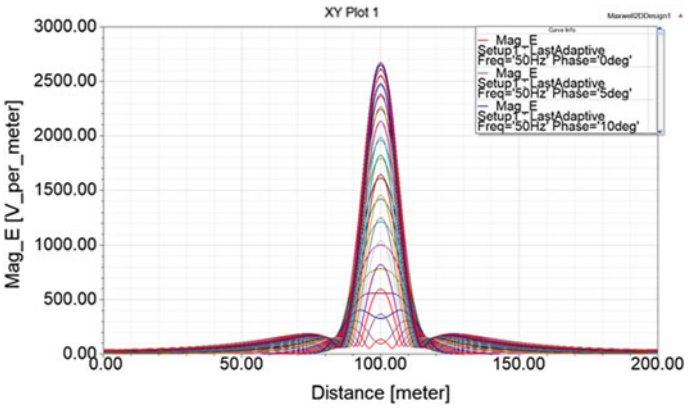


Fig. 9 Instantaneous electric field strength profiles at $h = 1$ m, for U phasing

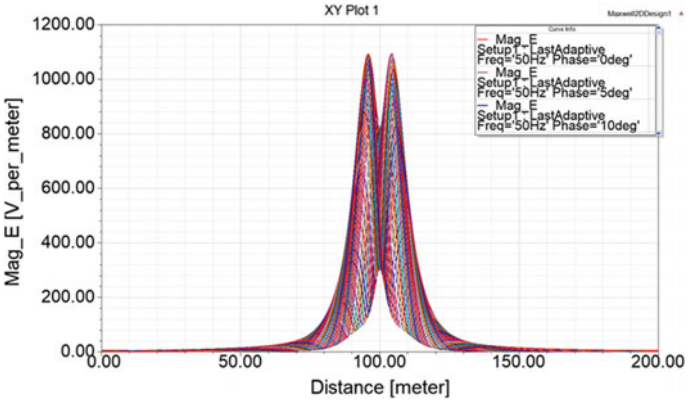


Fig. 10 Instantane-ous electric field strength profiles at $h = 1$ m, for T phasing

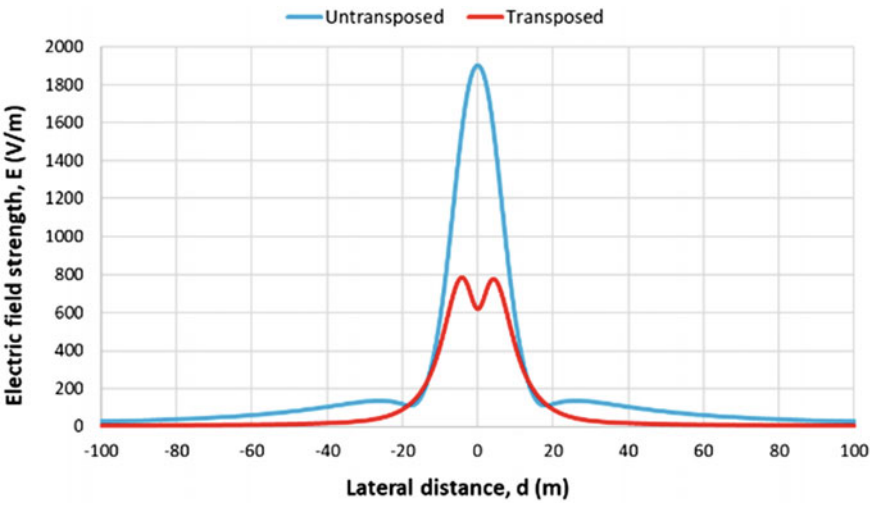


Fig. 11 Comparison between RMS electric field strength profiles obtained for untransposed and transposed phasing

2.2 Model Validation

The model validation has mainly been performed with the help of two interactive software tools based on analytical methods, *PowerELT* and *PowerMAG* [28], which are capable to produce accurate electric and magnetic field (lateral) profiles at any user-defined height above the ground level, together with 2D electric and magnetic field distributions in the transverse section of the OHVPL, in any rectangular plotting area also defined by user. Assuming the same power line geometry, and voltage and current information (amplitude and phase, respectively), Fig. 14 compares RMS

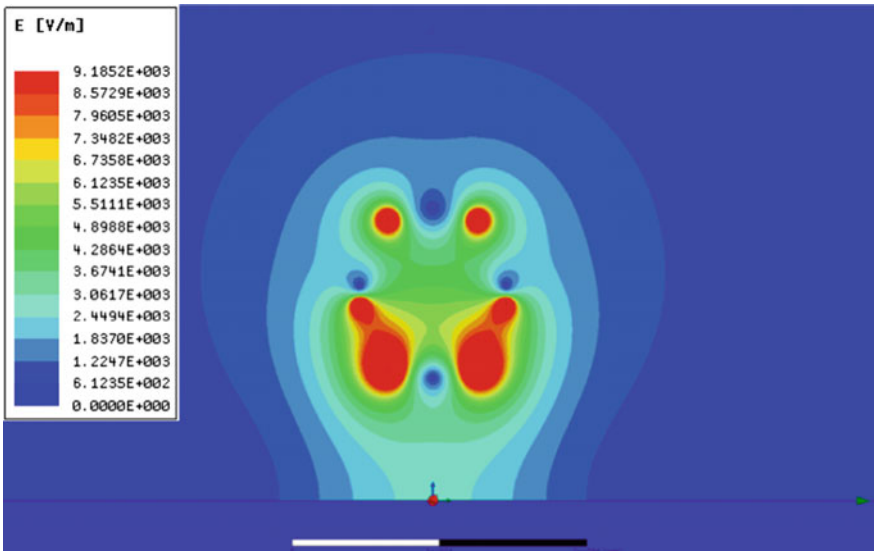


Fig. 12 The electric field strength distribu-tion around the 110 kV OHVPL for U phasing (t = 10.55 ms)

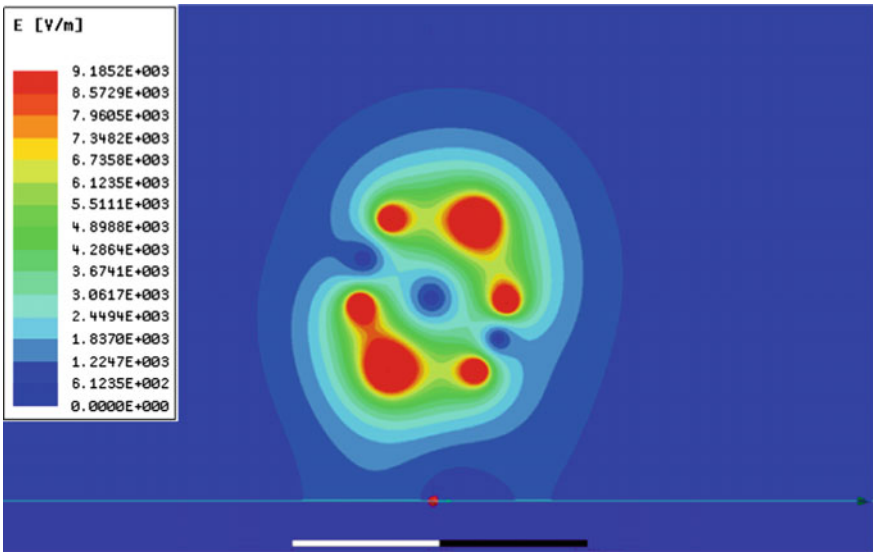


Fig. 13 The electric field strength distribu-tion around the 110 kV OHVPL for T phasing (t = 20 ms)

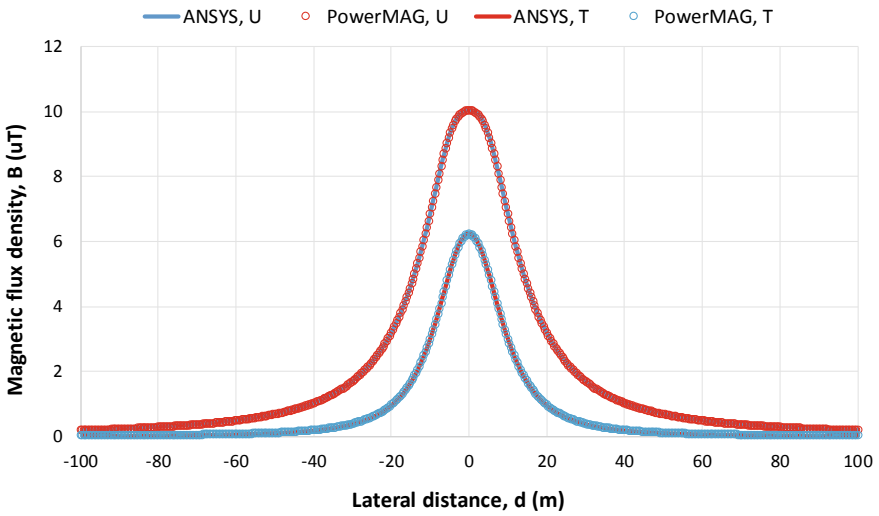


Fig. 14 Comparison between RMS magnetic flux density profiles obtained by numerical simulation and analytical computation

magnetic flux density profiles computed with the developed 2D ANSYS Maxwell model and PowerMAG software. Similarly, Fig. 15 compares RMS electric field strength profiles computed with the developed 2D ANSYS Maxwell model and

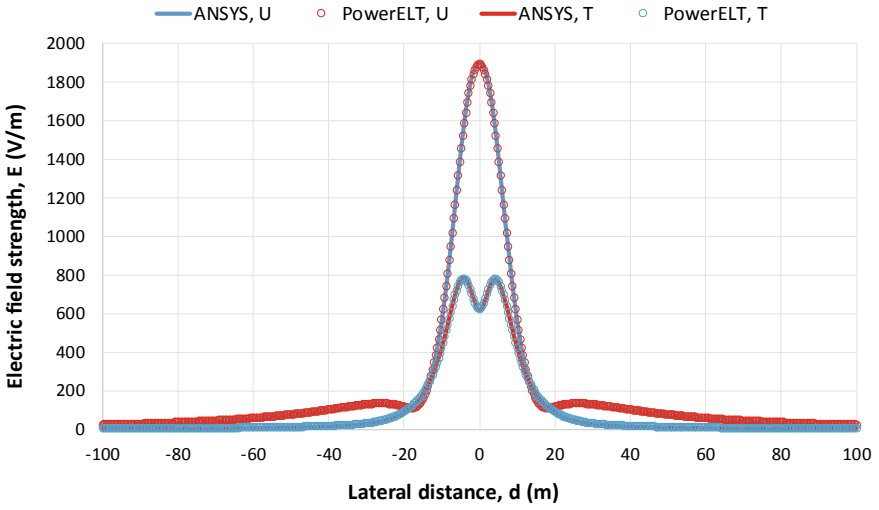


Fig. 15 Comparison between RMS electric field strength profiles obtained by numerical simulation and analytical computation

PowerELT software. As obvious, there is an excellent agreement between numerical and analytical results, regardless the phasing (U or T).

Similar comparisons have also been made for OHVPLs with higher nominal voltages (i.e., larger physical dimensions), each time obtaining perfect matching between simulated profiles. Exposure to ELF-EMFs from some of these lines, under various conditions, will be discussed in the subsequent section.

3 Finite Element Analysis of ELF-EMFs from Typical OHVPLs Used in the Romanian Power Grid

In Romania, OHVPLs are used for both power transmission and power distribution. Power transmission is achieved through a total length of 8759.4 km of OHVPLs, of which [40]: 3.1 km—750 kV, 4915.2 km—400 kV, 3875.6 km—220 kV and 40.4 km—110 kV, where 482.6 km serves as interconnection lines. In addition, power distribution operators make use of more than 20,000 km of OHVPLs operating at 110 kV. Thus, for assessing exposure to ELF-EMFs emitted by these OHVPLs, we have selected two double-circuit lines with nominal voltages of 110 kV and 220 kV, respectively, and a single-circuit line with nominal voltage of 400 kV, which can be considered typical. Finite element analysis of ELF-EMFs from these lines assumes the same conditions as in the described model, except that computations will be performed for three different ground clearances (minimum, average and maximum), as well as for maximum allowable current. However, because of the direct proportionality between current and magnetic flux density, the computed fields can easily be scaled down for more common loads.

3.1 ELF-EMFs from the 110 kV Double-Circuit OHVPL

In essence, the 110 kV double-circuit OHVPL subjected to investigations is the same used for model development, which has geometry dictated by suspension towers of Sn 110.252 type. As already mentioned, the two circuits are realized with standard 240/40 mm^2 ACSR conductors (21.7 mm exterior diameter), which have a maximum allowable current of 575 A (RMS). All other input data for this line are presented in Table 1, where d_i is the lateral distance from the OHVPL centerline to the conductor i, hi represents the height of the conductor i and h_g denotes the ground clearance of the line: 6 m, 9 m and 15.2 m, respectively.

a. *Magnetic field distribution*

Figure 16 presents RMS magnetic flux density profiles at the height $h = 1$ m, for untransposed phasing. The maximum magnetic field beneath the line (not necessarily at the centerline) varies from 5.48 μT for $h_g = 15.2$ m (at tower) to 18.5 μT for $h_g = 6$ m (at mid-span), which is generally below 18.5% of the ICNIRP limit for

Table 1 Input data for the 110 kV double-circuit OHVPL using towers of Sn 110.252 type (untransposed phasing)

Phase conductor	i	d_i [m]	hi [m]	$\underline{U_i}$ [kV]	$\underline{I_i}$ [A]
A	1	−3.05	h_g	63.51 ∠ 0°	575 ∠ 0°
B	2	−5.05	$h_g + 4.6$	63.51 ∠ −120°	575 ∠ −120°
C	3	−3.05	$h_g + 10.3$	63.51 ∠ 120°	575 ∠ 120°
A'	4	3.05	h_g	63.51 ∠ 0°	575 ∠ 0°
B'	5	5.05	$h_g + 4.6$	63.51 ∠ −120°	575 ∠ −120°
C'	6	3.05	$h_g + 10.3$	63.51 ∠ 120°	575 ∠ 120°

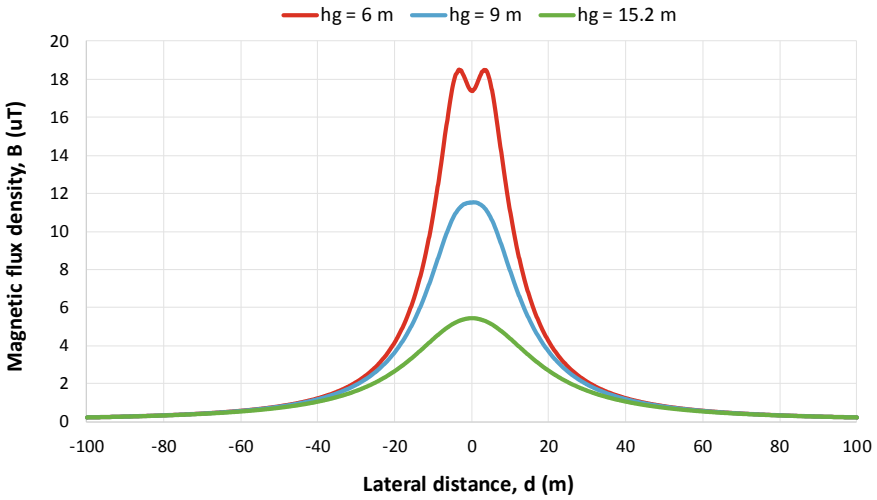


Fig. 16 RMS magne-tic flux density profiles for the 110 kV OHVPL, for untrasnsposed phasing and various ground clearances

general public, 100 μT. For transposed phasing (Fig. 17), the maximum magnetic field beneath the line varies from 2.12 μT for $h_g = 15.2$ m to 16.1 μT for $h_g = 6$ m, hence not exceeding 16.1% of the ICNIRP exposure limit. As it can easily be observed, starting with some distance from the centerline, the transposed phasing produces much lower magnetic field levels.

Considering the average clearance $h_g = 9$ m and a (more) typical loading of 325 A, Table 2 gives magnetic field levels at various distances from the centerline. Beneath the line, the magnetic flux density does not exceed 6.51 μT for U phasing and 4.03 μT for T phasing, while at the edge of the OHVPL corridor—18.5 m from the centerline, according to national regulations [41]—it decreases to 2.32 μT and 0.73 μT, respectively. The critical value of 0.4 μT—often used in epidemiological studies—is reached at a lateral distance of about 53.9 m and 24.5 m, respectively.

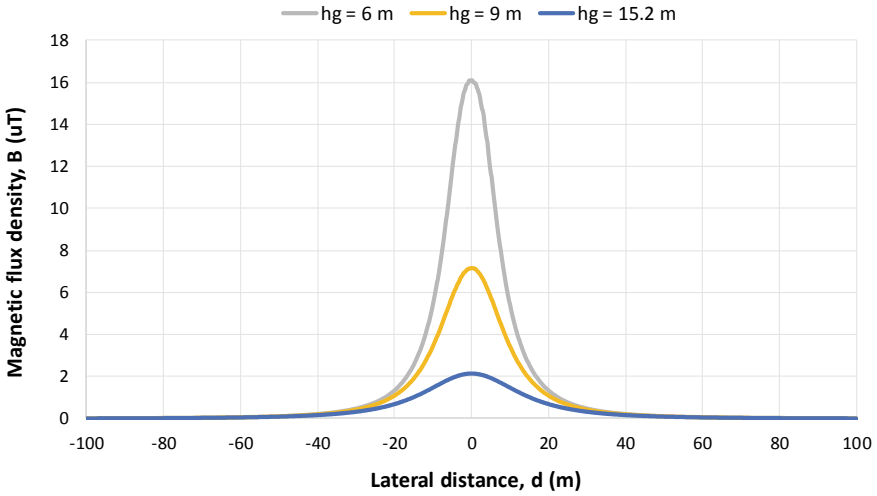


Fig. 17 RMS magne-tic flux density profiles for the 110 kV OHVPL, for transposed phasing and various ground clearances

Table 2 Typical magnetic field levels from the 110 kV OHVPL ($h_g = 9$ m, $I = 325$ A)

Phasing	B [μT], at various lateral distances				
	Maximum beneath line	*18.5 m*	*25 m*	*50 m*	*100 m*
U	6.51	2.32	1.49	0.46	0.14
T	4.03	0.73	0.38	0.067	0.011

b. *Electric field distribution*

The RMS electric field strength distribution at the height $h = 1$ m—for untransposed phasing—is illustrated in Fig. 18. The maximum electric field beneath the line (at the centerline) varies from 882.8 V/m at tower to 3017.2 V/m at mid-span, which is generally below 60.34% of the ICNIRP exposure limit for general public, 5000 V/m. For transposed phasing (Fig. 19), the maximum electric field levels beneath the line range from 243.2 V/m at tower to 1790.2 V/m at mid-span, hence not exceeding 35.8% of the exposure limit.

For the average clearance $h_g = 9$ m (Table 3), the electric field strength beneath the line reaches 1901.8 V/m for U phasing and 784.4 V/m for T phasing, while the electric field strength at the corridor edge reaches 110.3 V/m and 109.5 V/m, respectively. At the distance of 50 m from the centerline route, the electric field levels fall to only 77 and 11.3 V/m, respectively, similar to the lowest levels measured at 30 cm distance from household appliances.

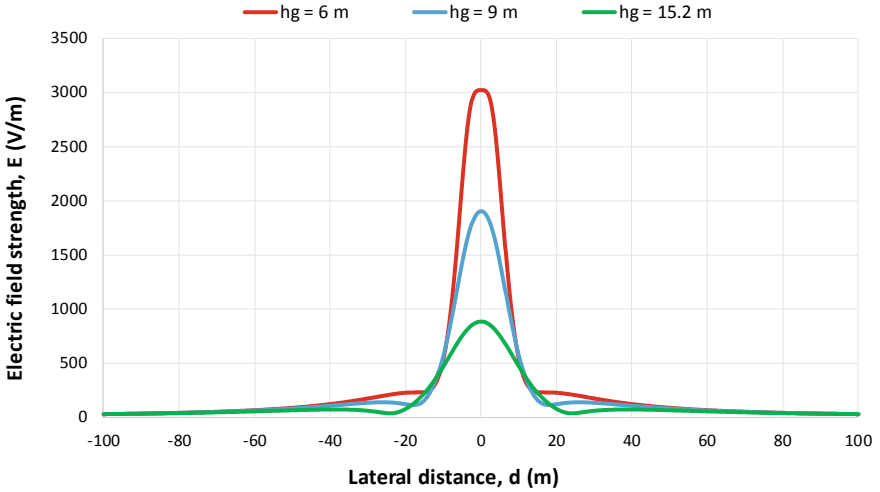


Fig. 18 RMS electric field strength profiles for the 110 kV OHVPL, for untransposed phasing and various ground clearances

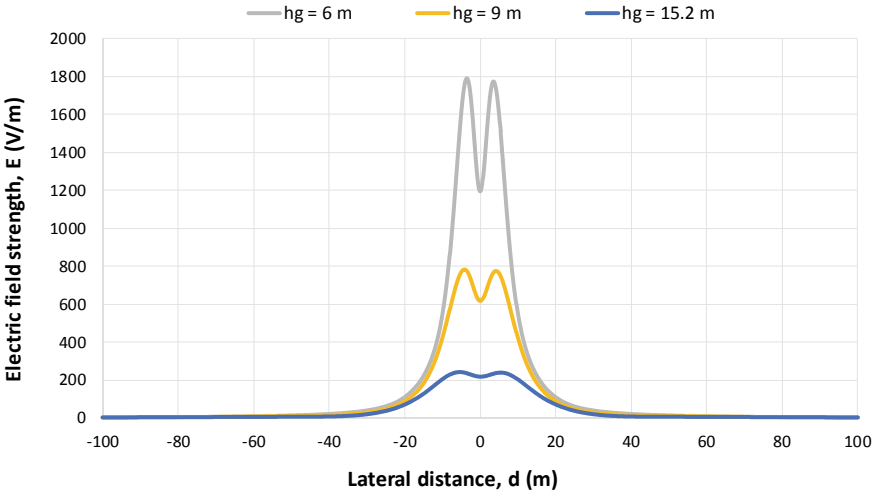


Fig. 19 RMS electric field strength profiles for the 110 kV OHVPL, for transposed phasing and various ground clearances

Table 3 Typical electric field levels from the 110 kV OHVPL ($h_g = 9$ m)

Phasing	E [V/m], at various lateral distances				
	Maximum beneath line	*18.5 m*	*25 m*	*50 m*	*100 m*
U	1901.8	110.3	134.4	77	25
T	784.4	109.5	47.9	11.3	3.3

3.2 ELF-EMFs from the 220 kV Double-Circuit OHVPL

The 220 kV double-circuit OHVPL selected for FE analysis has geometry dictated by suspension towers of Sn 220.202 type. The line is equipped with standard 450/75 mm^2 ACSR conductors (29.25 mm exterior diameter), for which the maximum allowable current is 975 A. Table 4 presents the geometrical data, as well as the voltage and current information used for computation, where the ground clearance h_g is taken 7 m, 11 m and 19 m, respectively. Because the geometry of this line is quite similar to the geometry of the 110 kV line, we expect similar electric and magnetic field distributions.

a. *Magnetic field distribution*

For untransposed phasing (Fig. 20), the maximum RMS magnetic flux density beneath the 220 kV OHVPL ranges from 7 μT at tower to 23.9 μT at mid-span, generally accounting for less than 24% of the ICNIRP exposure limit for general public. For transposed phasing (Fig. 21), the maximum RMS magnetic flux density along the half-span varies between 3.62 and 24.33 μT, hence not exceeding 24.4% of the ICNIRP exposure limit. And this time, much lower field levels can be observed at larger distances from the line for T phasing.

Table 5 gives magnetic field levels at various distances from the OHVPL centerline for the average clearance $h_g = 11$ m and a (more usual) loading of 200 A. Beneath the line, the magnetic flux density does not exceed 2.87 μT for U phasing and 2.35 μT for T phasing, while at the edge of the OHVPL corridor—27.5 m from the centerline [41]—it decreases to 0.93 μT and 0.42 μT, respectively. The critical value of 0.4 μT is reached at a lateral distance of about 45 m and 25.4 m, respectively. As we can see, because of the low load conditions, the typical exposure levels from this line are lower than those associated with the 110 kV OHVPL.

b. *Electric field distribution*

The electric field strength distribution for untransposed phasing is illustrated in Fig. 22. Beneath the 220 kV double-circuit OHVPL, the maximum field strength at the standard height $h = 1$ m ranges from 1452.8 V/m for $h_g = 19$ m to 4673.7 V/m for $h_g = 7$ m, which is very close to the ICNIRP limit for general public (93.5% of

Table 4 Input data for the 220 kV double-circuit OHVPL using towers of Sn 220.202 type (untransposed phasing)

Phase conductor	i	d_i [m]	hi [m]	$\underline{U}_i$ [kV]	$\underline{I}_i$ [A]
A	1	−5.00	h_g	$127 \angle 0°$	$975 \angle 0°$
B	2	−8.00	$h_g + 6.5$	$127 \angle -120°$	$975 \angle -120°$
C	3	−5.00	$h_g + 13$	$127 \angle 120°$	$975 \angle 120°$
A'	4	5.00	h_g	$127 \angle 0°$	$975 \angle 0°$
B'	5	8.00	$h_g + 6.5$	$127 \angle -120°$	$975 \angle -120°$
C'	6	5.00	$h_g + 13$	$127 \angle 120°$	$975 \angle 120°$

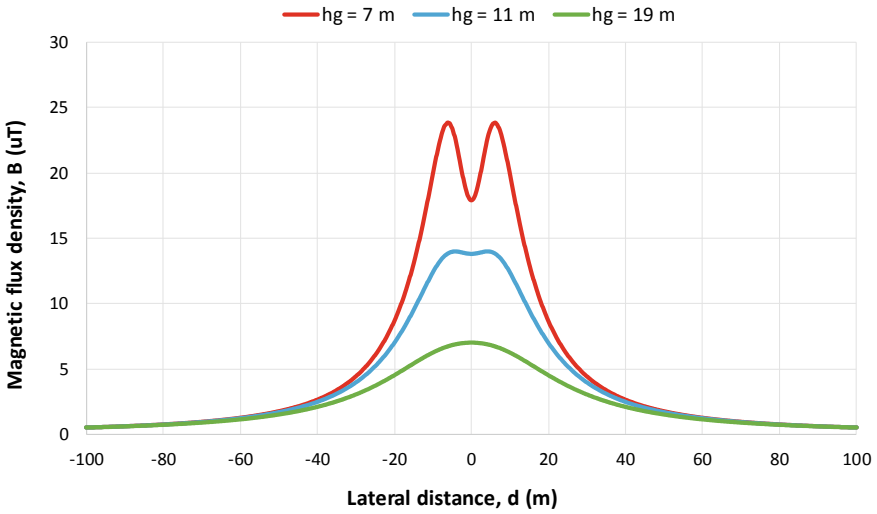


Fig. 20 RMS mag-netic flux density profiles for the 220 kV OHVPL, for untransposed phasing and various ground clearances

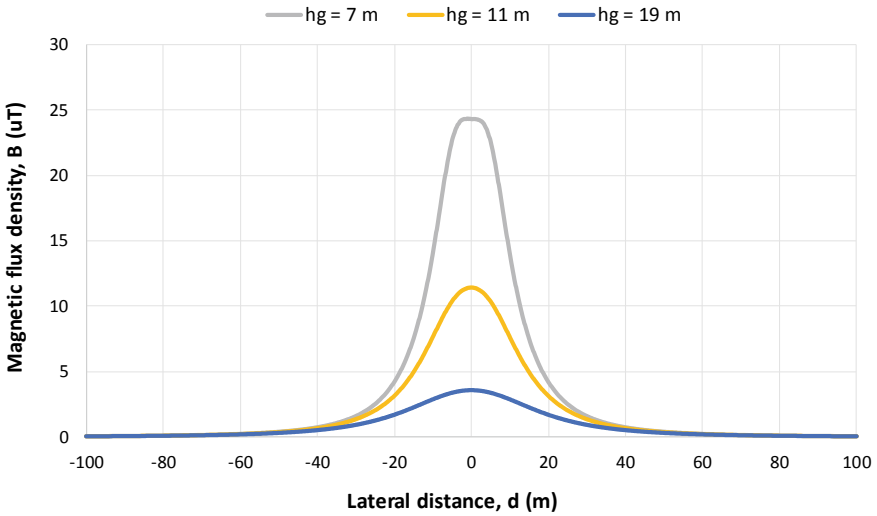


Fig. 21 RMS mag-netic flux density profiles for the 220 kV OHVPL, for transposed phasing and various ground clearances

Table 5 Typical magnetic field levels from the 220 kV OHVPL ($h_g = 11$ m, $I = 200$ A)

Phasing	B [μT], at various lateral distances				
	Maximum beneath line	*25 m*	*27.5 m*	*50 m*	*100 m*
U	2.87	1.07	0.93	0.35	0.11
T	2.35	0.42	0.34	0.076	0.01

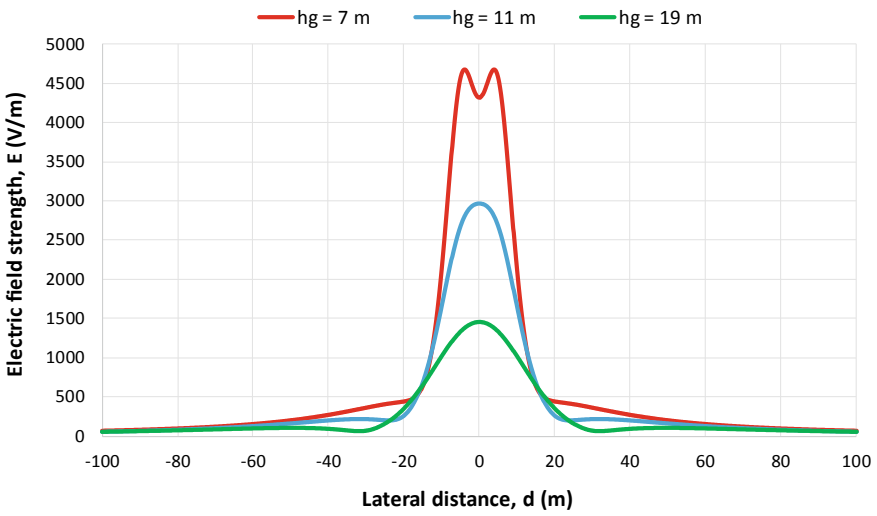


Fig. 22 RMS electric field strength profiles for the 220 kV OHVPL, for untransposed phasing and various ground clearances

the limit). For transposed phasing (Fig. 23), the maximum electric field strength at the same height lies in the range from 499.2 to 3653.9 V/m, which is below 73.1% of the ICNIRP exposure limit.

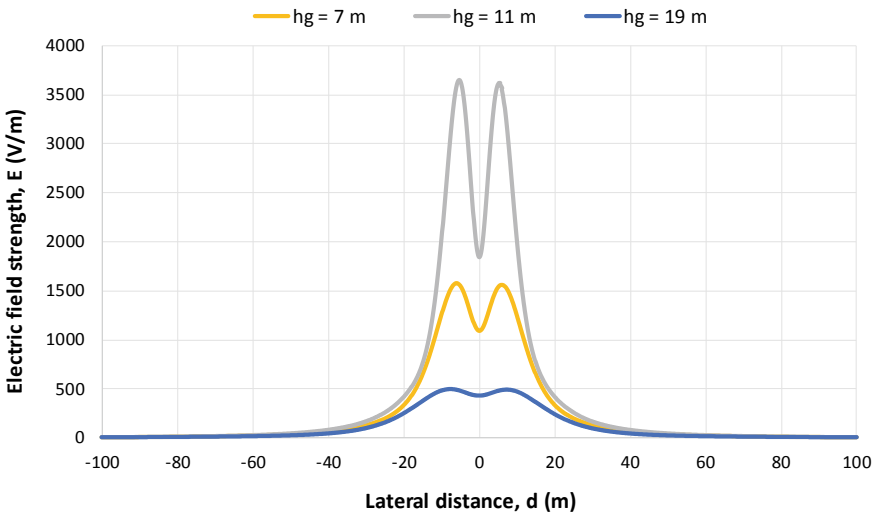


Fig. 23 RMS electric field strength profiles for the 220 kV OHVPL, for transposed phasing and various ground clearances

Table 6 Typical electric field levels from the 220 kV OHVPL ($h_g = 11$ m)

Phasing	E [V/m], at various lateral distances				
	Maximum beneath line	*25 m*	*27.5 m*	*50 m*	*100 m*
U	2969.5	207.6	215.9	161.7	60.5
T	1580	199	160	40	7.5

For the average clearance $h_g = 11$ m (Table 6), the electric field strength beneath the line reaches 2969.5 V/m for U phasing and 1580 V/m for T phasing, while the electric field strength at the corridor edge (27.5 m from the centerline) only reaches 215.9 V/m and 160 V/m, respectively. At 50 m from the OHVPL centerline, the electric field strength diminishes to 167.1 V/m and 40 V/m, respectively. As in the case of the 110 kV single-circuit OHVPL, such levels can also be measured at a distance of 30 cm from household appliances.

3.3 ELF-EMFs from the 400 kV Single-Circuit OHVPL

The last OHVPL selected for FE analysis is a 400 kV single-circuit line with geometry dictated by anchor portal towers of PAS 400.102 type. The line is equipped with two standard 450/75 mm^2 ACSR conductors per phase, with a distance between individual conductors of 0.4 m. The input data for this line are presented in Table 7, where the ground clearance h_g is taken 8.2 m, 12.6 m and 21.4 m, respectively. Magnetic field computations assume a maximum current of 1950 A.

a. *Magnetic field distribution*

The RMS magnetic flux density distribution at the height $h = 1$ m is illustrated in Fig. 24. As with the other investigated OHVPLs, the magnetic field beneath the 400 kV line does not exceed the ICNIRP exposure limit for general public, but at mid-span it can be as high as 57.4 μT, which represents more than half of this limit. Towards the tower, it falls to only 14.85% of the limit.

Table 7 Input data for the 400 kV single-circuit OHVPL using towers of PAS 400.102 type

Phase conductor	i	d_i [m]	hi [m]	$\underline{U_i}$ [kV]	$\underline{I_i}$ [A]
A	1	−11.50	h_g	$231 \angle 0°$	$975 \angle 0°$
	2	−11.10	h_g	$231 \angle 0°$	$975 \angle 0°$
B	3	−0.20	h_g	$231 \angle -120°$	$975 \angle -120°$
	4	0.20	h_g	$231 \angle -120°$	$975 \angle -120°$
C	5	11.10	h_g	$231 \angle 120°$	$975 \angle 120°$
	6	11.50	h_g	$231 \angle 120°$	$975 \angle 120°$

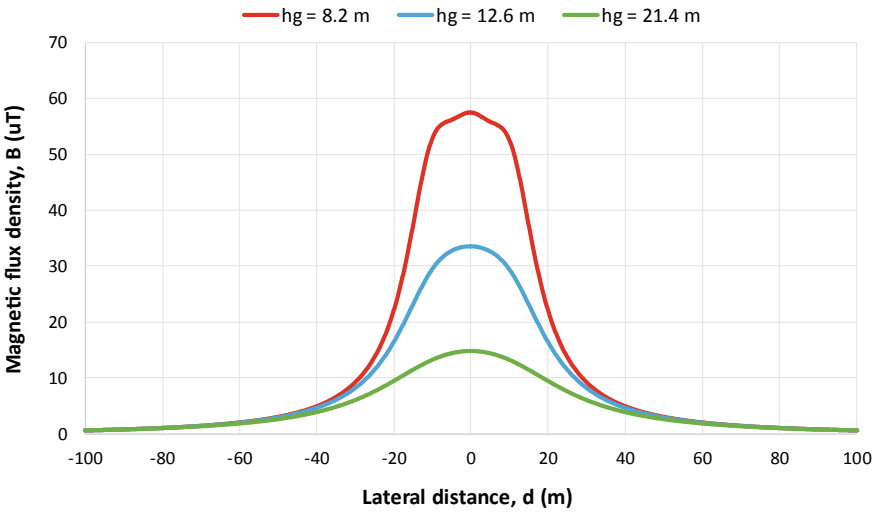


Fig. 24 RMS mag-netic flux density profiles for the 400 kV OHVPL, for various ground clearances

Table 8 Typical magnetic field levels from the 400 kV OHVPL ($h_g = 12.6$ m, $I = 450$ A)

B [μT], at various lateral distances				
Maximum beneath line	*25 m*	*37.5 m*	*50 m*	*100 m*
7.73	2.64	1.22	0.69	0.16

Once again, Table 8 gives magnetic field levels at various distances from the OHVPL centerline for the average clearance $h_g = 12.6$ m and a (normal) loading of 450 A. As it can easily be observed, the magnetic flux density at the OHVPL centerline is 7.73 μT, while at the edge of the OHVPL corridor—37.5 m from the centerline [41] —it decreases to only 1.22 μT. The critical value of 0.4 μT is reached at a lateral distance of about 65.1 m.

b. *Electric field distribution*

The RMS electric field strength distribution at the height $h = 1$ m is illustrated in Fig. 25. This time, the electric field at mid-span ($h_g = 8.2$ m) is about two times higher than the ICNIRP exposure limit for general public, namely 9145.3 V/m. Towards the tower ($h_g = 21.4$ m), the electric field strength falls to 1970.9 V/m (39.41% of the limit), but, as the ground clearance increases, the maximum field levels slightly move outside the line (for $h_g = 21.4$ m, the maximum field strength is recorded at a distance of 16 m from the OHVPL centerline).

Finally, Table 9 gives electric field levels for the average clearance $h_g = 12.6$ m. The electric field strength beneath the line reaches 4752 V/m, decreasing to 776.4 V/m at the corridor edge (37.5 m from the OHVPL centerline) and to 340.2 V/m at

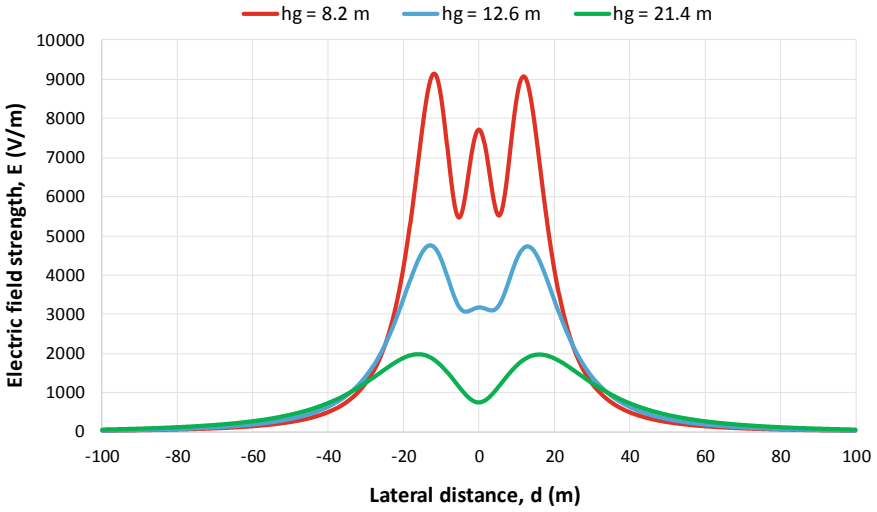


Fig. 25 RMS electric field strength profiles for the 400 kV OHVPL, for various ground clearances

Table 9 Typical electric field levels from the 400 kV OHVPL ($h_g = 12.6$ m)

E [V/m], at various lateral distances				
Maximum beneath line	*25 m*	*37.5 m*	*50 m*	*100 m*
4752	2212.7	776.4	340.2	44.1

50 m from the OHVPL centerline. At 100 m lateral distance, the electric field drops drastically, to only 44.1 V/m.

c. *Alternative computation approach*

All computations performed above assume that each sub-conductor of the considered 400 kV single-circuit OHVPL is modeled separately. However, this model can be simplified by replacing each of the three bundled conductors with an equivalent conductor of radius R_{eq}, given by [42]:

$$R_{eq} = \sqrt[N]{R \cdot d^{(N-1)}}, \tag{3}$$

where N stands for the number of sub-conductors in bundle, R represents the radius of a sub-conductor and d is the separation distance between sub-conductors. Equation (3) is applicable for up to three conductors per bundle, in our case leading to $R_{eq} = 76.485$ mm.

Figure 26 compares RMS magnetic flux density profiles computed with ANSYS Maxwell 2D by both approaches, as well as with PowerMAG software, which makes

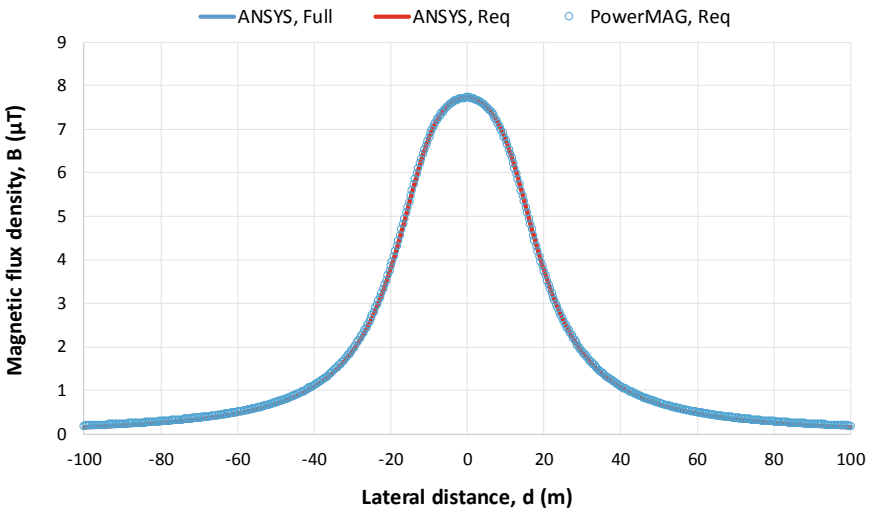


Fig. 26 Comparison between RMS magnetic flux density profiles obtained by the two approaches with ANSYS Maxwell 2D and PowerMAG software ($h_g = 12.6$ m, $I = 450$ A)

use of equivalent conductor model. As it can be observed, there is an excellent agreement between the three magnetic field profiles.

Similarly, Fig. 27 compares RMS electric field strength profiles computed with ANSYS Maxwell 2D by both approaches, as well as with PowerELT software, which

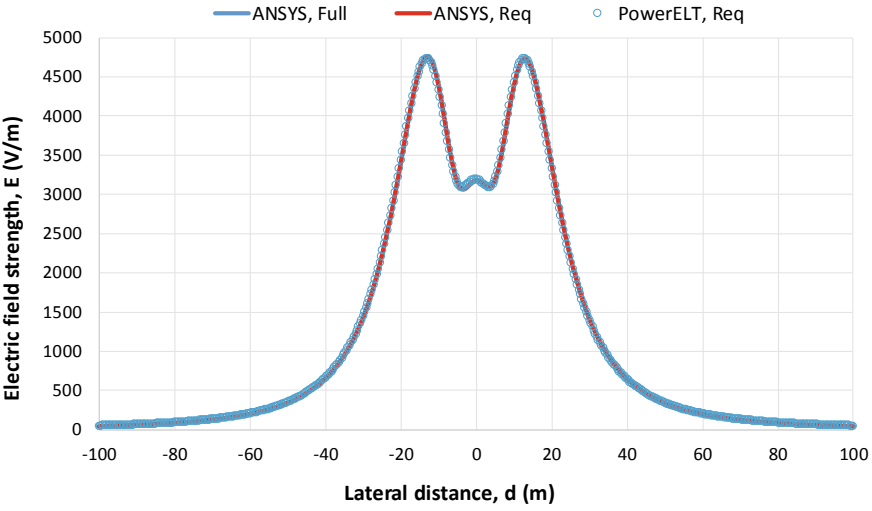


Fig. 27 Comparison between RMS electric field strength profiles obtained by the two approaches with ANSYS Maxwell 2D and PowerELT software ($h_g = 12.6$ m)

also makes use of equivalent conductor model. And this time, an excellent agreement between the three electric field profiles can be observed.

4 Conclusions

This chapter has been devoted to computing and analyzing ELF electric and magnetic fields emitted by typical configurations of OHVPLs used in Romania. All computations have been conducted using a 2D ANSYS Maxwell finite element model, strictly verified by quasi-static analytical methods. According to the obtained results, the highest exposure levels to ELF electric fields are associated with the 400 kV OHVPLs, directly beneath the line approaching the double of the ICNIRP limit for general public. As for ELF magnetic fields, the highest exposure levels are also associated with the 400 kV OHVPLs, but they are approaching only 60% of the ICNIRP limit. At the edge of the line corridor, the typical ELF electric and magnetic fields originating on the investigated OHVPLs are well below the specified limits, regardless the nominal voltage of the line. The computed ELF-EMF exposure levels are in line with those determined for similar OHVPLs in other countries.

References

1. Ahlbom A, Day N, Feychting M et al (2000) A pooled analysis of magnetic fields and childhood leukaemia. Br J Cancer 83:692–698
2. Greenland S, Sheppard AR, Kaune WT et al (2000) A pooled analysis of magnetic fields, wire codes, and childhood leukemia. Child Leuk-EMF Study Group Epidemiol 11:624–634
3. IARC Working Group (2002) IARC monographs on the evaluation of carcinogenic risks to humans series. Non-ionizing radiation, Part 1: static and extremely low-frequency (ELF) electric and magnetic fields, vol. 80. IARC Press, Lyon. https://monographs.iarc.fr/ENG/Monographs/vol80/mono80.pdf. Last accessed 28 Sept 2019
4. ICNIRP (1998) Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic fields (up to 300 GHz). Health Phys 74:494–522
5. IEEE (2002) C95.6-2002 IEEE standard for safety levels with respect to human exposure to electromagnetic fields 0 to 3 kHz, New York
6. Porsius JT, Claassen L, Smid T, et al (2014) Health responses to a new high-voltage power line route: design of a quasi-experimental prospective field study in the Netherlands. BMC Public Health 14(237):1–12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3975333/#B8. Last accessed 28 Sept 2019
7. Porsius JT, Claassen L, Smid T et al (2015) Symptom reporting after the introduction of a new high-voltage power line: a prospective field study. Environ Res 138:112–117
8. Sahbudin RKZ, Fauzi SA, Hitam S et al (2010) Investigation of electrical potential and electromagnetic field for overhead high voltage power lines in Malaysia. J Appl Sci 10(22):2862–2868
9. Elhabashi SM, Ehtaiba JE (2007) Electric fields intensity around the new 400 kV power transmission lines in libya. In: Proceedings of the 6th WSEAS international conference on circuits, systems, electronics, control and signal processing, Cairo, 29–31 Dec 2007, pp 390–398
10. EMFs.info (2019) Factors affecting the field from a power line. https://www.emfs.info/sources/overhead/factors/. Last accessed 28 Sept 2019

11. Vornicu S, Lunca E, Salceanu A (2019) ANSYS maxwell finite element model for 2D computation of the magnetic field generated by overhead high-voltage power lines. In: Proceedings of the 12th international conference and exhibition on electromechanical and energy systems (SIELMEN 2019), Chisinau, 10–11 Oct 2019, pp 382–385
12. Fuchs E, Masoum MAS (2008) Power quality in power systems and electrical machines. Elsevier Academic Press, Burlington
13. Olsen RG, Deno D, Baishiki RS et al (1988) Magnetic fields from electric power lines: theory and comparison to measurements. IEEE Trans Power Deliv 3(4):2127–2136
14. Vujević S, Lovrić T, Modrić T (2011) 2D computation and measurement of electric and magnetic fields of overhead electric power lines. In: Proceedings of the joint 3rd international workshop on nonlinear dynamics and synchronization (INDS'11) & 16th international symposium on theoretical electrical engineering (ISTET'11), Klagenfurt, 25–27 July 2011, pp 1–6
15. Kokoruš M, et al (2014) Analysis of the possible solutions for the reduction of electric and magnetic fields near 400 kV overhead transmission lines. In: Passerini G, Brebbia CA (eds) Environmental impact II. 2nd international conference on environmental and economic impact on sustainable development, Ancona, May 2014. WIT transactions on ecology and the environment, vol. 181. WIT Press, Southampton, p 225
16. Ellithy K, Al-Suwaidi S, Elsayed H (2011) Measuring human exposure to magnetic fields near EHV 400 kV GIS substation and power lines in state of qatar. In: Proceedings of the north American power symposium (NAPS 2011). Boston, 4–6 Aug 2011, pp 1–6
17. Medved D, Mišenčík L, Kolcun M, et al (2015) Measuring of magnetic field around power lines. In: Proceedings of the 8th international scientific symposium ELEKTROENERGETIKA 2015. Stará Lesná, 16–18 Sept 2015, pp 148–151
18. Vergara XP, Kavet R, Crespi CM et al (2015) Estimating magnetic fields of homes near transmission lines in the California power line study. Environ Res 140:514–523
19. Tourab W, Babouri A (2016) Measurement and modeling of personal exposure to the electric and magnetic fields in the vicinity of high voltage power lines. Saf Health Work 7(2):102–110
20. Filippopoulos G, Tsanakas DK (2005) Analytical calculation of the magnetic field produced by electric power lines. IEEE Trans Power Deliv 20(2):1474–1482
21. Moro F, Turri R (2008) Fast analytical computation of power-line magnetic fields by complex vector method. IEEE Trans Power Deliv 23(2):1042–1048
22. Ztoupis IN, Gonos IF, Stathopulos IA (2013) Calculation of power frequency fields from high voltage overhead lines in residential areas. In: Proceedings of the 18th international symposium on high voltage engineering. Seoul, 25–30 Aug 2013, pp 61–66
23. Milutinov M, Juhas A, Prša M (2007) Electric field strength and polarization of multi three-phase power lines. In: Proceedings of the 8th international conference on applied electromagnetics. Niš, 3–5 Sept 2007, pp 1–4
24. Milutinov M, Juhas A, Prša M (2009) Electromagnetic field underneath overhead high voltage power line. In: Proceedings of the 4th international conference on engineering technologies. Novi Sad, 28–30 April 2009, pp 1–5
25. Al Salameh MSH, Hassouna MAS (2010) Arranging overhead power line conductors using swarm intelligence technique to minimize electromagnetic fields. Prog Electromagn Res B 26:213–236
26. Gouda OE, Amer GM, Salem WA (2009) Computational aspects of electromagnetic fields near H.V. Transmission lines. Energy Power Eng 1(2):65–71
27. Lunca E, Istrate M, Salceanu A et al (2012) Computation of the magnetic field exposure from 110 kV overhead power lines. In: Proceedings of the 7th international conference on electrical and power engineering. Iasi, 25–27 Oct 2012, pp 628–631
28. Lunca E, Ursache S, Salceanu A (2018) Computation and analysis of the extremely low frequency electric and magnetic fields generated by two designs of 400 kV overhead transmission lines. Measurement 124:197–204
29. Razavipour SS, Jahangiri M, Sadeghipoor H (2012) Electrical field around the overhead transmission lines. World Acad Sci Eng Technol 6(2):168–171

30. Lunca E, Istrate M, Salceanu A (2013) Comparative analysis of the extremely low-frequency magnetic field exposure from overhead power lines. Environ Eng Manag J 12(6):1145–1152
31. Ali Rachedi B, Babouri A, Berrouk F (2014) A study of electromagnetic field generated by high voltage lines using COMSOL MULTIPHYSICS. In: Proceedings of the 2014 international conference on electrical sciences and technologies in Maghreb. Tunis, 3–6 Nov 2014, pp 1–5
32. Salceanu A, Lunca E, Paulet M (2017) Affordable evaluation of low frequency electric fields from the standpoint of directive 2013/35/EU. ACTA IMEKO 6(4):37–45
33. Lunca E, Vornicu S, Salceanu A et al (2018) 2D Finite element model for computing the electric field strength-rms generated by overhead power lines. J Phys Conf Ser 1065:1–4
34. EU (1999) 1999/519/EC: Council recommendation of 12 July 1999 on the limitation of exposure of the general public to electromagnetic fields (0 Hz to 300 GHz). https://publications.europa.eu/en/publication-detail/-/publication/9509b04f-1df0-4221-bfa2-c7af77975556/language-en. Last accessed 28 Sept 2019
35. Braicu SF, et al (2017) Evaluation of the electric and magnetic field near high voltage power lines. In: Vlad S, Roman NM (eds) MEDITECH 2016. 5th international conference on advancements of medicine and health care through technology. Cluj-Napoca, 2016. IFMBE proceedings, vol. 59, Springer International Publishing, p 141
36. Ghani SA, Ahmad Khiar MS, Chairul IS, et al (2014) Study of magnetic fields produced by transmission line tower using finite element method (FEM). In: Proceedings of the 2nd international conference on technology, informatics, management, engineering & environment. Bandung, 19–21 Aug 2014, pp 64–68
37. EMFs.info (2019) Phasing—how the phasing affects the field produced by an overhead line. https://www.emfs.info/sources/overhead/factors/phasing/. Last accessed 28 Sept 2019
38. Ponnle AA, Adedeji KB, Abe BT et al (2017) Variation in phase shift of phase arrangements on magnetic field underneath overhead double-circuit HVTLs: field distribution and polarization study. Prog Electromagn Res M 56:157–167
39. Lunca E, Ursache S, Salceanu A (2017) Characterization of the electric and magnetic field exposure from a 400 kV overhead power transmission line in Romania. In: Proceedings of the 22nd IMEKO TC4 international symposium and 20th international workshop on ADC modelling and testing. Iasi, 14–15 Sept 2017, pp 239–243
40. Transelectrica (2019) Power transmission grid. https://www.transelectrica.ro/en/web/tel/date-generale-management. Last accessed 28 Sept 2019
41. ANRE (2007) Technical norm regarding the delimitation of the protection and safety zones associated with energy capacities—Revision I of 09 March 2007
42. Grigsby LL (2012) Electric power generation, transmission, and distribution, 3rd edn. CRC Press, Boca Raton

Design and Finite Element Analysis of Permanent Magnet Synchronous Generator for Wind Turbine Application

Abdurrahman Yavuzdeger, Burak Esenboga, Firat Ekinci, and Tugce Demirdelen

Abstract Today, the demand for renewable energy sources is increasing day by day in order to reduce fossil fuels and meet the increasing energy demand. The fact that wind energy is suitable for energy production at continuous or low wind speed depending on geographical conditions increases its importance among eco-friendly energy sources. However, energy efficiency is one of the most important issues in the renewable energy field because energy production from these energy sources is constantly changing due to climate changes. Therefore, it is very important to use renewable energy sources efficiently and to enable innovative developments that will increase energy efficiency. In this chapter, a more efficient wind turbine alternator is modeled and analyzed in detail by using the ANSYS/Maxwell software program. The main objective of this chapter is to create an efficient alternator model used in both vertical and horizontal wind turbines. This alternator model is selected as a permanent magnet synchronous generator (PMSG) since there is no need for external excitation, smaller in size and easy to control. Firstly, the parameters are determined by using the mathematical model of the alternator. Secondly, the alternator is modeled and designed with the help of the design parameters such as pole pair, magnetizing inductance, the stator leakage, winding properties, number of turns and slots, etc. During the design process, all materials of the alternator are designed by taking into consideration of characteristic features of them. Finally, the designed alternator

A. Yavuzdeger (✉) · F. Ekinci
Energy Systems Engineering Department, Adana Alparslan Turkes Science and Technology University, Adana, Turkey
e-mail: ayavuzdeger@atu.edu.tr

F. Ekinci
e-mail: fekinci@atu.edu.tr

B. Esenboga · T. Demirdelen
Electrical and Electronic Engineering Department, Adana Alparslan Turkes Science and Technology University, Adana, Turkey
e-mail: besenboga@atu.edu.tr

T. Demirdelen
e-mail: tdemirdelen@atu.edu.tr

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_30

is electromagnetically analyzed thanks to ANSYS/Maxwell Electromagnetic Suit program which uses the Finite Element Method (FEM). Therefore, the electrical efficiency of the wind turbine alternator at different wind speeds is performed and the optimum design of the alternator is obtained. It is hoped that this study will guide for wind power plant operators and researchers interested in wind turbine design parameters.

Keywords Permanent magnet synchronous generator · Wind turbine · ANSYS/Maxwell · Finite element analysis

Nomenclature

A. Acronyms

PMSG	Permanent Magnet Synchronous Generator
FEM	Finite Element Method
FEA	Finite Element Analysis
DC	Direct Current
AC	Alternative Current
MEC	Magnetic equivalent circuit
ORPMSG	Outer Rotor Permanent Magnet Synchronous Generator
CFD	Computational Fluid Dynamics
GSA	Gravitational Search Algorithm
PSO	Particle Swarm Optimization
ANN	Artificial Neural Network
SQP	Sequential Quadratic Programming
NdFeB	Neodymium Iron Boron
RMS	Root Mean Square

B. Symbols/Parameters

N	Rated Speed
f	Frequency
rpm	Revolution Per Minute
p	Number of Poles
dq	Direct quadrature
V_{sd}	D-axis Stator Voltage
i_{sd}	D-axis Stator Current
φ_{sd}	D-axis Stator Flux
L_d	D-axis Inductance
V_{sq}	Q-axis Stator Voltage
i_{sq}	Q-axis Stator Current
φ_{sq}	Q-axis Stator Flux
L_q	Q-axis Inductance

R_s	Stator Winding Resistance
w_e	Angular Speed
ϕ_g	Magnetic Flux
D_{si}	Stator Inner Diameter
L_s	Length of the Stator

1 Introduction

Renewable energy is obtained from sources that can be reached from nature continuously or repeatedly. This type of energy, which can be easily produced from domestic sources, is seen as a solution that meets the requirements of the modern age. These resources such as solar, wind and wave energy are the natural resources that exist in nature and do not run out of time, so they are presented as important alternative energy sources to fossil fuels. Natural resources do not require many years to form. They also prevent environmental pollution. Moreover, they are preferred due to their sustainability characteristics compared to fossil fuels. Due to the increasing number of vehicles, the heating needs of the rapidly growing population, the growing volume of industrialization, and the increasing need for energy for many reasons, fossil fuels show signs of exhaustion earlier than expected. In addition, fuels such as coal, lignite and gasoline cause bad results environmentally when innovative technologies are not used. All these reasons show that alternative energy production is critical for the survival of people and nature. In order to prevent environmental pollution and fuel crisis, alternative energy sources come to the forefront for economic and continuous production. Considering the geographical locations of the countries, it is foreseen that efficient studies can be carried out in wind, solar, biomass and geothermal fields. Therefore, the use of renewable energy sources is increasing all over the world day by day.

From eco-friendly energy production applications, wind energy, which is free and abundant in nature, is the most preferred energy source among alternative energy sources. This energy source is a sustainable energy source where we can obtain highly efficient electrical energy without being dependent on external sources and environmentally friendly [1]. Wind energy, which is used in many fields, is mostly used in electricity generation today. Wind turbines are used to convert energy from wind energy to electrical energy. Wind turbines generally consist of propellers, tower, generator, gearbox and electric-electronic components. The wind energy is converted from the kinetic energy to mechanical energy in the rotating units of the machines by means of propellers. Then the speed of the movement is accelerated and this energy is transferred to the generators. The electricity generation is provided by the generators of the wind turbines.

In recent years, technological developments in wind energy systems and government supports are rapidly increasing the share of wind energy in the power system. This rapid growth in wind power has brought many studies and research topics

related to the inclusion of wind energy to the power system. Generators, which are key components of wind turbines, are selected from various generators with high efficiency to convert mechanical energy to electrical energy. These generators are DC, synchronous and asynchronous generators. DC generators are used in small power systems in the past. They have replaced by synchronous or asynchronous generators. These generators work together with power electronics elements that can easily convert direct current to alternating current with the help of converters. In medium and large power systems, synchronous and asynchronous generators are generally preferred. In this study, PMSG which is widely used in wind turbines is discussed. PMSG is recommended for wind turbine applications because it is self-excited without the need for any energy source. The most important advantage is that it can generate power at any speed. These generators have low maintenance costs. They are suitable for small and light applications. The generator speed can be controlled without the need for a gearbox. The stator of the PMSG is winded and permanent magnets are placed in the rotor. Considering these advantages of PMSG, this type of generator is widely used in power systems that produce electrical energy from wind power. It is one of the most important issues that generators, which are the main components of wind turbines, be designed with high efficiency in the renewable energy field. In a synchronous generator design, the proper arrangement of design parameters plays an important role in the generator's electrical efficiency. The generator efficiency is significantly affected by changing the vital parameters in the design. Therefore, in this study, the design and electromagnetic analysis of a PMSG are realized by using ANSYS/Maxwell program. The effect on the efficiency of the generator is examined by changing the basic parameters (such as rotor air gap, magnet type, stator and steel sheet type used in the rotor) of the synchronous generator. An efficient generator model for use in wind turbines is designed and analysed.

The major goal of this chapter is to present the novel design of a PMSG with maximum efficiency for a wind turbine application. Therefore, the parameters affecting the efficiency of the PMSG are analysed in detail. In the following sections, a comprehensive literature study on the PMSG is carried out. Topics focused on design the PMSG have been examined in the literature. The design and electromagnetic analysis of PMSG are carried out by using ANSYS/Maxwell program.

2 A Literature Review of PMSGs

In the design of the wind turbine, all the stages from wind wings to the transmission of the resulting power to the electrical grid are important. One of these steps and the main component of wind turbines are the generators, which contain many parameters and design knowledge necessary for high-efficiency design. Many studies have been done in the field of PMSG and some of these studies are as follows;

Widyan carried out the analysis, efficiency optimization, manufacturing and testing of radial flux low-speed high-efficiency PMSG. He designed a new rotor

configuration to reduce the cogging moment component due to slot and magnet interaction. In order to reduce the cogging moment present, special magnetic material was placed in the slot opening and the maximum value of the starting torque is decreased to 43% of the nominal torque [2]. Janne Nerg et al. presents a model showing the heat flow of a 3.35 MW PMSG. In this model, the main parameters of the generator were given and electromagnetic calculations were made by using a two-dimensional finite element method. Using computational fluid dynamics, performance analysis and heat flow calculations were performed. The results were confirmed by practical measurements [3]. In another study, the authors proposed a method consisting of three stages while designing the PMS generator with 20 kW power. Using magnets made of NdFeB materials, they have increased the permanent magnetization level up to 1.2 T [4]. The authors examined basic configuration such as material types, magnet types, electrical and mechanical parameters, and were specifically prepared to reduce the cogging effect of PMSG and improved slot shape efficiency. Different wire diameters were used to decrease copper losses and improved the thermal condition of the slot. Real-time performance simulation of PMSG was carried out by using finite element analysis (FEA) method. This PMSG provides minimal maintenance and easy installation and cooling benefits at low power wind turbine applications [5]. Osman et al. designed a prototype permanent magnet synchronous generator. There are four alternators on the same shaft in this work. It also consists of a control system that activates the stators at different wind speeds. In this way, both high speeds and low speeds can be adjusted to operate at high efficiency [6]. Arafat M. Y. et al. compared the inner rotor and outer rotor types in PMSG design. The outer rotor has been observed to be more efficient than the inner rotor, and these two generators have the same output power and speed [7]. Bouaziz O. et al. compared the performance of axial flux and radial flux PMSGs using ANSYS program and also, their estimated cost analyses for small size 550-W wind turbines. According to observed simulation results, it was presented that the axial flux PMSG was more efficient than the radial flux machine for microscale wind applications [8]. T. P. M. Bazzoa et al. proposed a multi-disciplinary permanent magnet synchronous generator design. This design includes six sub-models with fewer assumptions, hundreds of input and output variables. Therefore, the optimization results are precise and the accuracy is very high. To confirm the accuracy of these results, finite element simulations were performed. In addition, a Pareto was presented which shows the balance between generator cost and wind turbine energy production so that the manufacturer can select the most suitable one from the best possible solutions [9]. Santolo et al. used a hybrid system when designing a PMSG. This generator had a direct drive and flux switching design. Permanent magnet width, rotor tooth width and number, stator pole number were determined and cost and total harmonic distortion were reduced by using four-stage artificial neural network and multipurpose particle swing optimization in the design. A product suitable for low-power wind applications has been proposed [10]. The authors used a different design procedure when designing PMSG. Initially, the analytical method based on magnetic equivalent circuits (MEC) was used when designing physical and electromagnetic, thus providing a fast design. In later stages, the final results were obtained using the finite element method and RMXPRT. Air gap flux

density reached 0.8 T and increased by 5% in efficiency [11]. Erol et al. designed a new permanent magnet generator. This work had innovative flux topology and good sinusoidal voltage wave, magnetic flux, and current waveform. The generator was designed with three-dimensional FEA method with magnetostatic and magneto dynamic tools as rotor at two ends and stator in the middle. Preliminary studies and results this innovative approach is promising for the future [12]. Fang H. et al. put forward design procedure studies without increasing any material costs in order to increase the output power of permanent magnet synchronous alternator. The effect of the change of the mechanical pole arc coefficient on the output power of the PMSG was observed with simulation results [13]. The inner rotor and outer rotor permanent magnet design were made comparatively to GSA and GSA-PSO algorithms to optimize the design. GSA-PSO was found to be more efficient than GSA. The Outer Rotor PMSG (ORPMSG) weight was reduced by 32% and rotor weight by 13%, resulting in a design with fewer wind requirements. The overall productivity of ORPMSG increased by 2.19%. ORPMSG has been shown to be more advantageous in many areas [14]. In another study, Axial Flux Permanent Magnet Generator with single stator and double rotor was designed and its responses in constant speed wind and variable speed winds were presented whole system has been integrated and tested in MATLAB/Simulink environment. The calculated results and the results in simulation were given in table comparatively [15]. The axial flux coreless PMSG has two stators and three rotors. Permanent magnets were placed at the head and end of the middle rotor and have a speed of 500 rpm. Electromagnetic and structural analyses were performed by using three-dimensional finite element method [16]. In another study, Jin-Hyung Yoo et al. carry out a permanent magnet design for obtained low cogging torque. The edges of the permanent magnets were circularly optimized. Thus the cogging torque decreased by 40%, reducing the variation of the air gap reluctance [17].

The authors designed the modern generation double stator permanent magnet synchronous generator. Equations of DSPMSG were extracted, structural properties were explained and analyzed. The analyzed results were optimized by the analytical hierarchy process and simulated on the generator Maxwell based on the FEM. Furthermore, it was compared with a conventional permanent magnet generator and it was found that a motor with 45% more power was obtained low copper losses and a smaller structure [18]. Pedram et al. have designed two and four corrugated stator cores as natural cooling systems to create solutions to the temperature increase that occurs during the operation of the generators. The design also examined thermal analysis based on a three-dimensional FEA, and this study compared the normal stator core and the innovative cooling-channel stator core. The temperature had been measured 11 (°C) lower than the previous temperature [19]. Warda Gul et al. designed a 5 MW PMSG with a double stator and a single rotor. Their main goal in this design was to reduce the weight of large-scale generators. For this reason, optimization has been achieved by using a genetic algorithm and they have made necessary parameter calculations by using three-dimensional finite element analysis [20]. Sahib Khan et al. inner rotor and outer rotor PMSG have made comparative design. The generator parameter was calculated using D^2L method and analyzed by FEM. According to the

results, the outer rotor PMSG performs better thanks to the low flux connection and high output power [21]. Pablo Jaen-Sola et al. compared composite material and steel using ANSYS/Maxwell. The generator made of composite material has 72% lighter structure than steel. They also proposed mosaic structures except for traditional ones [22].

Following a literature survey, Table 1 summarizes the motor types, motor output power, the programs and methods used in the design and the applications.In this study, PMSG analysis and design, which is widely used in wind turbines and which is of high importance to design at high efficiency, has been carried out. Many parameters such as the material used in the stator and rotor, the shape of the slot type, slot-pole combination, magnet placement and winding types, which play a role in the efficiency of the generator, have been discussed. The design and electromagnetic analysis of a PMSG are realized by using ANSYS/Maxwell program. This program uses the Finite Element Method (FEM).

3 Design and Modelling of PMSG

There are vital critical parameters when designing an efficient generator for wind turbines. Determination of slot-pole combination, flux direction, rotor structure, magnet placement and winding types are some of these important parameters. Because the variations in this parameter affect the efficiency of the generators to be designed for wind turbines. The changes slot-pole combination which is one of the generator design parameters affect many output parameters, such as the generator momentary impact, efficiency, and the harmonic distribution in the induced voltage. Also, it is shown in Eq. 1 that the revolution of the generator is obtained by using the pole number of the machine and the system frequency.

$$N = \frac{120 \times f}{p} \tag{1}$$

where N is rated speed, f is the system frequency and p-value expresses the pole number of the generator. The number of poles is chosen as 50 so that the designed generator has an operating frequency of 50 Hz at 120 revolutions per minute (rpm). There are two types of rotors. Figure 1 presents these rotor types called the inner and outer rotors. When the literature is examined, internal rotor structures are generally used at permanent rotor synchronous generator design models. In this type, the rotor is placed inside the stator and the permanent magnets are located to the outer surface of the rotor. The inner rotor structure is not convenient for high-speed wind turbine practices. Because of the centrifugal force during rotation at high speed, the magnets are at risk of detaching from the rotor. However, because of having the low centrifugal force will not damage the magnets at the low-speed wind applications. In outer rotor structures, the stator is placed in the rotor and magnets are located on the inner part

Table 1 Various solar forecasting methods

Refs.	Motor type	Power	Analysis program or method	Application
[2]	PMSG	19.5 kW	Finite Element Analysis (FEA)	Cogging torque was reduced by placing soft magnetic material in slot openings
[3]	PMSG	3.35 MW	Computational Fluid Dynamics (CFD) and FEA	Heat flow calculations were performed
[4]	PMSG	20 kW	Artificial Neural Network (ANN) and 2D-FEA	Proposed three stages design method and increased the permanent magnetization level up to 1.2 T
[5]	Outer rotor PMG	400 W	ANSYS/Maxwell RMXPRT and Maxwell 2D	Cogging effect reduced and slot shape efficiency improved
[6]	Axial flux PMSG	250 W	ANSYS/Maxwell	Four alternators on the same shaft and a control system that activates the stators at different wind speeds
[7]	Outer and inner rotor radial flux PMG	400 W	ANSYS/Maxwell RMXPRT and Maxwell 2D	Comparison of the inner and outer rotor. The outer rotor was more efficient than the inner rotor. Cogging effect reduced
[8]	Axial Flux and Radial Flux PMSG	550 W	ANSYS/Maxwell	The comparative performance analysis of axial flux and radial flux PMSG. Axial flux PMSG was more efficient than radial flux PMSG

(continued)

Table 1 (continued)

Refs.	Motor type	Power	Analysis program or method	Application
[9]	PMSG	50 kW	Sequential Quadratic Programming (SQP) and finite element analysis	A multi-disciplinary PMSG design with optimization program and the balance between generator cost and wind turbine energy generation
[10]	PM flux switching generator	1.5 kW	Multi-objective particle swarm optimization and artificial neural network	Cost and total harmonic distortion were reduced by ANN and multipurpose particle swing optimization in the design
[11]	PMSG	100 W	ANSYS/Maxwell RMXPRT and magnetic equivalent circuits	Air gap flux density reached 0.8 T and increased by 5% in efficiency thanks to the design procedure followed
[12]	Axial flux PMSG	3.4 kW	Three-dimensional FEA	Innovative design as rotor at two ends and stator in the middle
[13]	PMSG	45 MW	ANSYS/Maxwell	The effect of the change of the mechanical pole arc coefficient on the output power of the PMSG without increasing any material costs
[14]	PMSG	500 KVA	MATLAB, GSA and PSO	The ORPMSG's weight was reduced by 32% and rotor weight by 13% with GSA and GSA-PSO
[15]	Axial flux PMSG	2 KVA	MATLAB/Simulink	Single stator and the double rotor was designed

(continued)

Table 1 (continued)

Refs.	Motor type	Power	Analysis program or method	Application
[16]	PMSG	4.5 kW	Three-dimensional FEA	PMSG has two stators and three rotors. Permanent magnets were placed at the head and end of the middle rotor
[17]	PMSG	120 kW	2D-FEA	The edges of the permanent magnets were circularly optimized. Thus the cogging torque decreased by 40%
[18]	PMSG	660 kW	ANSYS/Maxwell	Designed the modern generation double stator PMSG
[19]	Outer rotor PMSG	6 kW	Computational Fluid Dynamic (CFD) and 3D finite-element analysis	Designed two and four ducts stator cores as natural cooling systems
[20]	PMSG	5 MW	Genetic algorithm and FEA	Double stator and a single rotor generator designed and reduced the weight of large-scale generators
[21]	Outer and inner rotor axial flux PMG	2.5 kW	D^2L method and FEA	Comparison of the inner and outer rotor. The outer rotor is more efficient than the inner rotor
[22]	PMSG	100 kW and 3 MW	ANSYS/Maxwell	Compared to composite material and steel. The generator made of composite material has a 72% lighter structure than steel

of the rotor. This type of rotor structure allows more magnet placement in the rotor. The centrifugal force applied to the magnets by rotation of the outer rotor is in a direction to prevent the magnet from separating from the rotor. Therefore, the risk of magnets separating the rotor at high revolutions is eliminated. In this paper, the inner rotor type structure is used for the proposed PMSG model.

Biaxial (direct quadrature) equations are referred for the mathematical modeling of the proposed PMSG design [23]. The voltages, current, power torque, flux and

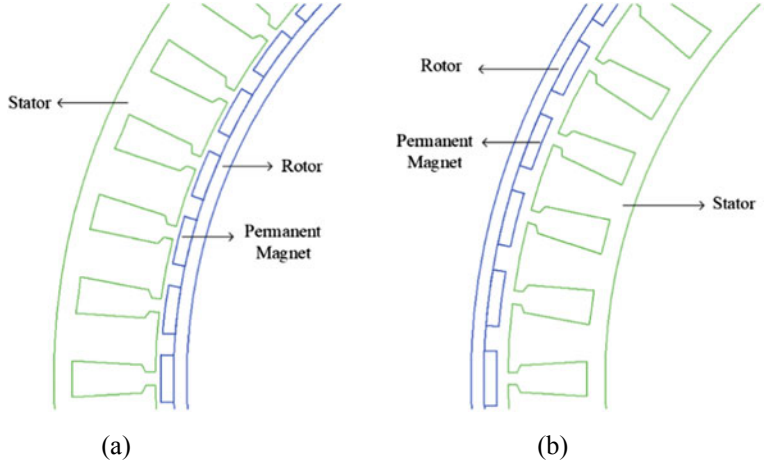


Fig. 1 **a** Inner rotor structure and **b** outer rotor structure

inductance expressions, which are shown in the following Eqs. 2–7, were converted into direct quadrature (dq) frames for analysis.

$$\varphi_{sd} = L_d \cdot i_{sd} + \varphi_m \tag{2}$$

$$\varphi_{sq} = L_q \cdot i_{sq} \tag{3}$$

$$V_{sd} = R_s \cdot i_{sd} + \frac{d\varphi_{sd}}{dt} - w_e \cdot \varphi_{sq} \tag{4}$$

$$V_{sq} = R_s \cdot i_{sq} + \frac{d\varphi_{sq}}{dt} - w_e \cdot \varphi_{sd} \tag{5}$$

$$P_e = \frac{3}{2}(V_{sd} \cdot i_{sd} + V_{sq} \cdot i_{sq}) \tag{6}$$

$$T_e = \frac{3}{2} p_p (i_q \cdot i_d (L_d - L_q) + \varphi_m \cdot i_q \tag{7}$$

In these equations, V_{sd} and V_{sq} denote dq axis stator voltage, i_{sd} and i_{sq} express the dq axis stator currents, φ_{sd} and φ_{sq} define the dq axis stator fluxes, L_d and L_q called as the dq axis inductances, R_s denotes the stator winding resistance and w_e states the alternator angular speed [24]. Also, the stator inner diameter and length of the stator are directly proportional to the magnetic flux as shown in Eq. 8.

$$\phi_g = B_{av} \times \left(\frac{\pi \times D_{si} \times L_s}{p} \right) \tag{8}$$

where ϕ_g is magnetic flux, D_{si} express the inner dimension of the stator and L_s denotes the stator length.

In the proposed permanent magnet synchronous generator, the number of poles and slots are determined 50 and 48 as shown in Fig. 2, respectively.

The winding type of the designed generator is determined as whole-coiled and 2 number of winding layers as shown in Fig. 3. Furthermore, the wire size was chosen 2.588 mm to obtain a higher stator fill factor ratio and more efficiency.

The most commonly used stator slot types in the permanent magnet synchronous generator design are as shown in Fig. 4. Each stator slot type has a different slot fill factor value. The high slot fill factor ratio allows the use of thicker conductors, which significantly reduces copper losses. The proposed stator slot type with a 75% stator fill factor ratio is as shown in Fig. 5. The dimensions of the proposed stator slot type

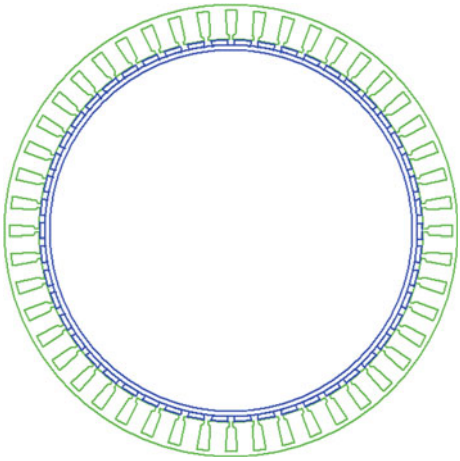


Fig. 2 The 48-slot and 50-pole generator structure

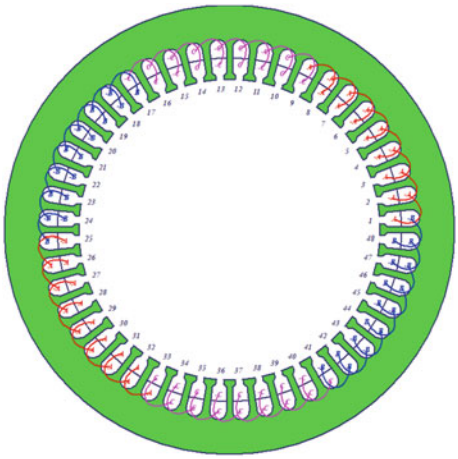


Fig. 3 Winding diagram of the PMSG

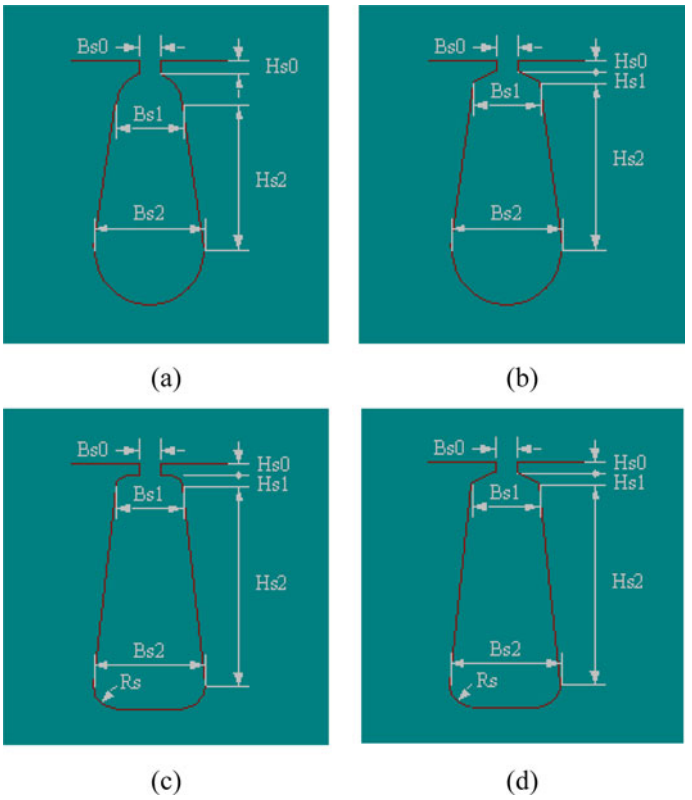


Fig. 4 Four different stator slot structure

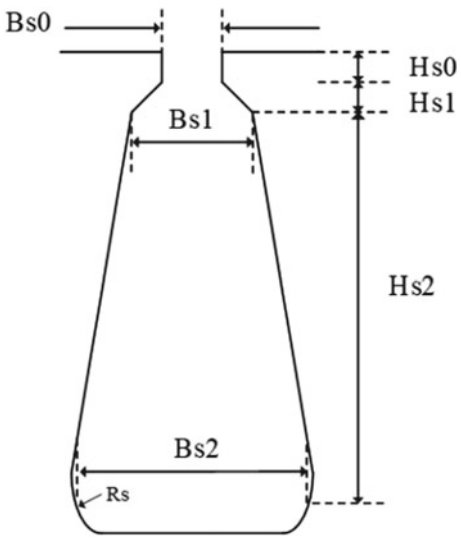


Fig. 5 The proposed slot shape design

Table 2 The dimensions of the proposed stator slot type

Hs0	3.2 (mm)
Hs1	1.8 (mm)
Hs2	26.4 (mm)
Bs0	5.2 (mm)
Bs1	10.5 (mm)
Bs2	13.1 (mm)
Rs	0 (mm)

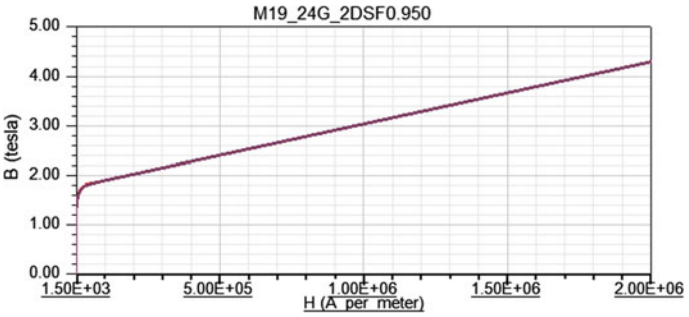


Fig. 6 B-H curve of M19_24G_2DSF0.950

are shown in Table 2. Rs value was selected as zero to obtain a high slot fill factor ratio.

The steel type of rotor and stator core are determined M19_24G_2DSF0.950. Figure 6 presents the curve characteristic of the magnetic properties of this material (B-H). Moreover, steel type of this material has 1,960,000 S/m bulk conductivity and 7267.5 kg/m^3 mass density.

Neodymium Iron Boron (NdFeB) magnetic material is widely used for generator design applications. NdFeB magnet is chosen as a permanent magnet in the proposed design. The bulk conductivity and the mass density are 625,000 S/m and 7550 kg/m^3 for this magnet, respectively.

In Table 3 the different design parameters of designed permanent magnet synchronous generator for wind turbine applications such as rated power, rated voltage rated speed, rotor position, number of poles, stator and rotor dimensions, efficiency and total net weight are presented.

Table 3 Characteristics structure of the permanent magnet alternator

Variables	Values
Nominal power (kW)	18
Nominal voltage (V)	380
Nominal speed (rpm)	120
Frequency (Hz)	50
Rotor position	Inner Rotor
Operating temperature	75 °C
Frictional loss	120 W
Nominal power factor	0.75
Pole number	50
Slot number	48
Outer size of the stator (mm)	500
Inner size of the stator (mm)	424
Outer size of the rotor (mm)	422
Inner size of the rotor (mm)	399
Efficiency (%)	90.5
Total net weight (kg)	111.941

4 Electromagnetic Transient Analysis

4.1 ANSYS/Maxwell Program

ANSYS/Maxwell is the industry-leading electromagnetic field simulator software for the analysis and modeling of electric motors, sensors, transducers and other electromagnetic and electromechanical devices. The nonlinear, time-varying motion of electromechanical components and their effect on drive circuit and control system design can be analyzed thanks to this program. Maxwell is a superior creative program that operates finite element analysis (FEA) method to solve the nonlinear and linear transient problems, time-dependent magnetic change for solid motion, AC electromagnetic, static magnetic, DC transmission, time-dependent electric exchange, expert design interfaces for electrical machines and converts and circuit and system simulation. The FEA method is a numerical process that can be performed to the many engineering problems. Stress analysis, heat transfer, fluid mechanics and electromagnetism problems for steady-state, variable-regime, linear, nonlinear states can be analyzed thanks to this useful method. The major principle of this method is to simplify and solve a complex problem. The solution region is divided into a plurality of simple, small, interconnected sub-regions called finite elements. Therefore, this method enables to analyze all particles separately. This case provides to increase the

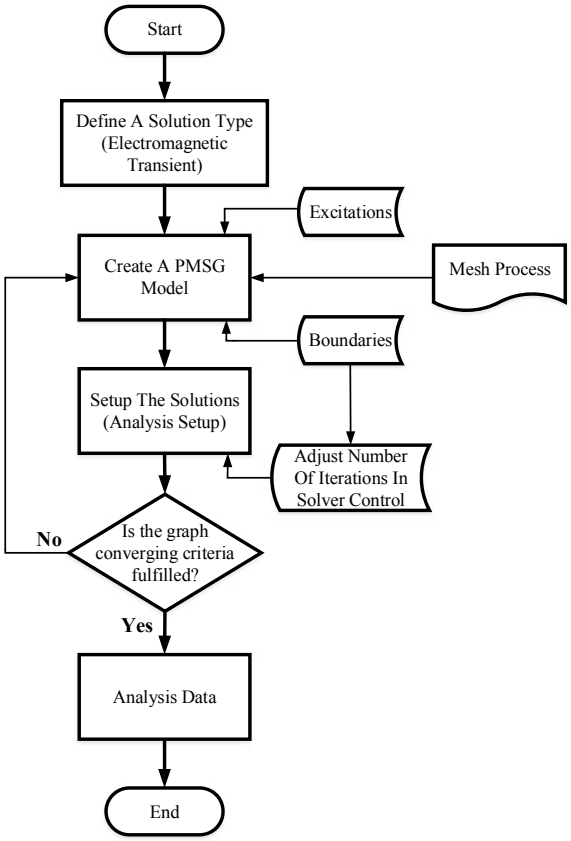


Fig. 7 ANSYS/Maxwell process flow for design and analysis of the PMSG

accuracy of the analysis results. In this study, ANSYS/Maxwell is used for electromagnetic transient analysis of the designed PMSG in order to observe the electromagnetic changes in the rotor and stator parts. The flow chart for the electromagnetic analysis of the PMSG is presented in Fig. 7.

4.2 Performance Results

In this chapter, a more efficient wind turbine alternator is modeled and analyzed in detail by using the ANSYS/Maxwell program. This program uses the Finite Element Method (FEM) that is used to solve engineering problems related to heat conduction, fluid mechanics, electrical and magnetic fields. FEM has known as an effective analysis method for solving common electromagnetic field problems. The first step in this method is to divide the problem into small substructures called finite elements. This process is called a mesh structure. The network structure consists of nodes and

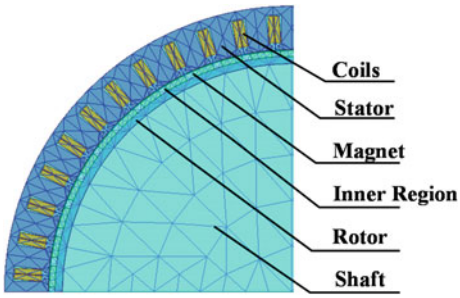


Fig. 8 Mesh structure on the stator, rotor and shaft of the PMSG

elements joining them. The desired size to be calculated in the solution of the problem is obtained by interpolation using the values at the joints. After the design process is completed, the mesh analysis structure is carried out on the stator, rotor and shaft of the PMSG. This operation is the key part of the analysis process because it allows the physical structures to be divided into sub-structures. The designed PMSG is divided into a hundred thousand (100,000) particles thanks to mesh operation so all particles are electromagnetically analyzed. The mesh operation on the transformer model is given in Fig. 8.

FEM calculations require quite long processes so there is various software on the computer that can perform these operations. Maxwell's equations are used to solve design problems in electrical machines. The equations are based on Ampere's law, Faraday's law of induction, Gauss' law for magnetism and Gauss' law. These equations are superior to solve electromagnetic field problems. Maxwell's four basic equations are given below.

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{9}$$

$$\nabla \times H = J + \frac{\partial D}{\partial t} \tag{10}$$

$$\nabla \cdot B = 0 \tag{11}$$

$$\nabla \times D = \rho \tag{12}$$

Maxwell's law in Eq. 9 describes Faraday's induction law. The time-dependent variation of the magnetic flux density passing through a surface area shows that the inverse polarity is equal to the electric field strength surrounding this area. Equation 10 shows that the electric current and the time-varying electric flux are equal to the magnitude of the magnetic field around an object. Equation 11 denotes that the magnetic flux entering any closed surface is equal to the magnetic flux output. Equation 12 states that the electric flux passing through a closed surface area is equal

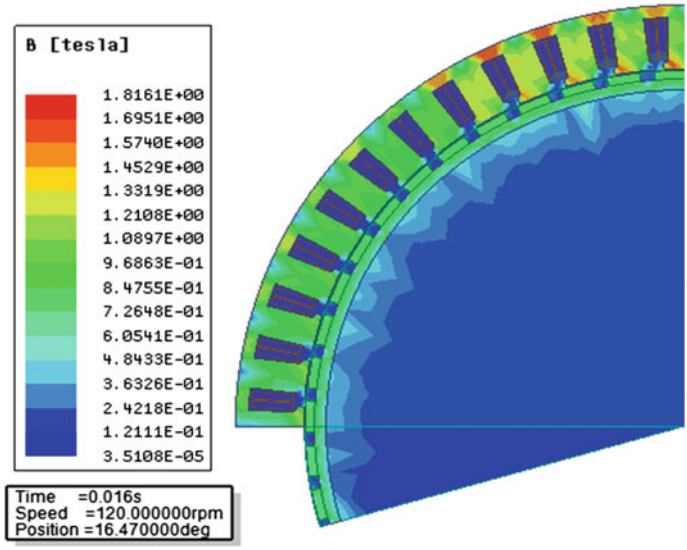


Fig. 9 Magnetic flux density analysis of PMSG

to the charge density in this area. These equations form the base of electromagnetic analysis of the electric machines in ANSYS/Maxwell.

The most important parts to be considered in electrical machine design are an air gap, stator yoke and rotor yoke magnetic flux changes. Figure 9 is shown that the magnetic flux density distribution of these parts is observed in detail thanks to electromagnetic analysis.

When the magnetic flux density distributions in Fig. 9 are examined, a magnetic flux density of 1.25 T in stator yoke, 1.21 T in stator coils and 1 T in rotor yoke are observed. In addition, when we examine some parts of the stator, magnetic flux density has reached around 1.8 T. In addition, Fig. 10 shows the distribution of magnetic flux lines in PMSG. It is seen that the stator region has higher magnetic flux lines. Flux lines that cannot complete a cycle through magnets are defined as leakage flux. These leakage fluxes show an increase in stator yoke and magnets.

The alternator is tested under the load condition by using ANSYS/Maxwell program. As a result of the alternator analysis, RMS voltages are approximately obtained as 230 V. The induced voltages and phase currents of the designed PMSG are shown in Figs. 11 and 12.

Another analysis result is the loss analysis of the PMSG. Figure 13 presents the copper losses of the PMSG. These losses occur in the rotor, stator and shaft parts of the designed PMSG.

It is observed in the analysis results that core losses are approximately 400 W. Copper losses are 1.4 kW under the load condition. Stray losses, friction and windage losses of the PMSG is obtained as 69 W. As a result of the loss analysis, the efficiency

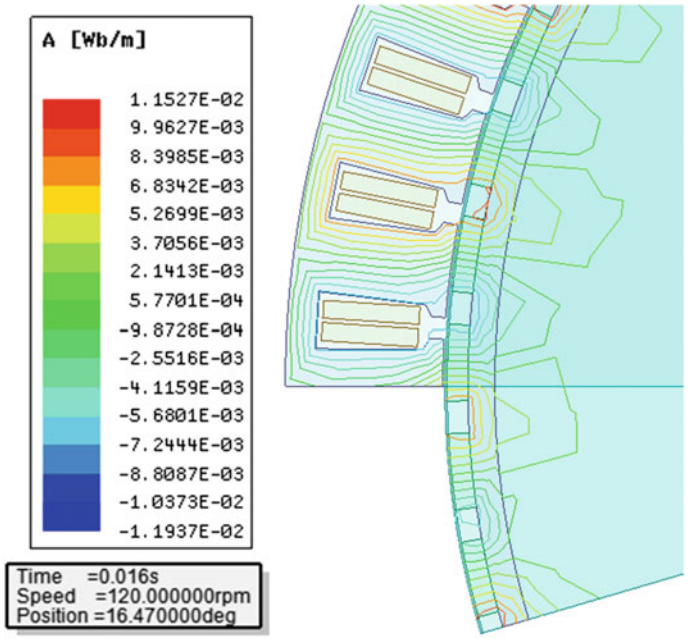


Fig. 10 Magnetic field lines analysis

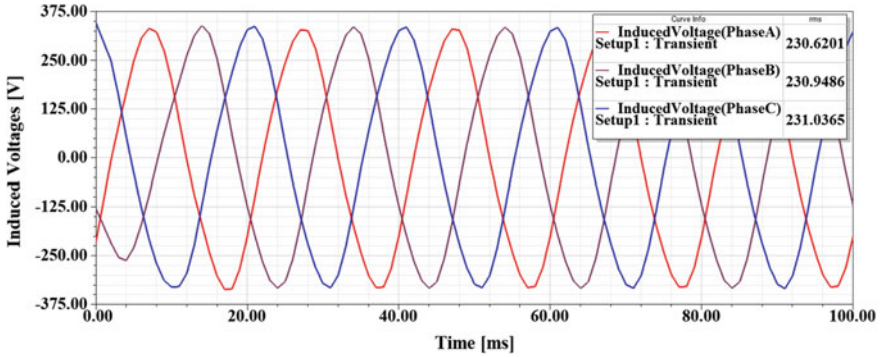


Fig. 11 Generated RMS output voltages of the designed PMSG

of the transformer is presented as 90.5%. 17.83 kW electric power is generated by 19.7 kW mechanical power shown in Fig. 14.

In this chapter, the proposed PMSG is designed for wind turbine applications in areas with low wind potential. The designed generator has significant advantages for these areas and also has higher efficiency than conventional generators. Such as the stator slot shape, stator fill factor and winding wire size critical parameters have been taken into consideration while achieving high efficiency. The characteristics of

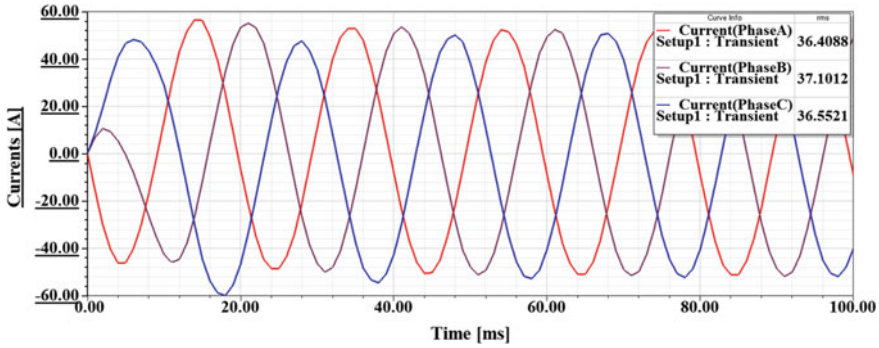


Fig. 12 Phase currents of the designed PMSG under the load condition

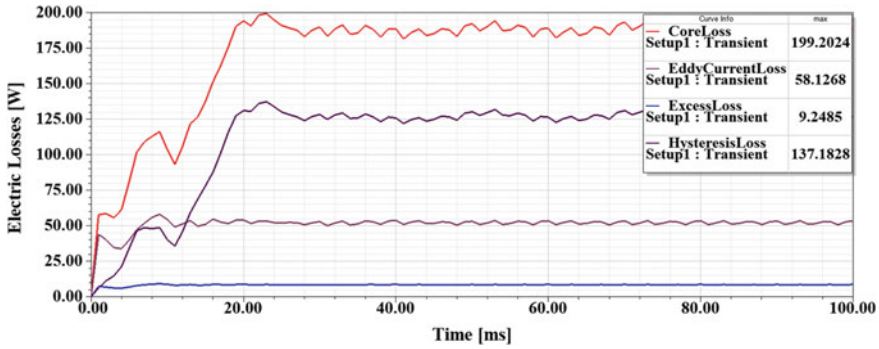


Fig. 13 Copper and core losses analysis of the PMSG

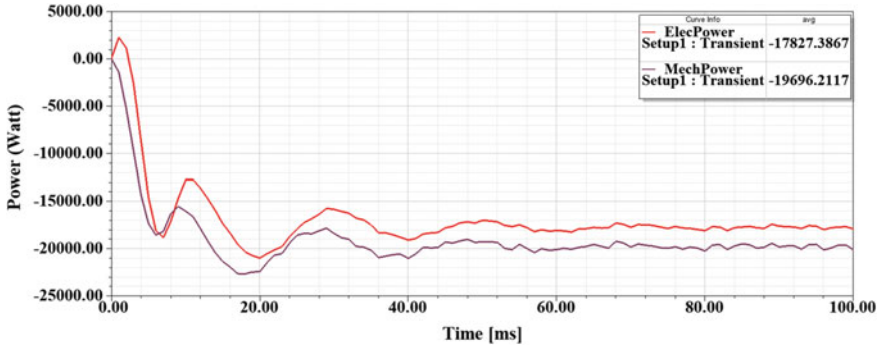


Fig. 14 Mechanical power and generated electric power of the PMSG

Table 4 The advantages and performance summary of the designed PMSG

Properties highlights of PMSG	Values
High electrical efficiency	90.5%
Convenient rotor type for low wind speed application	Inner rotor
Low cogging torque	50 poles/48 slots
The optimum level of the total weight	111.941 kg
Induced phase voltage to have low total harmonic distortion (THD)	380 V

the designed 18kW PMSG and the values obtained as a result of the analysis are presented in Table 4.

5 Conclusion

In this study, an 18kW permanent magnet synchronous generator is comprehensively designed for wind turbine applications. The simulation results of the proposed design were realized by ANSYS/Maxwell Electromagnetic Analysis Software. Firstly, the rated power, rated voltage, rated speed, rotor type, stator and rotor dimensions, stator slot shape of the proposed PMSG model were determined in detail during the design process. Subsequently, electromagnetic transient analysis of the PMSG was performed by using design parameters in order to obtain the maximum efficiency. The effect on the efficiency of the PMSG was observed by selecting suitable stator slot shapes, material types, electrical parameters, steel type stator and rotor core. Normally, it is difficult to converge the design results analytically due to multiple machine parameters and non-linear operation. Therefore, the use of FEM based analytic program is very important for machine design and analysis. Considering modeling and analysis results, it was seen that PMSG is designed efficiently and analyzed with high accuracy thanks to this program.

References

1. Ekinci F, Demirdelen T, Aksu IO et al (2019) A novel hybrid metaheuristic optimization method to estimate medium-term output power for horizontal axis wind turbine. Proc Ins Mech Eng Part A J Power Energy 0957650918821040
2. Widyan MS (2006) Design, optimization, construction and test of rare-earth permanent-magnet electrical machines with new topology for wind energy applications
3. Nerg J, Ruuskanen V (2013) Lumped-parameter-based thermal analysis of a doubly radial forced-air-cooled direct-driven permanent magnet wind generator. Mathemat Comput Simulat 90:218–229
4. Kiartzis S, Kladas A (2001) Deterministic and artificial intelligence approaches in optimizing permanent magnet generators for wind power applications. J Mater Process Technol 108:232–236

5. Arafat Y, Murshed M, Razzak MA (2015) Design and analysis of an outer rotor permanent magnet alternator for low-speed wind turbine. In: 2015 3rd international conference on green energy and technology (ICGET). IEEE, pp 1–7
6. Kalender O, Ege Y, Eskidere Ö et al (2015) A new axial flux permanent magnet synchronous alternator autonomously adapted to wind speeds. Measurement 69:87–94
7. Arafat MY, Murshed M, Hasan MM, Razzak MA (2016) Design aspects and performance analysis of inner and outer rotor permanent magnet alternator for direct driven low-speed wind turbine. In: 2016 2nd international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB). IEEE, pp 604–609
8. Bouaziz O, Jaafar I, Ammar FB (2016) 3D finite element modelling and comparative performance analysis between axial and radial flux machines for micro wind turbine application. In: 2016 international conference on electrical sciences and technologies in Maghreb (CISTEM). IEEE, pp 1–6
9. Bazzo T, Kolzer JF, Carlson R et al (2016) Multidisciplinary design optimization of direct-drive PMSG considering the site wind profile. Electric Power Syst Res 141:467–475
10. Meo S, Zohoori A, Vahedi A (2016) Optimal design of permanent magnet flux switching generator for wind applications via artificial neural network and multi-objective particle swarm optimization hybrid approach. Energy Convers Manage 110:230–239
11. Li X-M, Yang Z-X, Li Y-B et al (2016) Performance analysis of permanent magnet synchronous generators for wind energy conversion system. In: 2016 international conference on advanced mechatronic systems (ICAMechS). IEEE, pp 544–549
12. Kurt E, Gör H, Döner U (2016) Electromagnetic design of a new axial and radial flux generator with the rotor back-irons. Int J Hydrogen Energy 41:7019–7026
13. Fang H, Wang D (2016) A novel design method of permanent magnet synchronous generator from perspective of permanent magnet material saving. IEEE Trans Energy Convers 32:48–54
14. Puri V, Chauhan YK, Singh N (2017) A comparative design study and analysis of inner and outer rotor permanent magnet synchronous machine for power generation in vertical axis wind turbine using GSA and GSA-PSO. Sustain Energy Technol Assess 23:136–148
15. Laxminarayan SS, Singh M, Saifee AH, Mittal A (2017) Design, modeling and simulation of variable speed axial flux permanent magnet wind generator. Sustain Energy Technol Assess 19:114–124
16. Minaz MR, Celebi M (2017) Design and analysis of a new axial flux coreless PMSG with three rotors and double stators. Results Phys 7:183–188
17. Yoo J-H, Park C-S, Jung T-U (2017) Permanent magnet structure optimization for cogging torque reduction of outer rotor type radial flux permanent magnet generator. In: 2017 IEEE international electric machines and drives conference (IEMDC). IEEE, pp 1–6
18. Addin Yousefian H, Kelk HM (2018) A unique optimized double-stator permanent-magnet synchronous generator in high-power wind plants. Energy 143:973–979
19. Asef P, Perpina RB, Barzegaran MR (2018) An innovative natural air-cooling system technique for temperature-rise suppression on the permanent magnet synchronous machines. Electric Power Syst Res 154:174–181
20. Gul W, Gao Q, Lenwari W (2018) Optimal design of a 5MW double stator single rotor permanent magnet synchronous generator for offshore direct drive wind turbines using the genetic algorithm. In: 2018 21st international conference on electrical machines and systems (ICEMS). IEEE, pp 149–155
21. Khan S, Amin S, Bukhari SSH (2019) Design and comparative performance analysis of inner rotor and inner stator axial flux permanent magnet synchronous generator for wind turbine applications. In: 2019 2nd international conference on computing, mathematics and engineering technologies (iCoMET). IEEE, pp 1–7
22. Jaen-Sola P, McDonald AS, Oterkus E (2019) Lightweight design of direct-drive wind turbine electrical generators: a comparison between steel and composite material structures. Ocean Eng 181:330–341

23. Mihai AM, Benelghali S, Livadaru L et al (2012) FEM analysis upon significance of different permanent magnet types used in a five-phase PM generator for gearless small-scale wind. In: 2012 XXth international conference on electrical machines. IEEE, pp 267–273
24. Arumugam D, Logamani P, Karuppiah S, Thangaraj B (2017) Performance evaluation of PMSG for aircraft applications. Energy Procedia 117:385–392

Power and Energy System Modeling Based on Modified Tellegen Principle

Milan Stork and Daniel Mayer

Abstract The chapter deals with a problem of power and energy modeling of dynamic systems and their numerical solutions. The proposed approach is based on modified Tellegen's theorem well known from electrical circuits solving. The new of this approach is that it is based on the power calculation for different linear and nonlinear physical systems. It is supposed that system is described by state space equations and from these equations the power and energy are calculated. The results can be used for electrical circuits, but also for other real physical system e.g. mechanical systems, heat transfer etc. which are described by state space equations. Thus, mathematically as well as physically correct results are received. Certain known and often used system representation structures are used and their transformations. The theory is supported by solved examples. The mathematical origin, derivation and results of simulations are given in this chapter. It is important mark here a close relationship between first and second Kirchhoff's laws ensuring physical rightness and link to Tellegen's theorem. It is also important to note that inner product incorporate instantaneous power dissipated on resistors and instantaneous power on inductors and capacitors. At last, let's note that Tellegen's principle can be used not only to electrical circuits but as well any model of a physical correct system with lumped parameters. Therefore if the system is described accurately by state space equations, currents and voltages can be substituted by state space variables and it's derivations. In the chapter, the nonlinear differential equations of linear and nonlinear systems are numerically solved and optimization methods are used in some circumstances. The examples in chapter include also solution of chaotic systems. It is important to note that solution in this chapter is based on power and energy and therefore can

M. Stork (✉)
Department of Electronics and Information Technology/RICE, University of West Bohemia, Plzen, Czechia
e-mail: stork@kae.zcu.cz

D. Mayer
Department of Electrical and Computational Engineering, University of West Bohemia, Plzen, Czechia
e-mail: mayer@kte.zcu.cz

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_31

be useful also for modeling some devices used in renewable power systems e.g. batteries, capacitors and ultracapacitors, photovoltaic panels etc.

Keywords Scalar product · State space energy · Modified tellegen · Transformation

Symbols/Parameters

V	voltage
I	current
x_i	state space variable
E	energy
P	power
C	capacitor
L	inductor
R	resistor
T	transformation matrices
W	prescribed value
$(..)^T$	matrix or vector transposition

1 Introduction

In this part on the beginning the classical Tellegen theorem is described and will show that generalized theorem can be used for system described by state space equations.

Let N be a physically correct electrical circuit with the k lumped parameters. Parameters of N can be linear or nonlinear, hysteresis or non-hysteresis, time-constant or variable. The currents in the branches are $i_k(t)$ and the voltage on these branches are $v_k(t)$. Theorem 1. (Classic Tellegen's theorem). For branch currents $i_k(t)$ and branch voltages $v_k(t)$ holds true:

$$\left\langle i(t)^T, v(t) \right\rangle = \sum\nolimits_{k=1}^{b} i_k(t) v_k(t) = 0 \tag{1}$$

It is worth noticing a close relation between first and second Kirchhoff's laws (ensuring physical correctness of N) and Tellegen's theorem. It is also important to note that inner product according (1) include instantaneous power dissipated on resistors and instantaneous power on inductors and capacitors. Finally, let's note that Tellegen's theorem applies not only to electrical circuits but to any appropriately described model of a physical correct system with lumped parameters, for example mechanical, thermal, etc. [1–5]. The appropriately described system means

that matrix A of state space equations is antisymetric. The general, nonlinear or linear dissipative systems can be described by state space equations

$$\begin{aligned} \dot{x}(t) &= f[x(t)] + Bu(t), x(t_0) = x_0, \\ y(t) &= Cx(t) \end{aligned} \tag{2}$$

where x_0 is vector of initial conditions and with an state velocity vector field described by

$$f(x) = A(x)x \tag{3}$$

It has been shown in [6–10] that a special form of a dissipative state space system representation named as "dissipation normal form" exists and its structure can be described by the following matrices:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \cdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -a_{11} & a_2 & 0 & 0 & 0 & 0 \\ -a_2 & -a_{22} & a_3 & 0 & 0 & 0 \\ 0 & -a_3 & -a_{33} & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & a_{n-1} & 0 \\ 0 & 0 & 0 & -a_{n-1} & -a_{n-1,n-1} & a_n \\ 0 & 0 & 0 & 0 & -\alpha_n & -a_{n,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \cdots \\ b_{n-1} \\ b_n \end{bmatrix} \cdot u$$

$$y = C \cdot \boldsymbol{x} + D \cdot u \tag{4}$$

The matrix A must be anti-symmetric therefore

$$a_{ij} = -a_{ji} \quad where \quad i \neq j \tag{5}$$

It is important to note that a_{ij} can be not only constants, but also functions in nonlinear case. Energy E of the system described by anti-symmetric matrix (energy of state space variables) is

$$E = \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2 \ldots + x_n^2\right) = \frac{1}{2}\sum_{i=1}^{n} x_i^2 \tag{6}$$

Power is derived as derivation of energy

$$\begin{aligned} P &= \frac{d}{dt}\left(\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) = x_1\frac{dx_1}{dt} + x_2\frac{dx_2}{dt} \ldots + x_n\frac{dx_n}{dt} \\ &= x_1(-a_{11}x_1 + a_2x_2) + x_2(-a_2x_1 - a_{22}x_2 + a_3x_3) \\ &\quad + x_3(-a_3x_2 - a_{33}x_3 + a_4x_4) \ldots \end{aligned}$$

$$= -a_{11}x_1^2 + \underbrace{a_2x_1x_2 - a_2x_1x_2}_{0} - a_{22}x_2^2 + \underbrace{a_3x_2x_3 - a_3x_2x_3}_{0} \ldots$$

$$= -a_{11}x_1^2 - a_{22}x_2^2 \ldots - a_{nn}x_n^2 = -\sum_{i=1}^{n} a_{ii}x_i^2 \tag{7}$$

According Eqs. (6) and (7) it is possible determine stability of the system, because if $E > 0$ and $P < 0$. The energy function E can be used as a Lyapunov function. If trace of matrix $A < 0$ (sum of the diagonal numbers of $A < 0$) the system is stable (dissipative), if trace of matrix $A = 0$ the system is conservative, if trace of matrix $A > 0$ the system is antidissipative (unstable).

Therefore if the system is described according previous mentioned state space equations, currents and voltages can be substituted by state space variables and it's derivations according Eq. (8)

$$\left\langle x(t)^T, \frac{dx(t)}{dt} \right\rangle = x_1(t)\frac{dx_1(t)}{dt} + x_2(t)\frac{dx_2(t)}{dt} \ldots . + x_n(t)\frac{dx_n(t)}{dt} \tag{8}$$

It is to note that previous Eq. (8) is a close relation between different types of physical systems and Tellegen's theorem (terms in Eq. (8) are "power" terms). Therefore it is possible derive power or energy in some system by means of generalized Tellegen's theorem, which will shown in first example.

Notice: Usually the system is described by matrix which is not anti-symmetric; it can be transformed to anti-symmetric by similarity transformation, see the next simple example.

Example 1 The simple circuit according Fig. 1 can be described by voltage-current equations

$$\begin{aligned} \underbrace{L\frac{di_L}{dt}}_{v_L} &= -R_2 i_L + v_C \\ \underbrace{C\frac{dv_C}{dt}}_{i_C} &= -i_L + \frac{V_1 - v_C}{R_1} \end{aligned} \tag{9}$$

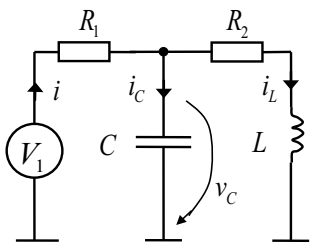


Fig. 1 The circuit used for example 1, $R_1 = 5\ \Omega$; $R_2 = 0.01\ \Omega$; $L = 0.5$ H; $C = 0.2$ F

Previous equation can be after some manipulation converted also to 2nd order differential equation.

$$\begin{aligned}\frac{d^2 i_L}{dt^2} &= -\frac{R_2}{L}\frac{di_L}{dt} + \frac{1}{L}\frac{dv_C}{dt} \\ \frac{dv_C}{dt} &= -\frac{i_L}{C} - \frac{1}{CR_1}v_C + \frac{V_1}{CR_1}\end{aligned} \tag{10}$$

The final, 2nd order differential equation is

$$\frac{d^2 i_L(t)}{dt^2} + \left(\frac{R_2}{L} + \frac{1}{CR_1}\right)\frac{di_L(t)}{dt} + \left(\frac{1}{CL} + \frac{R_2}{CLR_1}\right)i_L(t) = \frac{V_1(t)}{CLR_1} \tag{11}$$

After substituting values C, L, R of components

$$\frac{d^2 i_L(t)}{dt^2} + (0.02 + 1)\frac{di_L(t)}{dt} + (10 + 0.02)i_L(t) = 2V_1(t) \tag{12}$$

Roots are: $-0.5010 + 3.1227\text{i}$ and $-0.5010 - 3.1227\text{i}$. Equation (10) can be rewritten in state space form with state variables x_1 and x_2 as ($x_1 = i_L$, $x_2 = v_C$)

$$\begin{aligned}\frac{dx_1}{dt} &= -\frac{R_2}{L}x_1 + \frac{1}{L}x_2 \\ \frac{dx_2}{dt} &= -\frac{1}{C}x_1 - \frac{1}{CR_1}x_2 + \frac{V_1}{CR_1}\end{aligned} \tag{13}$$

The Eq. (13) in matrix forms is

$$\begin{bmatrix}\frac{dx_1}{dt} \\ \frac{dx_2}{dt}\end{bmatrix} = \underbrace{\begin{bmatrix}-\frac{R_2}{L} & \frac{1}{L} \\ -\frac{1}{C} & -\frac{1}{CR_1}\end{bmatrix}}_{A} \cdot \begin{bmatrix}x_1 \\ x_2\end{bmatrix} + \underbrace{\begin{bmatrix}0 \\ \frac{1}{CR_1}\end{bmatrix}}_{B} \cdot V_1 \tag{14}$$

For parts, $R_1 = 5\ \Omega$; $R_2 = 0.01\ \Omega$; $L = 0.5$ H; $C = 0.2$ F is

$$\begin{bmatrix}\frac{dx_1}{dt} \\ \frac{dx_2}{dt}\end{bmatrix} = \begin{bmatrix}-0.02 & 2 \\ -5 & -1\end{bmatrix} \cdot \begin{bmatrix}x_1 \\ x_2\end{bmatrix} + \begin{bmatrix}0 \\ 1\end{bmatrix} \cdot V_1 \tag{15}$$

Unfortunately, the matrix A is not anti-symmetric because of different values of C and L, therefore system according Eq. (14) must be transformed by state-space similarity transformation (the methods of similarity transformation will shortly described in next chapter). Transformed matrices are A_1, B_1 and C_1.

$$A_1 = T \cdot A \cdot T^{-1};\ B_1 = T \cdot B;\ C_1 = C \cdot T^{-1};\ \boldsymbol{x}_1 = T \cdot \boldsymbol{x};\ T \cdot T^{-1} = I \tag{16}$$

The transformation matrices T and T^{-1} for system according Eq. (14) are

$$T = \begin{bmatrix} \sqrt{L} & 0 \\ 0 & \sqrt{C} \end{bmatrix}; T^{-1} = \begin{bmatrix} \frac{1}{\sqrt{L}} & 0 \\ 0 & \frac{1}{\sqrt{C}} \end{bmatrix}; L \neq 0; C \neq 0 \tag{17}$$

Matrices after transformation are (now, matrix A_1 is anti-symmetrical)

$$A_1 = \begin{bmatrix} -\frac{R_2}{L} & \frac{1}{\sqrt{CL}} \\ -\frac{1}{\sqrt{CL}} & -\frac{1}{CR_1} \end{bmatrix} = \begin{bmatrix} -a_{11} & a_2 \\ -a_2 & -a_{22} \end{bmatrix}; B_1 = \begin{bmatrix} 0 \\ \frac{1}{R_1\sqrt{C}} \end{bmatrix} = \begin{bmatrix} 0 \\ b_2 \end{bmatrix} \tag{18}$$

where after transformation a_{ij} and b_2 are

$$a_{11} = \frac{R_2}{L}; a_2 = \frac{1}{\sqrt{CL}}; a_{22} = \frac{1}{CR_1}; b_2 = \frac{1}{\sqrt{C}R_1} \tag{19}$$

where relations between transformed state variables x_{1T}, x_{2T} and current i_L and voltage v_C are

$$x_T = T \cdot x = \begin{bmatrix} \sqrt{L} & 0 \\ 0 & \sqrt{C} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow x_{1T} = \sqrt{L} \cdot i_L; x_{2T} = \sqrt{C} \cdot v_C \tag{20}$$

Energy of state space variables is

$$E(t) = \frac{1}{2}\left(x_{1T}^2(t) + x_{2T}^2(t)\right) \tag{21}$$

Power is given as inner product according Eq. (22)

$$\begin{aligned} P(t) = \frac{dE(t)}{dt} &= \left\langle x_T(t)^T, \frac{dx_T(t)}{dt} \right\rangle = \begin{bmatrix} x_{1T} & x_{2T} \end{bmatrix} \left(A_1 \begin{bmatrix} x_{1T} \\ x_{2T} \end{bmatrix} + B_1 V_1 \right) \\ &= -a_{11}x_{1T}^2 + \underbrace{a_2 x_{1T} x_{2T} - a_2 x_{1T} x_{2T}}_{0} - a_{22}x_{2T}^2 + b_2 V_1 \cdot x_{2T} \end{aligned} \tag{22}$$

After substituting for a_{11}, a_2 and a_{22} from (19) into $P(t)$

$$P(t) = -\frac{R_2}{L}x_{1T}^2 + \underbrace{\frac{x_{1T}x_{2T}}{\sqrt{LC}} - \frac{x_{1T}x_{2T}}{\sqrt{LC}}}_{0} - \frac{x_{2T}^2}{CR_1} + \frac{V_1 \cdot x_{2T}}{\sqrt{C} \cdot R_1} \tag{23}$$

State space variables aren't real voltage and current; therefore for calculating real values the substitution must be done. Using Eqs. (20) and (23) powers are derived as

$$P(t) = -\frac{R_2}{L}x_{1T}^2 - \frac{x_{2T}^2}{CR_1} + \frac{V_1 \cdot x_{2T}}{\sqrt{C} \cdot R_1} = -\frac{R_2}{L}\left(\sqrt{L} \cdot i_L\right)^2 - \frac{\left(\sqrt{C} \cdot v_C\right)^2}{CR_1} + \frac{V_1 \cdot \left(\sqrt{C} \cdot v_C\right)}{\sqrt{C} \cdot R_1}$$
$$= -R_2 \cdot i_L^2 - \frac{v_C^2}{R_1} + \frac{V_1 \cdot v_C}{R_1} \tag{24}$$

Eq. (24) is rewritten for appropriate power description

$$\begin{aligned} P(t) &= -R_2 \cdot i_L^2 - \frac{v_C^2}{R_1} + \frac{V_1 \cdot v_C}{R_1} \\ &= -R_2 \cdot i_L^2 - \frac{v_C^2}{R_1} + \frac{V_1 \cdot v_C}{R_1} + \left(\frac{V_1 \cdot v_C}{R_1} - \frac{V_1 \cdot v_C}{R_1}\right) + \left(\frac{V_1^2}{R_1} - \frac{V_1^2}{R_1}\right) \\ &= -R_2 \cdot i_L^2 - \frac{v_C^2}{R_1} + 2\frac{V_1 \cdot v_C}{R_1} - \frac{V_1^2}{R_1} - \frac{V_1 \cdot v_C}{R_1} + \frac{V_1^2}{R_1} \\ &= V_1 \frac{V_1 - v_C}{R_1} - R_2 \cdot i_L^2 - \frac{(V_1 - v_C)^2}{R_1} \end{aligned} \tag{25}$$

Final result is

$$P(t) = \underbrace{V_1(t)\frac{V_1(t) - v_C(t)}{R_1}}_{P_i} \underbrace{-R_2 i_L^2(t) - \frac{(V_1(t) - v_C(t))^2}{R_1}}_{P_D} = 0 \tag{26}$$

The power $P(t)$ presented by (26) contain 3 powers. The input power P_i and two dissipations powers P_D (dissipation on R_1 and R_2). From this result is shown that result contain only dissipative powers. Result is not affected by energy storage elements L and C.

On the end of this chapter the results of simulation of circuit according Fig. 1 is presented. The circuit parts values are R_1= 5 Ω; R_2= 0.01 Ω; L = 0.5 H; C = 0.2 F.

On the first, circuit simulation started from initial condition (voltage on capacitor v_C= 1.41 V). In Fig. 2 the time response of state space variables i_L and v_C are shown. In Fig. 3 the state space energy of inductor, capacitor and total energy versus time is displayed. The time evolution of dissipated power on resistors R_1 and R_2 is shown in Fig. 4. In Fig. 5 the time response of energy of state variables (curve 1) and energy dissipated on resistors R_1 and R_2 (curve 2). Initial energy on capacitor $E_C = 0.5 \cdot C \cdot v_C^2 = 0.5 \cdot 0.2 \cdot 1.41^2 = 0.2$ is dissipated as heat on resistors (initial energy on inductor = 0).

The Figs. 6, 7, 8 and 9 shows the results of the same circuit, but the difference is that the circuit is driven by a constant voltage of 1 V with zero initial conditions. In contrast to Fig. 5, in Fig. 9, the energy dissipated on the resistors continues to increase as the circuit is energized by the DC voltage and the current continues to flow, while the energy at L and C stabilizes at the final values.

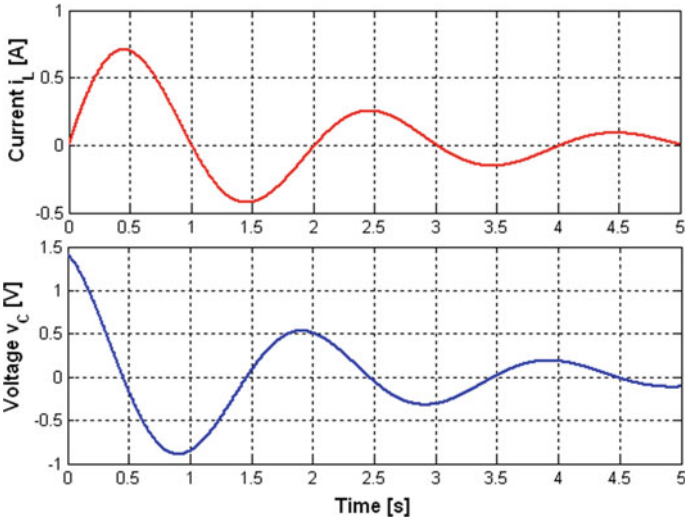


Fig. 2 The time evolution of state space variables i_L (top) and v_C (bottom) for circuit in Fig. 1. $V_1 = 0$, initial condition $v_C(0) > 0$

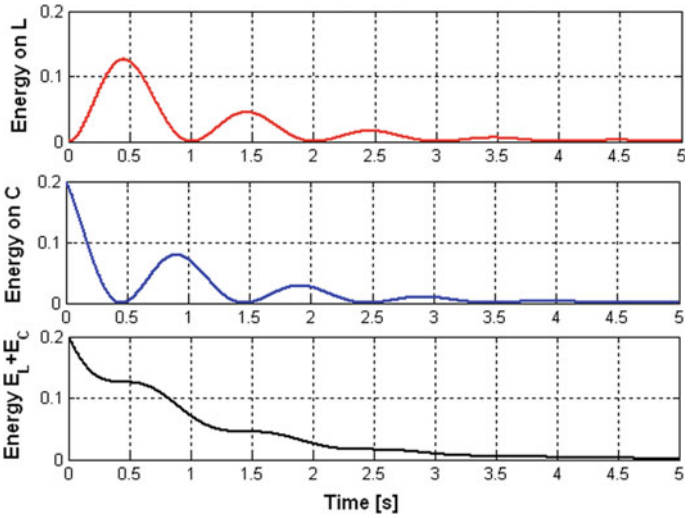


Fig. 3 The state space energy versus time for circuit according Fig. 1. From top to bottom: Energy on inductor L, energy on capacitor C and total energy $E_L + E_C$. $V_1 = 0$, initial condition $v_C(0) > 0$

The real RLC system (see Fig. 1) after similarity transformation was changed (see Eq. (20)—relations between state space variables and current and voltage) because are scaled. Moreover the coordinate transformation is usually complicated, especially

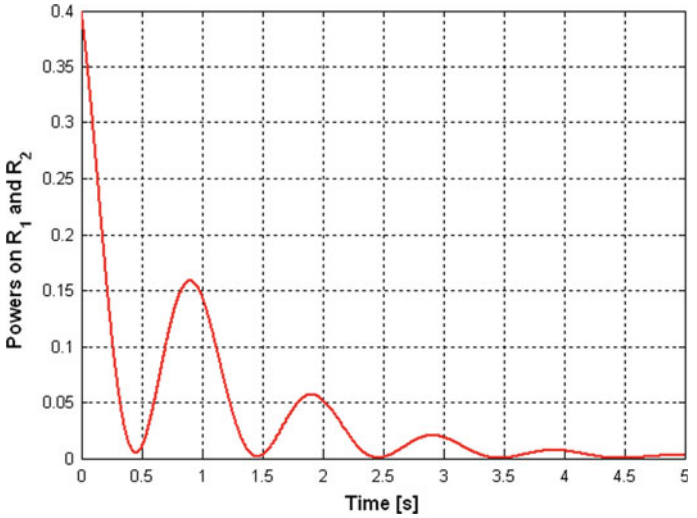


Fig. 4 The time evolution of power dissipated on resistors R_1 and R_2

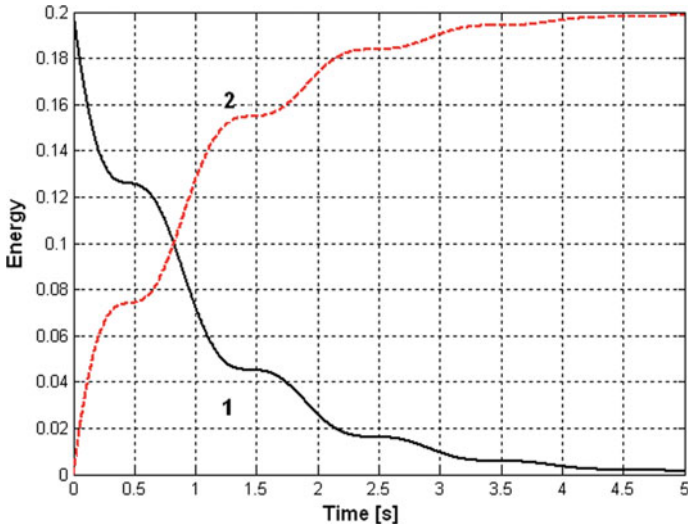


Fig. 5 The time evolution of energy of state variables (curve 1) and energy dissipated on resistors R_1 and R_2 (curve 2). $V_1 = 0$, initial condition $v_C(0) > 0$

for high order systems or nonlinear systems. Therefore the new approach—*modified Tellegen's approach* was derived.

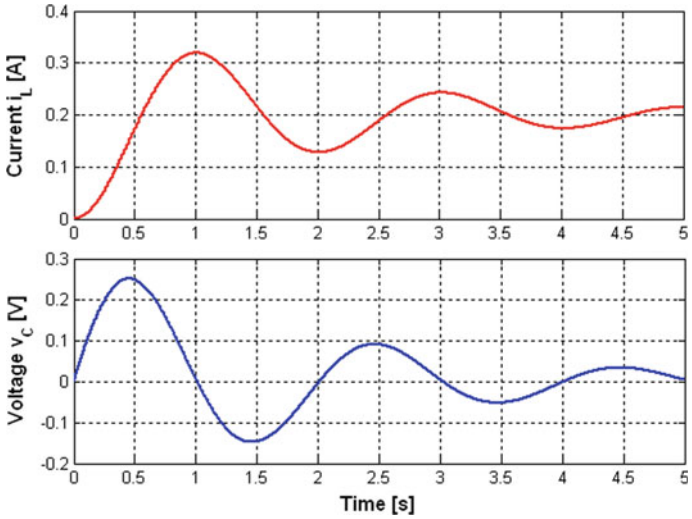


Fig. 6 The time evolution of state space variables i_L (top) and v_C (bottom) for circuit in Fig. 1. $V_1 = 1$, initial conditions $= 0$

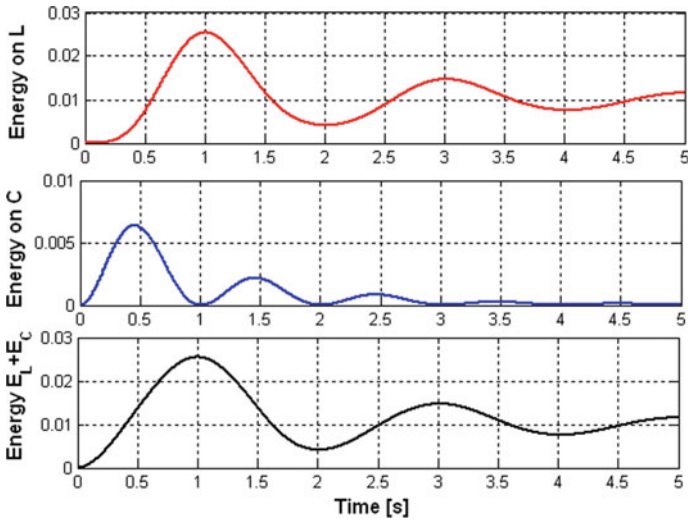


Fig. 7 The state space energy versus time for circuit according Fig. 1. From top to bottom: Energy on inductor L, energy on capacitor C, total energy $E_L + E_C$. $V_1 = 1$, initial conditions $= 0$

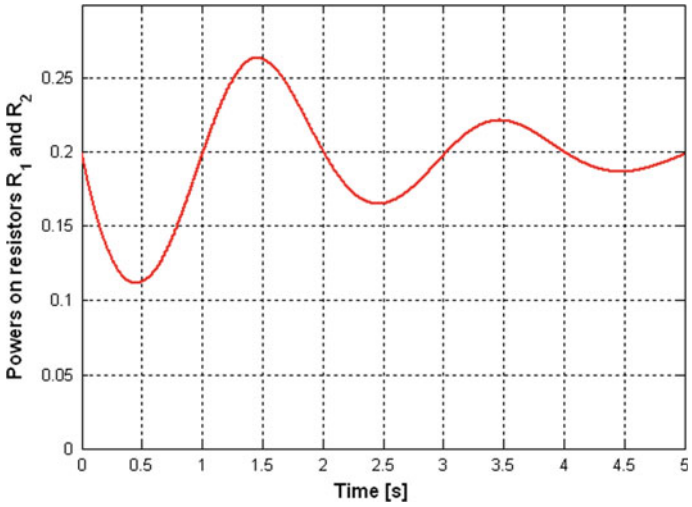


Fig. 8 The time evolution of power dissipated on resistors R_1 and R_2. $V_1 = 1$, initial conditions = 0

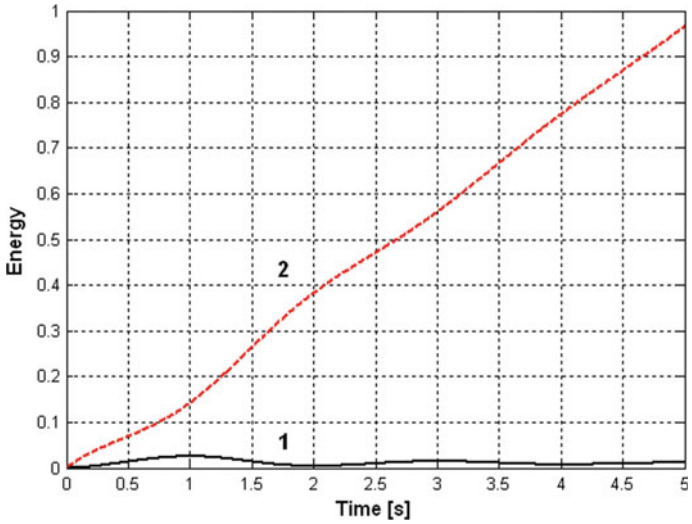


Fig. 9 The time evolution of energy of state variables (curve 1) and energy dissipated on resistors R_1 and R_2 (curve 2) for circuit according Fig. 1. $V_1 = 1$, initial conditions = 0

2 Modified Tellegen's Theorem

Suppose once more the circuit according Fig. 1, described by state space equations, Eq. (9) [11–14]. Energy of the state variables is

$$E(t) = \frac{1}{2}L \cdot i_L^2(t) + \frac{1}{2}C \cdot v_C^2(t) \tag{27}$$

Power is

$$P(t) = \frac{dE(t)}{dt} = L \cdot i_L \frac{di_L}{dt} + C \cdot v_C \frac{dv_C}{dt} \tag{28}$$

After substitution

$$\begin{aligned} P(t) = \frac{dE(t)}{dt} &= L \cdot i_L \left(-\frac{R_2}{L} i_L + \frac{1}{L} v_C \right) \\ &+ C \cdot v_C \left(-\frac{1}{C} i_L - \frac{1}{CR_1} v_C + \frac{V_1}{CR_1} \right) \\ &= -R_2 \cdot i_L^2 + \underbrace{i_L \cdot v_C - i_L \cdot v_C}_{0} - \frac{1}{R_1} v_C^2 + \frac{V_1}{R_1} v_C \end{aligned} \tag{29}$$

The modified relation for energy is given by the scalar product

$$E_M(t) = \frac{1}{2} \left\langle (F \cdot \boldsymbol{x}(t))^T, \boldsymbol{x}(t) \right\rangle \tag{30}$$

Modified mathematical relationship for power is

$$P_M(t) = \left\langle (F \cdot x(t))^T, \frac{dx(t)}{dt} \right\rangle \tag{31}$$

where $P_M(t)$ is dissipated power calculated by means of modified Tellegen's theorem. The main difference from the previous approach is that a diagonal square matrix F size of n is used that contains accumulation element values (values of L and C elements). For example 1, the state space equation is Eq. (9), which can be rewritten as

$$\begin{bmatrix} L\frac{dx_1}{dt} \\ C\frac{dx_2}{dt} \end{bmatrix} = \underbrace{\begin{bmatrix} -R_2 & 1 \\ -1 & -\frac{1}{R_1} \end{bmatrix}}_{A_M} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ \frac{1}{R_1} \end{bmatrix}}_{B_M} \cdot V_1$$

$$= \underbrace{\begin{bmatrix} L & 0 \\ 0 & C \end{bmatrix}}_{F} \cdot \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \underbrace{\begin{bmatrix} -R_2 & 1 \\ -1 & -\frac{1}{R_1} \end{bmatrix}}_{A_M} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ \frac{1}{R_1} \end{bmatrix}}_{B_M} \cdot V_1 \tag{32}$$

therefore matrix F, term $F_{11} = L$ (inductor value) and term $F_{22} = C$ (capacitor value).

$$F = \begin{bmatrix} F_{11} & 0 \\ 0 & F_{22} \end{bmatrix} = \begin{bmatrix} L & 0 \\ 0 & C \end{bmatrix} \tag{33}$$

Power $P_M(t)$ calculated from (31) is

$$\begin{aligned} P_M(t) &= \left\langle (F \cdot x(t))^T, \frac{dx(t)}{dt} \right\rangle = \left\langle \left(\begin{bmatrix} L & 0 \\ 0 & C \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)^T, \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \right\rangle \\ &= \left\langle \begin{bmatrix} Lx_1 & Cx_2 \end{bmatrix}, \begin{bmatrix} -\frac{R_2}{L}x_1 + \frac{1}{L}x_2 \\ -\frac{1}{C}x_1 - \frac{1}{CR_1}x_2 + \frac{V_1}{CR_1} \end{bmatrix} \right\rangle \end{aligned} \tag{34}$$

Final result of the previous equation is

$$\begin{aligned} P_M(t) &= \left\langle Fx(t), \frac{dx(t)}{dt} \right\rangle \\ &= -R_2 x_1^2(t) + \underbrace{x_1(t)x_2(t) - x_1(t)x_2(t)}_{0} - \frac{1}{R_1}x_2^2(t) + \frac{V_1(t)}{R_1}x_2(t) \\ &= -R_2 x_1^2(t) - \frac{1}{R_1}x_2^2(t) + \frac{V_1(t)}{R_1}x_2(t) \end{aligned} \tag{35}$$

where $x_1 = i_L$ and $x_2 = v_C$ and therefore result is the same as Eq. (24). For this calculation the similarity transformation was not used. The time evolution of signals, energy and power, all is the same as previous results, therefore Figs. 2, 3, 4, 5, 6, 7, 8 and 9 are exactly the same for this approach.

3 Similarity Transformation

The example of similarity transformation (coordinate transformation) was used in chapter 31.1 for second order system. In this section, the transformation of the state space system described by unbalanced anti-symmetric matrix A into a system with a balanced anti-symmetric matrix A_T is described. This is because the state space systems described by this type of matrix have some advantages. The principle of state transformation was presented in Eqs. (16–18) in subchapter 31.1.

For second order system with non-balanced anti-symmetric matrix A (different absolute values of a_{12} and a_{21}), the transformation to anti-symmetric is

$$A = \begin{bmatrix} a_{11} & a_{12} \\ -a_{21} & a_{22} \end{bmatrix};\ or\ A = \begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & a_{22} \end{bmatrix};\ a_{ij} \neq 0, i \neq j$$
$$T = \begin{bmatrix} 1/\sqrt{|a_{12}|} & 0 \\ 0 & 1/\sqrt{|a_{21}|} \end{bmatrix}; T^{-1} = \begin{bmatrix} \sqrt{|a_{12}|} & 0 \\ 0 & \sqrt{|a_{21}|} \end{bmatrix} \tag{36}$$

Transformations matrices are diagonal and therefore are simple calculate of inversion (reciprocal values on trace). For 3th order system

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ -a_{21} & a_{22} & a_{23} \\ 0 & -a_{32} & a_{33} \end{bmatrix} or\ A = \begin{bmatrix} a_{11} & -a_{12} & 0 \\ a_{21} & a_{22} & -a_{23} \\ 0 & a_{32} & a_{33} \end{bmatrix}$$
$$or\ A = \begin{bmatrix} a_{11} & -a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ 0 & -a_{32} & a_{33} \end{bmatrix} or\ A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ -a_{21} & a_{22} & -a_{23} \\ 0 & a_{32} & a_{33} \end{bmatrix}$$
$$T = \begin{bmatrix} \frac{1}{\sqrt{|a_{12}|}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{|a_{21}|}} & 0 \\ 0 & 0 & \sqrt{\left|\frac{a_{23}}{a_{21} \cdot a_{32}}\right|} \end{bmatrix}; a_{ij} \neq 0, i \neq j \tag{37}$$

For 4th order system (only transformation matrix is shown):

$$T = \begin{bmatrix} \frac{1}{\sqrt{|a_{12}|}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{|a_{21}|}} & 0 & 0 \\ 0 & 0 & \sqrt{\left|\frac{a_{23}}{a_{21} \cdot a_{32}}\right|} & 0 \\ 0 & 0 & 0 & \sqrt{\left|\frac{a_{23} \cdot a_{34}}{a_{21} \cdot a_{32} \cdot a_{43}}\right|} \end{bmatrix} \tag{38}$$

For 5th order system the trace of transformation matrix T is

$$T_{11} = \frac{1}{\sqrt{|a_{12}|}}; T_{22} = \frac{1}{\sqrt{|a_{21}|}}; T_{33} = \sqrt{\left|\frac{a_{23}}{a_{21} \cdot a_{32}}\right|};$$
$$T_{44} = \sqrt{\left|\frac{a_{23} \cdot a_{34}}{a_{21} \cdot a_{32} \cdot a_{43}}\right|}; T_{55} = \sqrt{\left|\frac{a_{23} \cdot a_{34} \cdot a_{45}}{a_{21} \cdot a_{32} \cdot a_{43} \cdot a_{54}}\right|} \tag{39}$$

For nth order system the trace of transformation matrix T_{nn} is

$$for\ n > 2\ T_{nn} = \sqrt{\left| \frac{\prod_{i=2}^{n-1} a_{i,i+1}}{\prod_{i=2}^{n} a_{i,i-1}} \right|}; a_{ij} \neq 0; i \neq j \tag{40}$$

4 Modeling Based on Modified Tellegen's Theorem

Different types of dynamic systems, linear or nonlinear, e.g. mechanical, liquid and thermal etc. can be described by state space equations or differential equations. The equations can be transformed into a different shape and, moreover, in the form of electrical systems that can be well solved and simulated. In addition to the usual solution, e.g. the time course of the state variable, it is possible to use the approach described here to determine or estimate the power and energy in the system. The modified Tellegen's theorem is used for power and energy estimation of differential equations. For system modeling it is possible use 2 types of electrical ladder structure models: Model which can describe system with complex conjugate poles or system with only real poles M1 (RLC structures, oscillating systems) and model which is used for systems with only real poles M2 (RC or RL structures, non oscillating systems), circuit diagram booth of the models are displayed in Figs. 10, 11 and 12 [15–19].

It would seem unnecessary to use 2 models when the first type can be used for circuits with both for complex and also for real poles, but the second model is

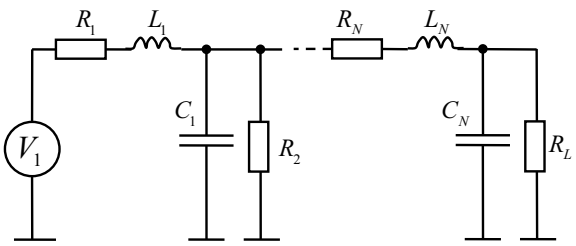


Fig. 10 The electrical model which is possible use for modeling systems with complex conjugate poles and also only real poles—M1

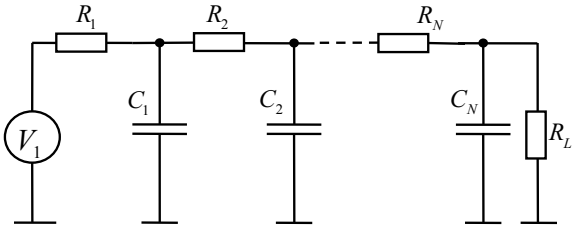


Fig. 11 The electrical model which is possible use for modeling systems with real poles only—M2-C

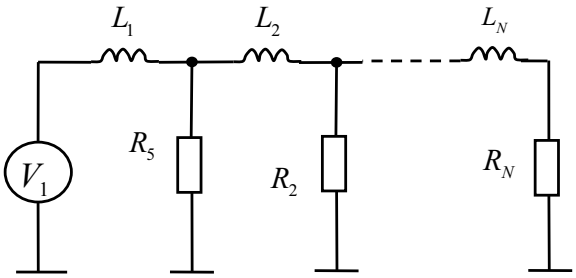


Fig. 12 The electrical model which is possible use for modeling systems with real poles only—M2-L

justified in the case of thermal systems that often leads to a three-diagonal matrix or in symmetrical three-diagonal matrix.

The state space matrices for model M1 is Eq. (41), for model M2-C it is Eq. (49). Equation for model M2-L is similar, but in this work only M2 (see Fig. 11) is used below. Models are demonstrated on 4th order circuits, but can be easily extended to lower or higher order. The state space equations for model M1:

$$\begin{bmatrix} \frac{di_{L1}}{dt} \\ \frac{dv_{C1}}{dt} \\ \frac{di_{L2}}{dt} \\ \frac{dv_{C2}}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{R_1}{L_1} & -\frac{1}{L_1} & 0 & 0 \\ \frac{1}{C_1} & -\frac{1}{C_1 R_2} & -\frac{1}{C_1} & 0 \\ 0 & \frac{1}{L_2} & -\frac{R_3}{L_2} & -\frac{1}{L_2} \\ 0 & 0 & \frac{1}{C_2} & -\frac{1}{C_2 R_4} \end{bmatrix} \cdot \begin{bmatrix} i_{L1} \\ v_{C1} \\ i_{L2} \\ v_{C2} \end{bmatrix} + \begin{bmatrix} \frac{1}{L_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} V_1 \tag{41}$$

therefore rewritten as some general differential 4th order system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -a_{11} & -a_{12} & 0 & 0 \\ a_{21} & -a_{22} & -a_{23} & 0 \\ 0 & a_{32} & -a_{33} & -a_{34} \\ 0 & 0 & a_{43} & -a_{44} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u \tag{42}$$

where

$$\begin{aligned} a_{11} &= \frac{R_1}{L_1}; a_{12} = \frac{1}{L_1}; a_{21} = a_{23} = \frac{1}{C_1}; a_{22} = \frac{1}{C_1 R_2}; \\ a_{32} &= a_{34} = \frac{1}{L_2}; a_{33} = \frac{R_3}{L_2}; a_{43} = \frac{1}{C_2}; a_{44} = \frac{1}{C_2 R_4}; \end{aligned} \tag{43}$$

with following property (for linear systems).

$$\sum_{j=1}^{n} a_{ij} = 0; i = 2, 3 \ldots n-1; i \neq j \tag{44}$$

where matrix F of energy storage elements for M1 is

$$F = \begin{bmatrix} L_1 & 0 & 0 & 0 \\ 0 & C_1 & 0 & 0 \\ 0 & 0 & L_2 & 0 \\ 0 & 0 & 0 & C_2 \end{bmatrix} = \begin{bmatrix} F_{11} & 0 & 0 & 0 \\ 0 & F_{22} & 0 & 0 \\ 0 & 0 & F_{33} & 0 \\ 0 & 0 & 0 & F_{44} \end{bmatrix} \tag{45}$$

Dissipation power P_M for model M1 is calculated as

$$\begin{aligned} P_M = \left\langle (Fx)^T, \frac{dx}{dt} \right\rangle &= L_1 x_1 \left(-\frac{R_1 x_1}{L_1} - \frac{x_2}{L_1} + \frac{V_1}{L_1} \right) \\ &+ C_1 x_2 \left(\frac{x_1}{C_1} - \frac{x_2}{C_1 R_2} - \frac{x_3}{C_1} \right) + L_2 x_3 \left(\frac{x_2}{L_2} - \frac{R_3 x_3}{L_2} + \frac{x_4}{L_2} \right) \\ &+ C_2 x_4 \left(\frac{x_3}{C_2} - \frac{x_4}{C_2 R_4} \right) \end{aligned} \tag{46}$$

and result after some manipulation is

$$\begin{aligned} P_M &= -R_1 x_1^2 + V_1 x_1 - \frac{x_2^2}{R_2} - R_3 x_3^2 - \frac{x_4^2}{R_4} \\ &= \underbrace{V_1 x_1}_{Pi} \underbrace{-R_1 x_1^2 - \frac{x_2^2}{R_2} - R_3 x_3^2 - \frac{x_4^2}{R_4}}_{Pd} \end{aligned} \tag{47}$$

where P_i is input power and P_d is dissipated power on resistors.

The expression for energy of state space variables is

$$\begin{aligned} E_M(t) &= \frac{1}{2} \langle (F \cdot \boldsymbol{x}(t))^T, \boldsymbol{x}(t) \rangle = \frac{1}{2} \left(F_{11} x_1^2 + F_{22} x_2^2 + F_{33} x_3^2 + F_{44} x_4^2 \right) \\ &= \frac{1}{2} \left(L_1 i_{L1}^2 + C_1 v_{C1}^2 + L_2 i_{L2}^2 + C_2 v_{C2}^2 \right) \end{aligned} \tag{48}$$

For model M2 relationships can be described by

$$\begin{bmatrix} \frac{dv_{C1}}{dt} \\ \frac{dv_{C2}}{dt} \\ \frac{dv_{C3}}{dt} \\ \frac{dv_{C4}}{dt} \end{bmatrix} = \begin{bmatrix} \frac{-1}{C_1}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) & \frac{1}{C_1 R_2} & 0 & 0 \\ \frac{1}{C_2 R_2} & \frac{-1}{C_2}\left(\frac{1}{R_2} + \frac{1}{R_3}\right) & \frac{1}{C_2 R_3} & 0 \\ 0 & \frac{1}{C_3 R_3} & \frac{-1}{C_3}\left(\frac{1}{R_3} + \frac{1}{R_4}\right) & \frac{1}{C_3 R_4} \\ 0 & 0 & \frac{1}{C_4 R_4} & \frac{-1}{C_4}\left(\frac{1}{R_4} + \frac{1}{R_5}\right) \end{bmatrix} \cdot \begin{bmatrix} v_{C1} \\ v_{C2} \\ v_{C3} \\ v_{C4} \end{bmatrix} + \begin{bmatrix} \frac{1}{C_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} V_1 \tag{49}$$

therefore modified as some general differential 4th order system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -a_{11} & a_{12} & 0 & 0 \\ a_{21} & -a_{22} & a_{23} & 0 \\ 0 & a_{32} & -a_{33} & a_{34} \\ 0 & 0 & a_{43} & -a_{44} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot u \tag{50}$$

where

$$a_{11} = \frac{1}{C_1}\left(\frac{1}{R_1} + \frac{1}{R_2}\right); a_{12} = \frac{1}{C_1 R_2}; a_{21} = \frac{1}{C_2 R_2}; a_{22} = \frac{1}{C_2}\left(\frac{1}{R_2} + \frac{1}{R_3}\right) = a_{21} + a_{23};$$
$$a_{23} = \frac{1}{C_2 R_3}; a_{32} = \frac{1}{C_3 R_3}; a_{33} = \frac{1}{C_3}\left(\frac{1}{R_3} + \frac{1}{R_4}\right) = a_{32} + a_{34}; a_{34} = \frac{1}{C_3 R_4};$$
$$a_{43} = \frac{1}{C_4 R_4}; a_{44} = \frac{1}{C_4}\left(\frac{1}{R_4} + \frac{1}{R_5}\right) \tag{51}$$

with property (for linear systems)

$$\sum_{j=1}^{n} a_{ij} = 0; i = 2, 3 \ldots n-1 \tag{52}$$

In the next part are given examples for calculation of dissipated power using previous theoretical derivations.

5 Examples of Power and Energy Calculation Based on Modified Tellegen's Theorem

In this section there are examples to calculate dissipated power or energy using modified Tellegen theorem. It is important to note that the calculated power is related to dissipated power and for energy according to (30) is concern only the energy of state variables.

Example 2 Differential equation which describe some dynamical system is

$$x^{(4)} + 1.5x^{(3)} + 3x^{(2)} + 2\dot{x} + x = 0 \tag{53}$$

Characteristic polynomial is

$$\lambda^4 + 1.5\lambda^3 + 3\lambda^2 + 2\lambda + 1 = 0 \tag{54}$$

with roots of the polynomial: $-0.3178 + 1.3537$i; $-0.3178 - 1.3537$i; $-0.4322 + 0.5748$i; $-0.4322 - 0.5748$i. Because of the roots are complex, it is necessary to use the M1 model.

The following matrix can be derived (with 1 dissipative element), which has the same characteristic polynomial and meets the requirements of the M1 model:

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \\ \frac{dx_4}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{d_1}{q_1} & -\frac{1}{q_1} & 0 & 0 \\ \frac{1}{q_2} & 0 & -\frac{1}{q_2} & 0 \\ 0 & \frac{1}{q_3} & 0 & -\frac{1}{q_3} \\ 0 & 0 & \frac{1}{q_4} & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} \frac{1}{q_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} V_1 \tag{55}$$

where d_i are dissipative elements and q_i are energy storage elements. The model M1 for differential Eq. (53) is Eq. (55) (matrix in Eq. (56) has the same characteristic polynomial as Eq. (53)):

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -1.5 & -0.384 & 0 & 0 \\ 4.342 & 0 & -4.342 & 0 \\ 0 & 0.169 & 0 & -0.169 \\ 0 & 0 & 3.553 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0.384 \\ 0 \\ 0 \\ 0 \end{bmatrix} V_1 \tag{56}$$

From the values of a_{ij}, the values of the model elements R, L, C can be calculated using the following relations

$$a_{11} = \frac{R_1}{L_1} = R_1 a_{12} = 1.5; a_{12} = \frac{1}{L_1} = 0.384; a_{21} = a_{23} = \frac{1}{C_1} = 4.342;$$
$$a_{22} = 0; a_{32} = a_{34} = \frac{1}{L_2} = 0.169; a_{33} = 0; a_{43} = \frac{1}{C_2} = 3.553; a_{44} = 0; \tag{57}$$

therefore part values are $L_1 = 2.6$; $R_1 = 0.577$; $C_1 = 0.23$; $L_2 = 5.92$; $C_2 = 0.28$. Results of time responses of 4 state space variables, dissipated power and energy (system run only from initial condition v_{C1}= 10 V) are shown in Figs. 13, 14 and 15.

The previous example was solved for only one dissipation element. In the following example, two dissipation elements are used, with the second element requiring a predetermined dissipation value of 0.5. Thus value of a_{44} is 0.5. The state space equations are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -1.0 & -0.521 & 0 & 0 \\ 2.878 & 0 & -2.878 & 0 \\ 0 & 0.174 & 0 & -0.174 \\ 0 & 0 & 2.878 & -0.5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0.521 \\ 0 \\ 0 \\ 0 \end{bmatrix} V_1 \tag{58}$$

For Eq. (58) $R\,L\,C$ part of electrical system values are $L_1 = 1.92$; $R_1 = 1.92$; $C_1 = 0.3475$; $L_2 = 5.747$; $C_2 = 0.3475$; $R_2 = 5.752$. Results of time responses of 4 state space variables, dissipated power and energy are shown in Figs. 16, 17 and 18.

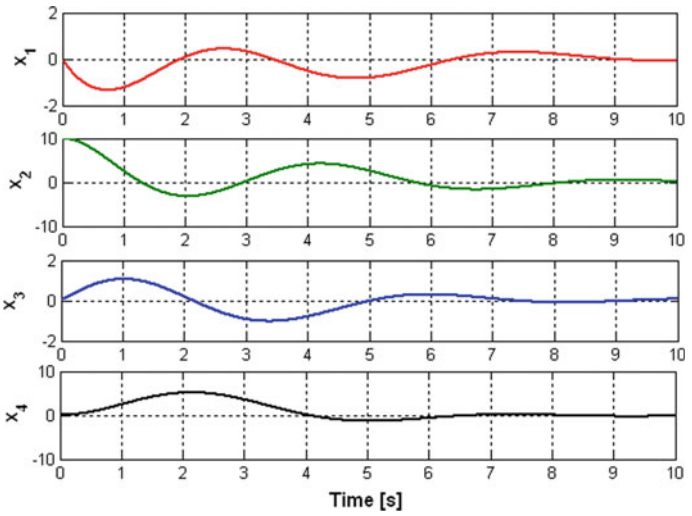


Fig. 13 The time response of state space variables of Eq. (56) running from initial condition $v_{C1}=$ 10 V

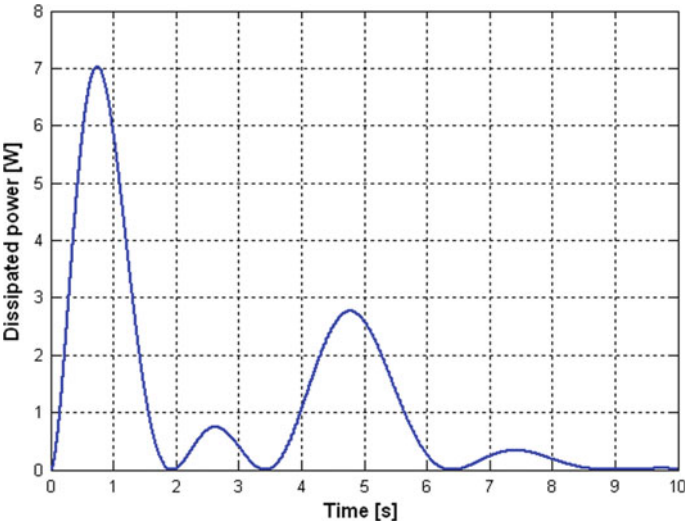


Fig. 14 The time evaluation of dissipated power of Eq. (56) running from initial condition $v_{C1}=$ 10 V

In Fig. 17 the time evolution of power is shown. There are the dissipated power on resistors (positive), power from source (negative) and sum of booth power which is zero, because after transition time input power + output power = 0.

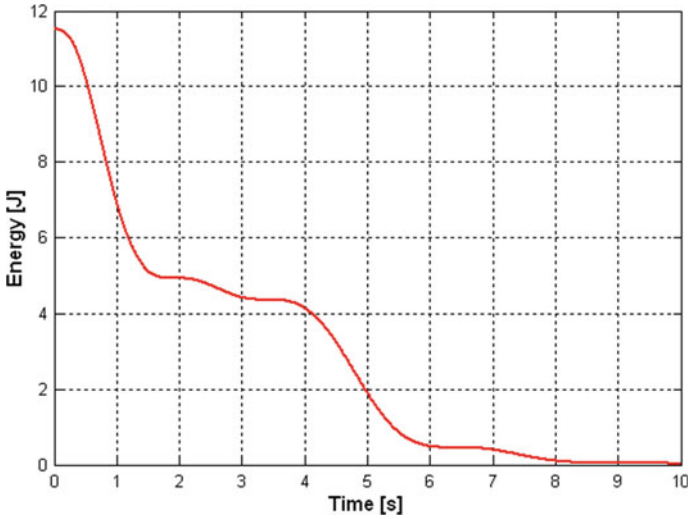


Fig. 15 The time evaluation of state space energy of Eq. (56) running from initial condition $v_{C1}=$ 10 V

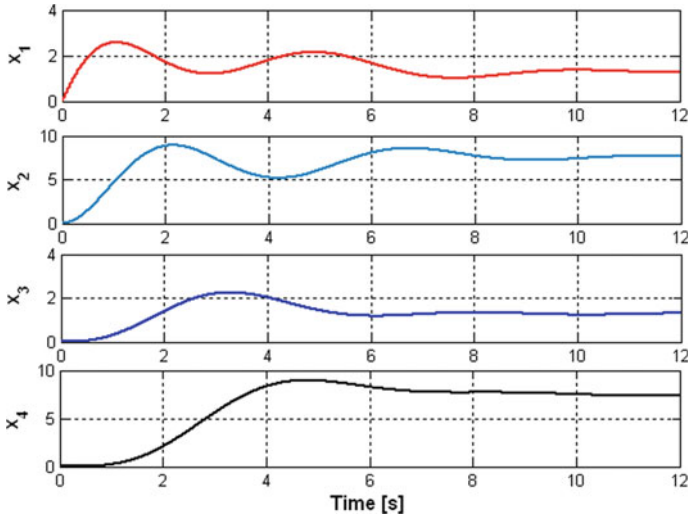


Fig. 16 The time response of state space variables of Eq. (58) running from source $V_1=$ 10 V and zero initial condition

Example 3 Consider following differential equation:

$$x^{(4)} + 8x^{(3)} + 20x^{(2)} + 16\dot{x} + 2x = 0 \tag{59}$$

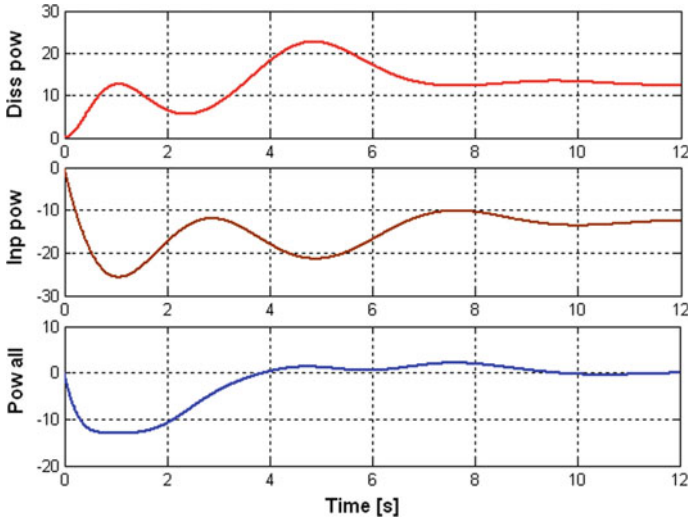


Fig. 17 The time evaluation of power of Eq. (58) running from source $V_1 = 10$ V and zero initial condition. From top to bottom—dissipated power on resistors, power of input source, total power

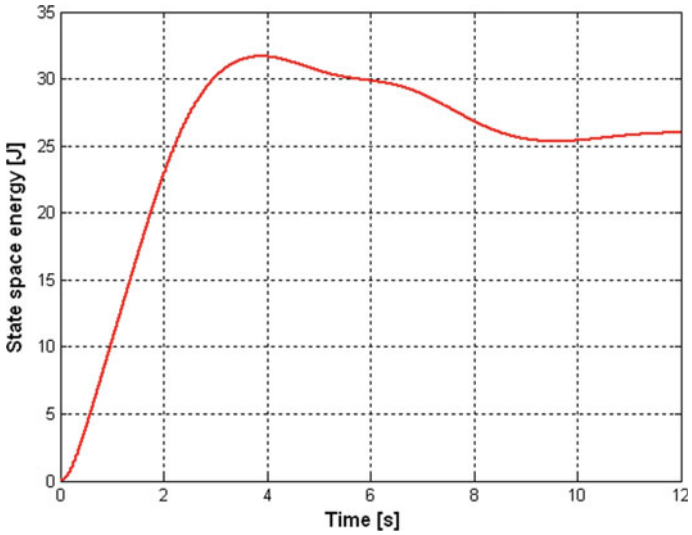


Fig. 18 The time evaluation of state space energy of Eq. (58) running from source $V_1 = 10$ V and zero initial condition

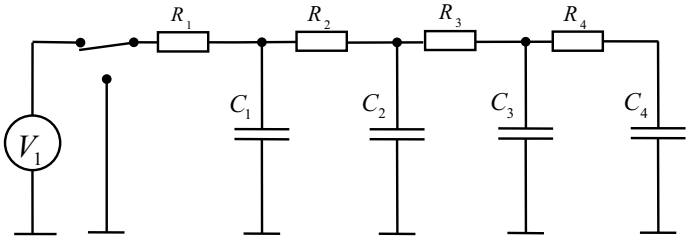


Fig. 19 The circuit diagram of resistor-capacitor ladder structure—switched for charging and discharging

Characteristic polynomial is

$$\lambda^4 + 8\lambda^3 + 20\lambda^2 + 16\lambda + 2 = 0 \tag{60}$$

with roots of the polynomial: -3.848; -2.765; -1.235; -0.152. Because of the all roots are real, it is possible to use the M2 model. The following state space system can be derived:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -2.55 & 0.552 & 0 & 0 \\ 0.59 & -1.18 & 0.59 & 0 \\ 0 & 0.53 & -1.053 & 0.53 \\ 0 & 0 & 3.21 & -3.21 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0.552 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot V_1 \tag{61}$$

It is possible find values of resistors and capacitors from Eqs. (51) and (61) for chosen $C_1 = 1.5$ (C_1 is parameter). Other calculated values are: $R_1 = 0.3333$; $R_2 = R_3 = R_4 = 1.2082$; $C_2 = 1.4005$; $C_3 = 1.5716$; $C_4 = 0.2576$; Electrical circuit is shown in Fig. 19. This circuit will solve for charging—discharging.

$$F = \begin{bmatrix} C_1 & 0 & 0 & 0 \\ 0 & C_2 & 0 & 0 \\ 0 & 0 & C_3 & 0 \\ 0 & 0 & 0 & C_4 \end{bmatrix} = \begin{bmatrix} 1.5 & 0 & 0 & 0 \\ 0 & 1.4 & 0 & 0 \\ 0 & 0 & 1.57 & 0 \\ 0 & 0 & 0 & 0.258 \end{bmatrix} \tag{62}$$

therefore according M2 (49), (61) and (62) P_M is

$$\begin{aligned} P_M &= \left\langle (F \cdot x)^T, \frac{dx}{dt} \right\rangle \\ &= C_1 \cdot x_1 \cdot \frac{dx_1}{dt} + C_2 \cdot x_2 \cdot \frac{dx_2}{dt} + C_3 \cdot x_3 \cdot \frac{dx_3}{dt} + C_4 \cdot x_4 \cdot \frac{dx_4}{dt} \end{aligned} \tag{63}$$

after adjustments (63) according to the Eq. (49)

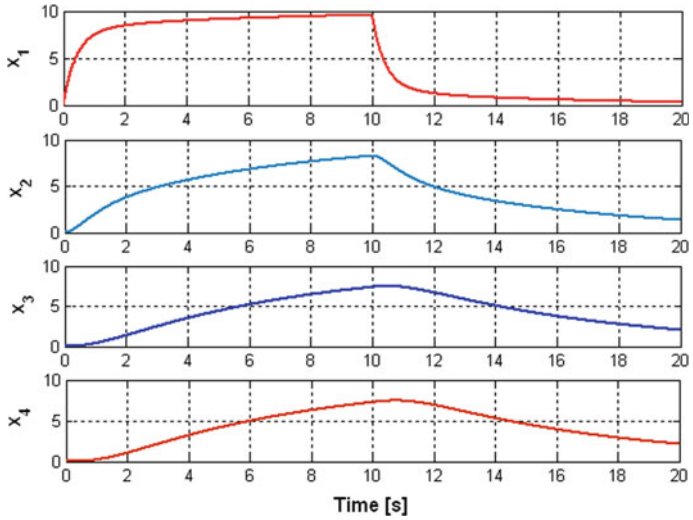


Fig. 20 The time evaluation of state space variables of Eq. (61), charging and discharging

$$P_M = V_1 \frac{V_1 - v_{C1}}{R_1} - \frac{(V_1 - v_{C1})^2}{R_1} - \frac{(v_{C1} - v_{C2})^2}{R_2} - \frac{(v_{C2} - v_{C3})^2}{R_3} - \frac{(v_{C3} - v_{C4})^2}{R_4} \tag{64}$$

Energy of the state space variables is calculated according Eq. (30).

In Fig. 20 the time evolutions of state space variables are shown. From *time* = 0 until 10, the circuit is charged from power supply $V_1 = 10$ V. After (*time* > 10), the switch, connect R_1 to ground and circuit is discharged. Power and energy lost on resistors (R_1, R_2, R_3, R_4) is shown in Fig. 21.

Notice: Energy lost on resistors was calculated as integral of power.

State space energy, i.e. energy stored in capacitors (C_1, C_2, C_3, C_4) is shown in Fig. 22.

Example 4 The second order system with nonlinear resistor, see Fig. 23. Differential equations for this system are

$$\begin{aligned} \frac{dx_1}{dt} &= -\frac{\overbrace{-\left(W - x_2^2\right)}^{R_N}}{L} x_1 + \frac{1}{L} x_2 = -\frac{-\left(3 - x_2^2\right)}{0.5} x_1 + \frac{1}{0.5} x_2 \\ \frac{dx_2}{dt} &= -\frac{1}{C} x_1 - \frac{1}{CR} x_2 = -\frac{1}{0.2} x_1 - \frac{1}{0.2 \cdot 5} x_2 \end{aligned} \tag{65}$$

where $L = 0.5$; $C = 0.2$; $R = 5$; R_N is nonlinear resistor and $W = 3$ (W is prescribed value). The simulation starts from initial condition $x_2 = v_C = 1.412$.

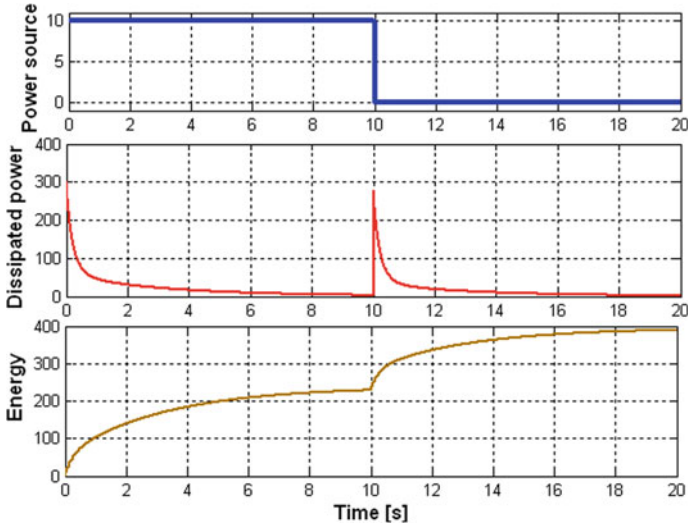


Fig. 21 The power and energy lost on resistors during the charging and discharging. Top—power supply voltage, middle—dissipated power, bottom—dissipated energy for one cycle charging and discharging

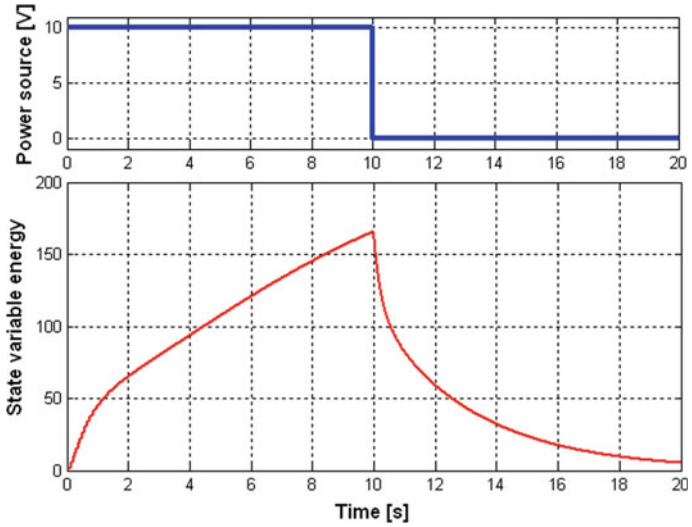


Fig. 22 The time evolution of state space energy during the charging and discharging. Top—power supply voltage, bottom—energy on capacitors for one cycle charging and discharging

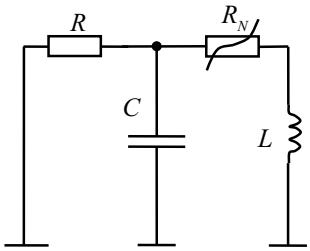


Fig. 23 The circuit diagram of nonlinear system, $L = 0.5$; $C = 0.2$; $R = 5$, R_N nonlinear resistor

$$R_N = -\left(W - x_2^2\right) \tag{66}$$

Nonlinear resistor R_N is positive or negative according

$$\begin{aligned} x_2^2 > W &\Rightarrow positive\,resistance \\ x_2^2 < W &\Rightarrow negative\,resistance \end{aligned} \tag{67}$$

When the value of x_2 is decreasing (smaller then desired value) the R_N is negative, therefore system according Eq. (65) is anti-dissipative therefore value of x_1 and x_2 is increasing, whereas if x_2 is higher then desired value W, R_N is positive, system is dissipative and values of x_1 and x_2 are decreasing. Equation (65) can be rewritten as

$$\begin{aligned} \frac{dx_1}{dt} &= -\frac{-\left(W - x_2^2\right)}{L} x_1 + \frac{1}{L} x_2 = -a_{11} x_1 + a_{12} x_2 \\ \frac{dx_2}{dt} &= -\frac{1}{C} x_1 - \frac{1}{CR} x_2 = -a_{21} x_1 - a_{22} x_2 \end{aligned} \tag{68}$$

Energy of system is

$$E_M = \frac{1}{2}\left(L \cdot x_1^2 + C \cdot x_2^2\right) \tag{69}$$

Power is

$$P_M = \frac{-\left(W - x_2^2\right)}{a_{12}} \cdot x_1^2 + \frac{a_{22}}{a_{21}} x_2^2 \tag{70}$$

Results of simulations are shown in Figs. 24, 25, 26 and 27 using the implementations according Figs. 28, 29 and 30 (in the appendix).

In Fig. 24 the time evolution of state variables are shown. Thanks to positive/negative nonlinear resistance R_N the system oscillates. Amplitude of signal x_2 is approx. equal of W, therefore 3 ($W = 3$). In Fig. 25 the energy of state space variables and total energy is displayed. The time evolution of nonlinear resistance R_N and power lost on resistors R_N and R respectively is shown in Figs. 26 and 27.

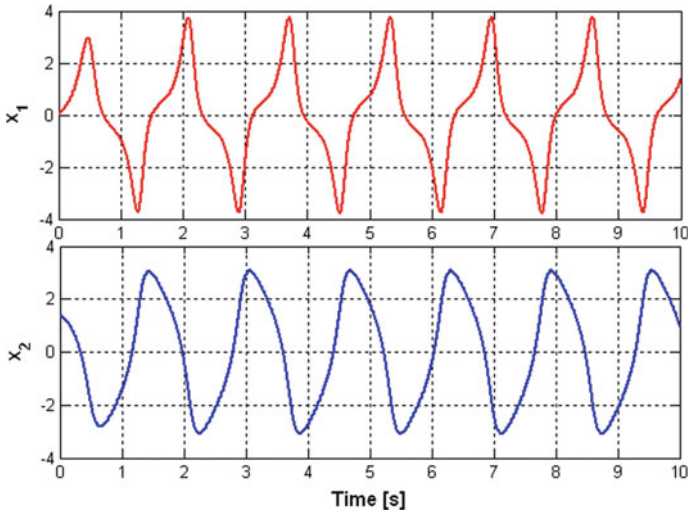


Fig. 24 The time evolution of state space variables of nonlinear system. Top—$x_1(t)$, bottom—$x_2(t)$

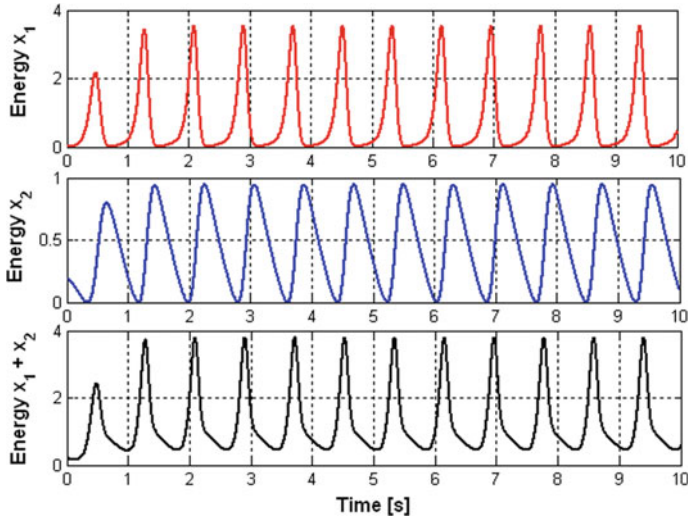


Fig. 25 The time evolution of state space energy of nonlinear system. Top—energy of $x_1(t)$, middle—energy of $x_2(t)$, bottom—total energy

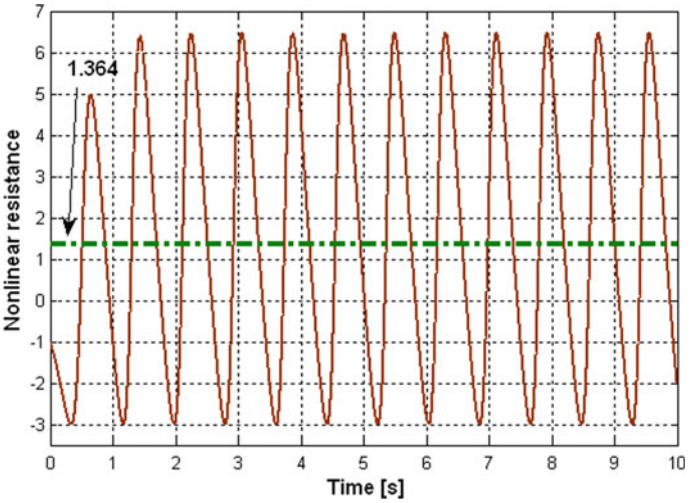


Fig. 26 The time evolution of nonlinear resistance. Mean value is 1.364

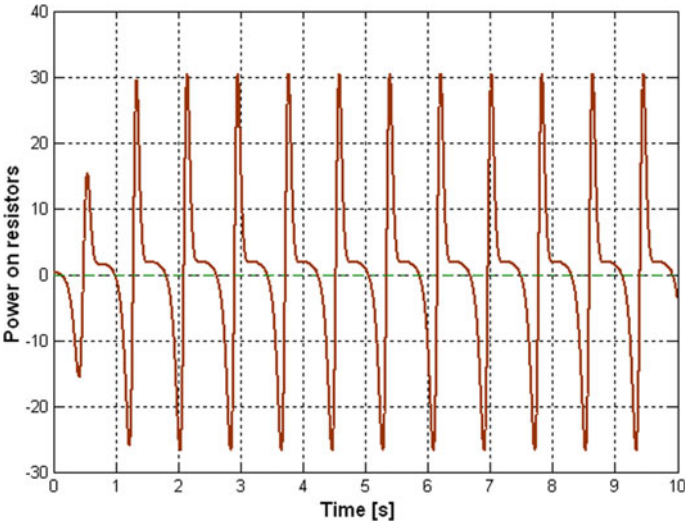


Fig. 27 The time evolution of power on resistors. Mean value -0.0368

6 Conclusions

In this chapter, a new approach to power and energy calculation has been presented in systems described by equations of state or differential equations. It was shown that 2 types of models can be used for linear differential equations, the model for oscillating systems and the model for thermal systems. The one example of nonlinear system was also solved. The theoretical derivations were confirmed by the results obtained by solving the examples. Solutions of the examples were confirmed as differential equations in MATLAB, as models in MATLAB_Simulink and also as circuits solved in SPICE. The main advantage is that presented approach can be used for different types of linear or nonlinear physical systems.

Acknowledgements This work was supported by Department of Electronics and Information Technology, University of West Bohemia, Plzen, Czech Republic and by the Ministry of Education, Youth and Sports of the Czech Republic under the project OP VVV Electrical Engineering Technologies with High-Level of Embedded Intelligence, CZ.02.1.01/0.0/0.0/18_069/0009855 and by the Internal Grant Agency of University of West Bohemia in Plzen, the project SGS-2018-001.

Appendix

This appendix presents another look on nonlinear circuit described by Fig. 23. Such circuit can be (according (65)) drawn as controlled feedback system. The result is shown in Fig. 28. In the Fig. 29 the system from Fig. 28 is redrawn as MATLAB_Simulink model. In Fig. 30 the MATLAB_Simulink model is extended for calculation of power dissipated on resistors.

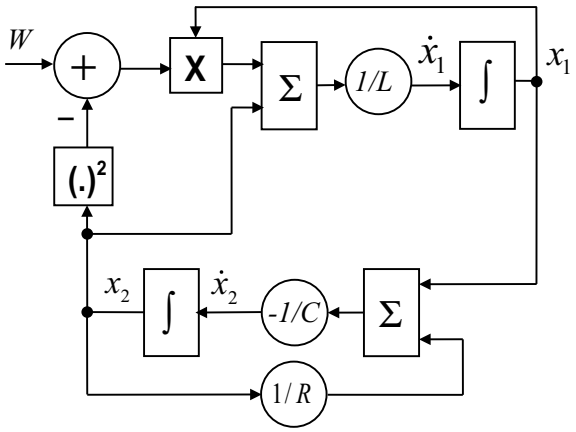


Fig. 28 The feedback model of circuit with nonlinear resistor (see Fig. 23), $L = 0.5$; $C = 0.2$; $R = 5$ and W is prescribed value

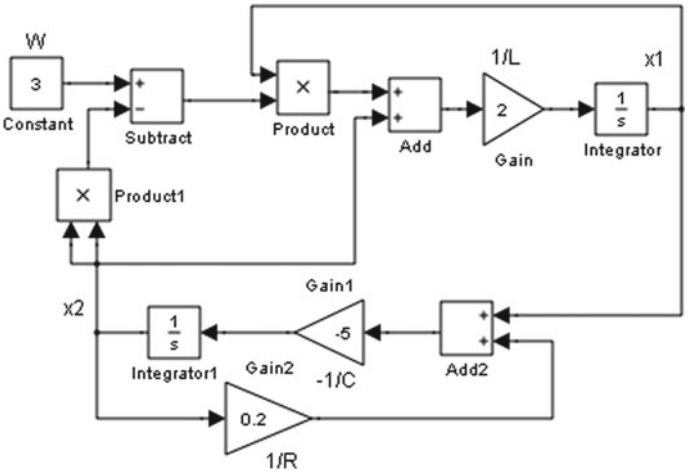


Fig. 29 The MATLAB_Simulink model of nonlinear circuit (Fig. 23) derived from system shown in Fig. 28. Prescribed value $W = 3$

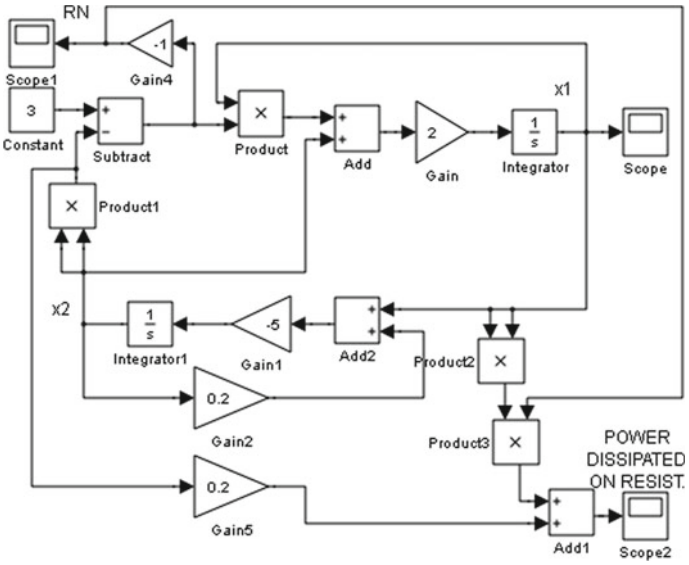


Fig. 30 The MATLAB_Simulink model of nonlinear circuit (Fig. 23) with calculation of power dissipated on resistors R_N and R

References

1. Kailath T (1980) Linear systems. Prentice Hall, Upper Saddle River, NJ
2. Khalil HK (2002) Nonlinear systems, 3rd edn. Prentice-Hall, Upper Saddle River, NJ
3. Ruehli AE, Johnson TA (1999) Circuit analysis computing by waveform relaxation, Wiley, New York, vol 3
4. Ruehli AE (1987) Circuit analysis, simulation and design, New York: Elseviers
5. Lewis FL (1992) Applied optimal control and estimation. Prentice-Hall, Englewood-Cliffs, NJ
6. Ogata K (2002) Modern control engineering, 4th edn. Prentice Hall, Upper Saddle River, NJ
7. Hrusak J, Mayer D, Stork M (2006) New approach to non-linear instability and chaos Based on generalized tellegen's principle, WMSCI Orlando, Florida, International Institute of Informatics and Systemics, USA, ISBN 980-6560-67-1, 199–206
8. Hrusak J, Mayer D, Stork M (2006) Dissipativity, minimality and physical correctness of state space representations, CITSA 2006. Orlando, Florida, USA, International Institute of Informatics and Systemics, ISBN 980-6560-83-3, 136–141
9. Stork M, Hrusak J, Mayer D (2006) Chaos and strange behavior detection of continuous and digital nonlinear systems based on internal system energy autocovariance, WMSCI Orlando, Florida, International Institute of Informatics and Systemics, USA, 228–233
10. Mayer D, Hrusak J, Stork M (2013) On state-space energy based generalization of Brayton–Moser topological approach to electrical network decomposition. Computing, Springer. https://doi.org./10.1007/s00607-012-0280-2
11. Van der Schaft AJ, Maschke BM (1995) The hamiltonian formulation of eergy conserving physical systems with external ports. Archiv fuer Elektronik und Uebertragungstechnik 49(5/6):362–371
12. Williams RL, Douglas A (2007) Linear state-space control systems, Wiley, Inc. ISBN: 978-0-471-73555-7
13. Zhou K (1995) Robust and optimal control. Prentice-Hall, Upper Saddle River, NJ
14. Jordan DW, Smith P (2007) Nonlinear ordinary differential equations, an introduction for scientists and engineers, 4th edn. Oxford University Press, ISBN 978–0–19–920824–1
15. Yan YW, Chung KL (1994) Fast algorithm for solving special tridiagonal systems. Computing, vol 52, pp 203–211
16. Al-Khaleel M, Gander M, Ruehli J (2014) A mathematical analysis of optimized waveform relaxation for a small RC circuit. Appl Numeri Math 75(2014):61–76
17. Gander MJ, Ruehli AE (2004) Optimized waveform relaxation methods for RC type circuits, IEEE Trans Circuit Syst—I: Regular papers 51(4)
18. Langtangen HP, Pedersen GK (2016) Scaling of differential equations, Springer, ISBN 978-3-319-32725-9
19. Shampine LF, Gladwell I, Thompson S (2003) Solving ODEs with MATLAB, Cambridge University Press

Self-tuning Yaw Control Strategy of a Horizontal Axis Wind Turbine Based on Machine Learning

Aitor Saenz-Aguirre, Ekaitz Zulueta, Unai Fernandez-Gamiz, Jose Antonio Ramos-Hernanz, and Jose Manuel Lopez-Guede

Abstract The design procedure of a Machine Learning (ML) based yaw control strategy for a Horizontal Axis Wind Turbine (HAWT) is presented in the following chapter. The proposed yaw control strategy is based on the interaction of three different Artificial Intelligence (AI) techniques to design a ML system: Reinforcement Learning (RL), Artificial Neural Networks (ANN) and metaheuristic optimization algorithms. The objective of the designed control strategy is to achieve, after a training stage, a fully autonomous performance of the wind turbine yaw control system for different input wind scenarios while optimizing the electrical power generated by the wind turbine and the mechanical loads due to the yaw rotation. The RL algorithm is known to be able to learn from experience. The training process could be carried out online with real-time data of the operation of the wind turbine or offline, with simulation data. The use of an ANN to store the data of the matrix $Q(s, a)$ related to the RL algorithm eliminates the large scale data management and simplifies the operation of the proposed control system. Finally, the implementation

A. Saenz-Aguirre (✉)
Nuclear Engineering and Fluid Mechanics Department, University of the Basque Country, Eibar, Spain
e-mail: aitor.saenz@ehu.eus

E. Zulueta · J. M. Lopez-Guede
Automatic Control and System Engineering Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain
e-mail: ekaitz.zulueta@ehu.eus

J. M. Lopez-Guede
e-mail: jm.lopez@ehu.eus

U. Fernandez-Gamiz
Nuclear Engineering and Fluid Mechanics Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain
e-mail: unai.fernandez@ehu.eus

J. A. Ramos-Hernanz
Electrical Engineering Department, University of the Basque Country, Vitoria-Gasteiz, Araba, Spain
e-mail: josean.ramos@ehu.eus

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_32

of a metaheuristic optimization algorithm, in this case a Particle Swarm Optimization (PSO) algorithm, allows calculation of the optimal yaw control action that responds to the compromise between the generated power increment and the mechanical loads increase due to the yaw actuation.

Keywords Wind turbine control · Yaw control · Reinforcement learning · Artificial neural network · Optimization · Pareto front

Abbreviations and Acronyms

ML	Machine Learning
HAWT	Horizontal Axis Wind Turbine
AI	Artificial Intelligence
RL	Reinforcement Learning
ANN	Artificial Neural Network
PSO	Particle Swarm Optimization
LCOE	Levelized Cost of Energy
MLP-BP	MultiLayer Perceptron with Back Propagation
MDP	Markov Decision Process
DP	Dynamic Programming
MC	Monte Carlo
TD	Temporal Differences
PoF	Pareto optimal Front
PID	Proportional Integral Derivative
FAST	Fatigue, Aerodynamics, Structure and Turbulence
NREL	National Renewable Energies Laboratory
MSE	Mean Squared Error
DM	Decision Making

Nomenclature

θ_{wind}	Direction of the wind
$\theta_{nacelle}$	Orientation of the nacelle
θ_{yaw}	Yaw angle
Ω_{yaw}	Yaw rotational speed
s	State of the RL algorithm
a	Action of the RL algorithm
r	Immediate reward of the RL algorithm
γ	Discount factor
$Q(s, a)$	Expected long-term reward matrix in RL algorithm
$Q\_P(s, a)$	Expected long-term power gain reward matrix in RL algorithm

$Q\_M(s, a)$	Expected long-term mechanical moment reward matrix in RL algorithm
$Q\_P(s(t), a(t))$	Expected long-term power gain reward function in RL algorithm
$Q\_M(s(t), a(t))$	Expected long-term mechanical moment reward function in RL algorithm

1 Introduction

The gradual depletion of the fossil fuels and the atmospheric pollution originated by their combustion have brought an important growth of the renewable energy generation systems. Nowadays, the most important renewable energy generation source is the wind energy. Many studies showing the positive tendency of the wind energy can be found in the literature. For example, according to some studies presented by Rosales-Asensio et al. [1], the sustainable power production with wind origin in Denmark achieved a 40% of the power produced in the country in 2015. This same value was quite smaller in Spain, with a 17% in 2015, but having raised from a 10.4% in 2007. More recent studies elaborated by WindEurope [2] show remarkable increments in the wind energy installed power in 2018 especially in four countries: a 29% in Germany, a 16% in the United Kingdom, a 13% in France and a 6% in Sweden.

The power generation increase in wind energy systems is tightly related to the investigation work carried out to reduce the Levelized Cost of Energy (LCOE) of the wind turbines, which encourages capital investment in the sector, as explained in the work of Nyanteh et al. [3]. One main topic of this research work is the development of advanced control strategies to optimize the performance of the wind turbines [4–9].

In this chapter, the design procedure of a yaw control system of a Horizontal Axis Wind Turbine (HAWT) based on Machine Learning (ML) is presented. The objective of the ML based control strategy developed in this chapter is to achieve a fully autonomous performance of the yaw system of the wind turbine based on its own experience, which could be acquired via an offline training, i.e., when the wind turbine is paused, or an online training, i.e., during operation of the wind turbine. An offline training process is proposed in this chapter. However, a continuous online training process with real data acquired during operation of the wind turbine to continuously learn from experience could be implemented as well. The MLP-BP is used to store the data of the matrices $Q(s, a)$ related to the RL algorithm and manage them as continuous functions, $Q(s(t), a(t))$. This process avoids quantification and large data management problems. The combination of an RL strategy and an ANN is widely known as Deep Reinforcement Learning [10, 11]. As observed in the works of Saenz-Aguirre et al. [5, 8], an increment of the power generated by the wind turbine with a considerable reduction of the mechanical loads due to the yaw rotation is expected to be achieved.

This chapter is structured as follows: the objectives and applications of the proposed yaw control strategy are presented in Sect. 2. Section 3 details the theoretical basis of the different Artificial Intelligence (AI) techniques used to design the ML system. The design procedure of the yaw control system based on ML is exposed in Sect. 4. Finally, Sect. 5 presents the conclusions.

2 Objectives and Applications

The main factor that determines the power output of a wind turbine is the wind incident to its rotor. However, the wind is originated as a result of very complex meteorological processes, which, as stated by Bivona et al. [12], are very complex to model, and can, thus, suffer from unpredictable important variations. Some wind gusts can even exceed the safe wind speed operation range of the wind turbine and endanger its correct performance. To avoid this issue, a control system is implemented in the wind turbines.

The control system of a wind turbine is formed by different strategies aimed to regulate the rotational speed of the rotor in the whole range of operating points of the wind turbine. As a result of these strategies, the power output of the wind turbine is predefined for the whole range of wind speed values in which the turbine operates. The curve that relates the power output of the wind turbine with the wind speed is known as the power curve. The power curve of the NREL 5 MW wind turbine, presented in the work of Jonkman et al. [13], is illustrated in Fig. 1.

The main control objective in the partial power zone, plotted in blue color in Fig. 1, is to maximize the power the wind turbine extracts from the wind, which can be expressed as in Eq. (1).

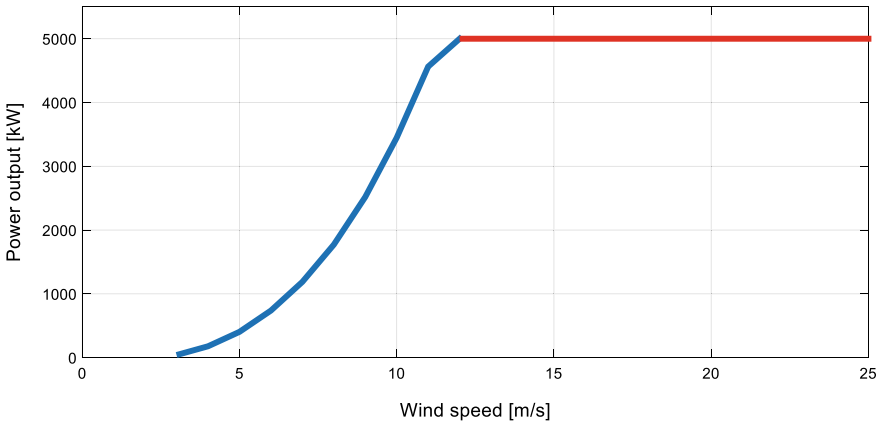


Fig. 1 Power curve of the NREL 5 MW wind turbine

$$P_{opt} = \frac{1}{2} \cdot \rho \cdot C_P \cdot A \cdot v^3 [W] \tag{1}$$

where ρ [kg/m^3] is density of the air, C_P [-] is the power coefficient, A [m^2] is the area covered by the rotor and v [m/s] is the wind speed.

However, in order to express the real power the wind turbine extracts from the wind, the misalignment between the incident wind and the rotor must be considered, commonly known as the yaw angle. The expression is shown in Eq. (2).

$$P = P_{opt} \cdot \cos^3\left(\theta_{yaw}\right)[W] \tag{2}$$

where θ_{yaw} [deg] is the yaw angle.

As it can be observed in Eq. (2), a correct alignment of the wind turbine with the direction of the incident wind can make the power generated by the wind turbine increase considerably. The control system that allows a correct alignment of the wind turbine with respect to the incident wind is the yaw control. A detailed explanation about the yaw control system of a 5 kW wind turbine is introduced in the work of Yücel and Özder [14].

On the other hand, as a result of the high inertia values of the mechanical components that participate in the yaw rotation, remarkable mechanical loads arise in different elements of the wind turbine. The physical effect that explains these loads is known as the gyroscopic effect. An study of possible control strategies aimed to attenuate the high mechanical loads resulting from the gyroscopic effect are presented in [15, 16]. Additionally, an analysis of the mechanical loads generated as a consequence of the yaw rotation is presented in the work of Shariatpanah et al. [17].

As a result, an adequate design of the yaw control strategy allows not only maximization of the power generated by the wind turbine, but also reduction of the mechanical loads in several elements of the wind turbine, and, thus, to increment its lifetime.

The objectives of the proposed yaw control strategy are:

- Achieve a fully autonomous and self-tuning yaw control strategy to be implemented in the wind turbine.
- Design a control strategy based on ML that can continuously learns from its own experience.
- Selection of the optimal yaw control action (maximal power and minimal loads possible) for every possible scenario of the wind turbine operation.

The main applications of the designed yaw control strategy are:

- Increment of the power produced by the wind turbine, with the consequent enhancement of its efficiency, and the reduction of the LCOE.
- Reduction of the mechanical loads originated as a result of the yaw rotation, with the consequent increment of the lifetime of the mechanical components of the wind turbine, and the reduction of the LCOE.

3 Machine Learning and Artificial Intelligence Techniques

The AI is the science that studies the projection of the human intelligence in technological machines. In other words, the AI is the science that analyses the possibility to develop smart behavior patterns in technological machines. Nowadays, with the technological advances in the field of the informatics and the existence of very large amounts of data to be processed, the AI is on the focus of the research work.

The field of the AI is composed by numerous different techniques, which, in general, have been developed to emulate the human intelligence or decision making capability, as it is explained in the work of Wang et al. [18]. The most important AI techniques are the RL, ANNs, Fuzzy Logic, bio-inspired or metaheuristic optimization algorithms and Bayesian Networks. Each AI technique serves to a determined goal and could be used individually or in interrelation with other AI techniques.

One of the most important features that offers the AI is the capability of the systems to learn automatically. This feature of self-learning is commonly known as ML, as it is explained in detail in the work of Fadlullah et al. [19]. The ML has undergone an important boom after the development of the ANNs, which are able to continuously learn from very large amounts of data. RL is another type of ML, in which the systems learns to make the best decisions in a given environment by using its own experience.

With the technological boom and the increasing processing capability of the processors a new learning method known as Deep Learning [19] has been born, in which new and amplified configurations of ANNs are used for the ML process. In the same way, the Deep Reinforcement Learning [19] method has also been created, which combines the use of the RL algorithm and ANNs to store the matrix $Q(s, a)$ related to the RL algorithm.

The self-tuning ML based yaw control strategy presented in this chapter makes use of three different AI techniques: RL, ANN and metaheuristic optimization algorithms. This section is structured as follows: the theoretical background of the RL is explained in Sect. 3.1. Section 3.2 analyses the theory behind the ANNs. And, finally, the theoretical basis of the optimization algorithms is introduced in the Sect. 3.3.

3.1 Reinforcement Learning

RL [10, 20–22] is an AI technique, corresponding to a type of ML, in which a determined system learns from the experience of its own interaction with the environment in which it is placed. As it is stated in the work of Jagodnik et al. [20], the training process of a RL algorithm is achieved by trial and error with the objective of maximizing a reward function defined numerically and by mapping of situations to actions.

A pipeline with the basic operating principle of a RL algorithm is presented in Fig. 2. A defined agent which is in a determined environment receives information of

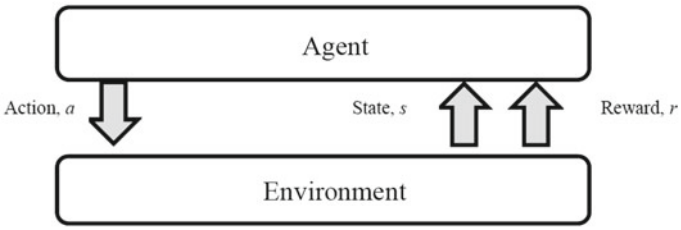


Fig. 2 Basic pipeline of a RL algorithm [5]

its state ($s \in S$) and decides to take the action ($a \in A$). As a result of this action, the agent receives information of its new state and the immediate reward of the action ($r \in R$). The objective of the RL algorithm is to find a map of states to actions, known as policy, to maximize the long-term reward in different situations. In other words, the RL controller selects the future actions with regard to the experiences of a whole range of actions in predefined states. The experiences are obtained by trial and error by interaction with a dynamic environment, as exposed in the work of Kaelbling et al. [23].

The main elements of a RL algorithm are:

- **State** ($s \in S$): Defines the state of an agent that is placed in a determined environment.
- **Action** ($a \in A$): Defines the action taken by an agent that is in a defined state ($s \in S$) in a determined environment.
- **Reward** ($r \in R$): Defines the immediate reward received by an agent that takes a certain action ($a \in A$) in a given state ($s \in S$).
- **Policy** (π): It is a mapping of the actions ($a \in A$) to the states ($s \in S$). Thus, it defines the behavior of the agent.
- **Long-term reward** (R_t): Indicates the long term reward received by the agent if a certain action ($a \in A$) in a given state ($s \in S$) is taken. The long-term reward is the value to be maximized.

The long-term reward R_t of a RL algorithm can be numerically calculated in different ways. The most widely-used expression is based on the addition of the immediate rewards ($r \in R$) received by the agent during a determined period of time and using a discount factor γ, as it is shown in Eq. (3).

$$R_t = \sum_{k=t}^{t+T} \gamma_k \cdot r_{t+k+1} \tag{3}$$

where the discount factor γ is set to $0 < \gamma < 1$.

From now on, in order to refer to the function that indicates the long-term reward R_t expected by the agent a new expression is shown in Eq. (4).

$$E\left(\sum_{k=t}^{t+T} \gamma_k \cdot r_{t+k+1}\right) \tag{4}$$

One important aspect related to the RL algorithms is that the environment in which the agent is placed is defined as a Markov Decision Process (MDP). This means that the environment transitions are independent on past states and exclusively depend on the current state ($s \in S$) and the action taken ($a \in A$). Therefore, the expressions of the state and reward transitions are presented in Eq. (5) and Eq. (6), respectively.

$$p^a_{ss'} = p\left\{s_{t+1} = s' | s_t = s, a_t = a\right\} \tag{5}$$

$$R^a_{ss'} = E\left\{r_{t+1} | s_t = s, a_t = a, s_{t+1} = s'\right\} \tag{6}$$

The policy π followed by the agent defines the mapping of actions to states and, thus, dictates the criteria to take determined actions. Hence, the policy π defines the probability to select each action ($a \in A$) in each determined state ($s \in S$). As a result, the expected long-term reward with respect to the current state ($s \in S$) and the policy π followed, known as $V^\pi(s)$, and the expected long-term reward with respect to the current state ($s \in S$), the current action ($a \in A$) and the policy π followed, known as $Q^\pi(s, a)$, can be numerically calculated as shown in Eq. (7) and Eq. (8), respectively.

$$V^\pi(s) = E_\pi\{R_t | s_t = s\} = E_\pi\left\{\sum_{k=t}^{t+T} \gamma_k \cdot r_{t+k+1} | s_t = s\right\} \tag{7}$$

$$Q^\pi(s, a) = E_\pi\{R_t | s_t = s, a_t = a\} = E_\pi\left\{\sum_{k=t}^{t+T} \gamma_k \cdot r_{t+k+1} | s_t = s, , a_t = a\right\} \tag{8}$$

The optimal values of both $V^\pi(s)$ and $Q^\pi(s, a)$ can be expressed as in Eqs. (9) and (10).

$$V(s) = \max(V^\pi(s)) \tag{9}$$

$$Q(s, a) = \max(Q^\pi(s, a)) \tag{10}$$

The objective of the RL algorithm is to find the optimal mapping of actions to states so that the value of the $Q(s, a)$ expressed in Eq. (10) is maximized for each par of state ($s \in S$) and action ($a \in A$). To that end, there are 3 different methods to solve a MDP process: Dynamic Programming (DP), Monte Carlo (MC) method and Temporal Differences (TD). In the following lines an explanation on each one of them is introduced.

– **Dynamic Programming**

The DP method, explained in detail in the works of Bertsk et al. [24–26], is based on the knowledge of a model of the environment in which the agent is placed. That means that the state transitions $p^a_{ss'}$, see Eq. (5), and the reward transitions $R^a_{ss'}$, see Eq. (6), can be calculated analytically. As a result, the value of $V^\pi(s)$ and $Q^\pi(s, a)$ can also be represented analytically using Bellman equations, as shown in Eqs. (11) and (12).

$$V^\pi(s) = \sum_a \pi(s, a) \sum_{s_{t+1}} p^a_{ss'} \cdot [R^a_{ss'} + \gamma \cdot V^\pi(s_{t+1})] \tag{11}$$

$$Q^\pi(s, a) = \sum_a \pi(s, a) \sum_{s_{t+1}} p^a_{ss'} \cdot [R^a_{ss'} + \gamma \cdot Q^\pi(s_{t+1}, a_{t+1})] \tag{12}$$

The numerically calculated values of $V^\pi(s)$ and $Q^\pi(s, a)$ are used to perform an iterative algorithm in which every action ($a \in A$) of every possible state ($s \in S$) is considered and the policies π that maximize the value of $Q(s, a)$ are to be found.

One of the biggest drawbacks of this method is the computational cost, since for the calculation of each policy π calculations related to a great number of states and actions have to be performed.

– **Monte Carlo method**

The MC method [27, 28] is based on the assumption that a model of the environment is unknown, and thus, its performance depends on the experimental data. Since the model is unknown, the values of the state transitions $p^a_{ss'}$, see Eq. (5), and the reward transitions $R^a_{ss'}$, see Eq. (6), and as a result, the values of $V^\pi(s)$ and $Q^\pi(s, a)$ cannot be analytically computed, so they are calculated as an average of the experimentally obtained reward values.

The objective is to try to calculate the value of $Q^\pi(s, a)$ for all the state-action pairs and find the policies π that maximize the value of $Q(s, a)$. To that end, usually stochastic policies that have probabilities greater than 0 to consider each state ($s \in S$) and action ($a \in A$) are implemented.

– **Temporal Differences**

The TD method is a combination of DP and MC methods having the advantages associated to each one of them. It is based on analytical calculation, like the DP method, but, like the MC method, it does not depend on a model of the environment. In this method, the calculations to continuously learn are performed between successive predictions instead of between predictions and the final value. Hence, the convergence is faster and the computational cost is remarkably reduced. The two principal TD based algorithms are Q-Learning, explained in detail in the works of Watkins et al. [29, 30], and SARSA, introduced in the work of Adam et al. [31].

The principal difference between both methods is the calculation of the values of $Q(s, a)$. In the Q-Learning algorithm the state and actions are quantified and a

matrix is obtained as a result of mapping a $Q(s, a)$ value to each state-action par. However, in SARSA, the function $Q(s, a)$ is considered as an exponential moving average continuous function.

The calculation of the $Q(s, a)$ in SARSA algorithm can be expressed as shown in Eq. (13).

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha \cdot \left[r + \gamma \cdot Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)\right] \quad (13)$$

The calculation of the $Q(s, a)$ in Q-Learning algorithm can be expressed as shown in Eq. (14).

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha \cdot \left[r + \gamma \cdot max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)\right] \quad (14)$$

3.2 Artificial Neural Networks

ANNs correspond to a branch of the AI intended to mimic the performance of a biological brain. Biological brains are composed by millions of neurons distributed in layers and widely interconnected between them. Through these interactions between neurons the information flow from one neuron to another occurs. Furthermore, the information flow happens always in one direction, which can be either forwards or backwards. ANNs [32–34], which try to emulate this behavior, are digital systems with a variable number of neurons distributed in a structure similar to that of biological networks and with a similar functionality.

According to the work of Yang [35], the first standard artificial neuron design was introduced by W. McCulloch and W. Pitts in 1943 and, after that, they have undergone an important development. Nowadays they are very precious especially for their good performance in parallel processing, distributed memory alongside the number of neurons and the adaptability to the environment and the generalization capability.

ANNs are a compound of a variable number of neurons distributed in different ways and with a different type of interconnections. An individual neuron, shown in Fig. 3, is the smallest element of an ANN and presents the following structure:

The main elements of an artificial neuron are:

- **Inputs** (x_j): Define the inputs to the neuron.
- **Input weights** (w_j): Define the weights of each input to the neuron.
- **Propagation rule** (h_i): It defines the combination of the different inputs of the neuron before the activation function. The most common propagation rule is the linear combination of the product of each input and its weight. Moreover, usually another parameter commonly expressed as θ is added. Therefore, the propagation rule can be mathematically expressed as shown in Eq. (15).

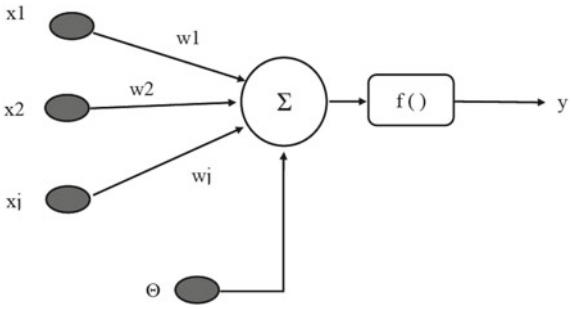


Fig. 3 Neuron of an ANN

$$h(x_1, \ldots, x_j, w_1, \ldots, w_j) = \sum_{j=1}^{n} w_j \cdot x_j - \theta \tag{15}$$

– Activation function (f_i): The activation function defines the activation state of the neuron. Additionally, it represents the output of the neuron.

If it is an on/off neuron, the activation function of the neuron can be expressed as in Eq. (16).

$$y = \begin{cases} 1 \quad if \sum_{j=1}^{n} w_j \cdot x_j \geq \theta \\ 0 \quad if \sum_{j=1}^{n} w_j \cdot x_j < \theta \end{cases} \tag{16}$$

However, when a continuous output of the neuron is desired, usually a sigmoid function [36] is used as the activation function, as shown in Eq. (17).

$$f(x) = \frac{1}{1 + e^{-\beta \cdot x}} \tag{17}$$

where the value associated to the exponential factor is $\beta > 0$.

ANNs are formed by compound of a variable number of neurons in different structures and interconnection patterns. The neurons are divided in layers., usually in a standard ANN there are 3 different neuron layers: The input layer (contains the input neurons, which are in number the same as the inputs of the ANN), the hidden layer (contains the processing neurons) and the output layer (contains the output neurons, which are in number the same as the outputs of the ANN). The number of hidden layers and the number of neurons in each hidden layer is adaptable and can be modified by the designer of the ANN.

The training algorithms of the ANNs are the responsible for making the ANN learn from its input values. There are two main ANN training method groups: The supervised learning and the unsupervised learning. As it is exposed in the work of Chen et al. [32], the supervised learning adjusts the values of the weights related

to the interconnection between neurons with the objective of minimizing the error existent between the output of the ANN and the reference output. The most important application of the supervised learning is for regressions or modelling of systems and one of the most used examples of supervised learning is the BackPropagation algorithm. The unsupervised learning does not need an output reference and the ANN is trained with numerous input patterns to explore the relation between them and categorize them. The most important application of the unsupervised learning is the clustering of data. A combination of supervised and unsupervised learning methods in a hybrid learning strategy is also possible.

3.3 Optimization Algorithms

Optimization algorithms are techniques designed and aimed to find the maximum/minimum or optimal solution of a determined function or problem. First optimization algorithms were introduced in the twentieth century. Nowadays, optimization algorithms are applied to a grand variety of applications. As it is explained in the work of Yang et al. [35], one of the biggest application fields of the optimizations algorithms is the industrial engineering world, where the reduction of costs, the increment of the efficiency and the optimization of the industrial processes have become of capital importance.

An important group inside the optimization algorithms is the bio-inspired or metaheuristics algorithms, which are inspired in natural processes to solve optimization problems. The metaheuristic algorithms [37–39] have been widely studied in the literature. In the following lines the metaheuristic optimization algorithm used in the design process exposed in this chapter and the multivariable optimization is explained in detail.

– **Particle Swarm Optimization**

The PSO [40] algorithms are metaheuristic optimization algorithms inspired in the behavior of a group of particles, referred as swarm, in a search space and evolving towards an optimal solution. As it is explained in the work of Khan and Singh [38], this algorithm is widely used due to its high robustness, small number of tunable parameters and its easy implementation.

As introduced in the work of Khan and Singh [38], each particle is a possible solution to the optimization problem, and is associated with a position vector $x_{i,t}$ and a velocity vector $v_{i,t}$. Exactly as in the case of the GAs, in a PSO algorithm there must be a fitness function that evaluates the specification fulfillment of each particle and provides them with a fitness values.

The velocity and position update of each particle is calculated with the following expressions presented in Eq. (18) and Eq. (19), respectively.

$$v_{i,t+1} = H \cdot v_{i,t} + \varphi_1 \cdot \left(x\_opt_{i,t} - x_{i,t}\right) + \varphi_2 \cdot \left(x\_global\_opt_{i,t} - x_{i,t}\right) \quad (18)$$

$$x_{i,t+1} = x_{i,t} + v_{i,t+1} \cdot \Delta t \quad (19)$$

where H [kg m^2] is the inertia constant of the system, φ_1 [-] is the exploitation factor, φ_2 [-] is the exploration factor, $x\_opt_{i,t}$ [m] is the best solution of the particle and $x\_global\_opt_{i,t}$ [m] is the best solution of the whole swarm.

As it can be observed in Eq. (18), the velocity of each particle is computed with regard to the personal best fitness obtained by that particle and the global best fitness obtained by the whole swarm. By modifying factors φ_1 [-] and φ_2 [-] the exploration and exploitation capability of the algorithm can be configured. Furthermore, the inertia constant H [kg m^2] defines the movement capacity of the particles.

The execution of a PSO could be summarized in the following 5 steps:

(1) **Initialization**. The swarm population is randomly formed.
(2) **Evaluation**. The fitness of each individual particle is evaluated.
(3) **Modification**. The best position of each particle, the best position of the whole swarm and each particle's velocity are computed.
(4) **Update**. Move each particle to the new position.
(5) **Termination**. Steps 2 to 4 are repeated until a termination condition has been satisfied.

– **Multiobjective optimization. Pareto optimal Front**

A multiobjective optimization [41–43] problem is that in which more than one objective is to be optimized. In contrast to single-objective optimization problems, in multiobjective cases there is not only one unique optimal solution, but a set of optimal solutions that respond to the trade-off or compromise necessity between the objectives to be optimized.

The concept of optimization of multiobjective problems was generalized in the work of Pareto [44] in 1896. In these type of problems a solution is dominated if there is any other solution that has a better (higher or lower depending on the context of the optimization problem) fitness value for all the objectives to be optimized. If there is no such a solution. The set of non-dominated solutions is known as the Pareto optimal Front (PoF). A PoF of a double-objective optimization problem is illustrated in Fig. 4.

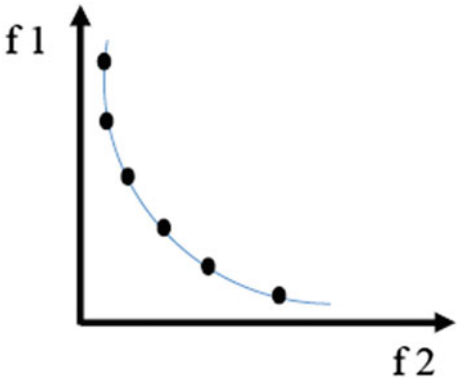


Fig. 4 PoF of a double-objective optimization problem

4 Machine Learning Based Wind Turbine Yaw Control

An adequate alignment of the wind turbine rotor with respect to the incoming wind by means of the yaw system of the wind turbine enables increment of the power generation at cost of an increase of the mechanical loads in different elements of the wind turbine, especially the yaw bearings. Hence, an adequate design and tuning of the yaw control system is of great importance to both optimize the power generation of the wind turbine and ensure its safe operation.

Usually, classical control structures based on PIDs have been used for the design of the yaw control strategy of the wind turbine [17, 45]. However, these classical control structures show some drawbacks in form of "wind up" of the integral action and posterior big oscillations, which can result in an undesired increment of the mechanical loads. As a result, some advanced control strategies for the yaw angle control of a wind turbine are proposed in the literature [5, 8, 46–48].

In this chapter, with the objective of achieving an improved performance of the yaw control system of a wind turbine, a ML based wind turbine yaw control system is exposed. A block diagram of the proposed ML based yaw control strategy is presented in Fig. 5.

The proposed yaw control system is based on the following AI techniques:

- A RL algorithm that learns from its own experience and enables the wind turbine to select the optimal decision in each scenario of its operation.
- An ANN to store the data of the matrix $Q(s, a)$ of the RL algorithm.
- A PSO and PoF based optimization algorithm to select the set of optimal actions that respond to the compromise necessity between the power increment and the mechanical loads associated to the yaw rotation.

This section is structured as follows: the design procedure of the RL algorithm applied to the ML based yaw control is explained in Sect. 4.1. Section 4.2 presents the structure and design process of the MLP-BP neural network. The design of the

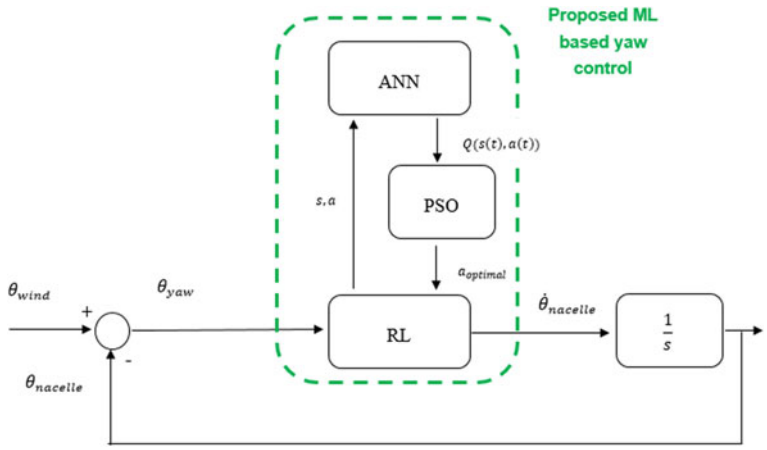


Fig. 5 Pipeline of the proposed ML based yaw control

PSO and PoF based algorithm is explained in Sect. 4.3. Finally, the Decision Making (DM) algorithm is exposed in Sect. 4.4.

4.1 Yaw Control RL

The RL algorithm developed for the yaw control of a HAWT presents multiple state, action and immediate reward variables. The objective of the multivariable structure is an improved characterization of the system in the most accurate way possible. To that end, 2 state variables, 2 action variables and 2 immediate reward variables are considered in the proposed RL algorithm.

The states s are:

– **StateYawA** [deg]: This state defines the orientation difference between the wind incident to the rotor and the nacelle of the wind turbine is shown in Eq. (20).

$$\theta_{yaw} = \theta_{wind} - \theta_{nacelle} \tag{20}$$

– **StateWindS** [m/s]: This state defines the wind speed value incident to the rotor.

The actions a are:

– **ActionYawK** [-/s]: This action defines the proportional gain associated to the yaw rotational speed controller. The expression to calculate the yaw rotational speed is shown in Eq. (21).

$$\Omega_{yaw} = ActionYawK \cdot \theta_{yaw} \tag{21}$$

– **ActionYaw** [deg]: This action defines the limit associated to the yaw rotation. In some cases, due to mechanical actuator problem or safety issues, the yaw rotation of the nacelle is limited to a certain value. The expression to note the rotation range allowed by this action is shown in Eq. (22).

$$\Delta\theta_{yaw} \in [-\text{ActionYaw} \ , \ \text{ActionYaw} \ ] \tag{22}$$

The immediate rewards r are:

– **RewardP** [%]: This immediate reward defines the power gain achieved by the wind turbine when a certain yaw action is performed. The expression to compute this immediate reward is shown in Eq. (23).

$$RewardP = \frac{P\_control - P\_no\_control}{P\_no\_deviation} \cdot 100 \tag{23}$$

As it can be observed in Eq. (23), in order to calculate the power gain 3 different scenarios related to the yaw actuation of the wind turbine are considered. The scenario P_control refers to the scenario in which the yaw control of the wind turbine is active and the nacelle of the wind turbine rotates to the yaw command provided by the yaw control and at the provided yaw speed value. The scenario, P_no_control refers to the scenario in which the yaw control of the wind turbine is not active, and, thus, the orientation of the wind turbine nacelle is fixed. Finally, the scenario P_no_deviation refers to the scenario in which the nacelle of the wind turbine is perfectly aligned with the direction of the wind incoming to the rotor.

– **RewardM** [N m]: This immediate reward defines the value of the mechanical moment in the yaw bearings. The value of this immediate reward has been defined with the mechanical moment in the yaw bearings because it has been found as the most critical mechanical load when performing a yaw rotation. Different mechanical load values, or even a weighted average of them, could be considered as the immediate reward to be considered by the proposed ML based yaw control algorithm.

As it was stated in Sect. 3.1 of this chapter, in a RL algorithm the calculation of the values $Q(s, a)$ for each state-action par is associated to the long-term reward considering a discount factor γ, see Eq. (3). In the RL algorithm proposed in this chapter there are 2 different immediate rewards r. Therefore, 2 different matrices $Q(s, a)$ will result in the algorithm. The expression to calculate the matrices $Q(s, a)$ using the immediate rewards r is shown in Eq. (24).

$$Q(s, a) = \sum_{i=0}^{i=T} r_{t+i} \cdot \gamma^i \tag{24}$$

The expression in Eq. (24) is applied to both the immediate rewards r considered in the ML based yaw control algorithm presented in this paper and the expression of

both matrices $Q(s, a)$ are obtained and presented in Eqs. (25) and (26). The discount factor γ is set to 1 in both cases because it is considered that all the values in the time horizon are equally important.

$$Q\_P(s,a) = \frac{\frac{1}{T}\int_t^{t+T}(P\_control - P\_no\_control)\cdot dt}{\frac{1}{T}\int_t^{t+T} P\_no\_deviation \cdot dt} \cdot 100[\%] \tag{25}$$

$$Q\_M(s,a) = \int_t^{t+T} RewardM(t)\cdot dt[N\cdot m] \tag{26}$$

After definition of the states s, actions a, immediate rewards r and the expressions of the matrices $Q\_P(s, a)$ and $Q\_M(s, a)$, simulations of the performance of the wind turbine to obtain data for the training process of the RL algorithm are carried out. The simulations are carried out with the aeroelastic code FAST [49] and the wind turbine model NREL 5 MW, presented in the work of Jonkman et al. [13], both designed by the National Renewable Energies Laboratory (NREL) in the USA.

The objective of the training process of the RL algorithm is to obtain the data related to all possible actuation scenarios associated to the yaw control of the wind turbine. Thus, in the design process presented in this chapter, an offline training process of the wind turbine with all the possible considered wind speed values and the yaw control actions is proposed. Thus, simulations with StateWindS = 3:1:17 [m/s], StateYawA = −90:10:90 [deg], ActionYawK = 0.1:0.1:1 [-/s] and ActionYaw = −90:10:90 [deg] have been carried out with the aeroelastic code FAST. The values of the matrices $Q\_P(s, a)$ and $Q\_M(s, a)$ are calculated with the data obtained from the simulations, see Eqs. (25) and (26).

4.2 Yaw Control MLP-BP

A MLP-BP neural network is designed to store the data of the matrices $Q\_P(s, a)$ and $Q\_M(s, a)$ corresponding to the RL algorithm. The objective of storing these matrices as continuous functions $Q\_P(s(t), a(t))$ and $Q\_M(s(t), a(t))$ is to eliminate the necessity of large amount of data management, which could result problematic in the implementation of the control strategy in the control system of the wind turbine, due to memory issues. Additionally, with the use of an ANN to store the data of the RL algorithm, the replacement policy of the RL algorithm is incorporated, since the ANN learns from the new calculated values. This aspect is of great importance if an online training of the RL algorithm during operation of the wind turbine is implemented. In that case, the ANN continuously learns from new calculated values and the accuracy of the functions $Q\_P(s(t), a(t))$ and $Q\_M(s(t), a(t))$ increase.

The selected topology of the ANN designed to store the data of the matrices $Q\_P(s, a)$ and $Q\_M(s, a)$ is a MLP-BP. The designed MLP-BP neural network presents 4 inputs and 2 outputs. A pipeline with the input and outputs of the designed MLP-BP neural network is presented in Fig. 6. Internally, the MLP-BP presents a structure

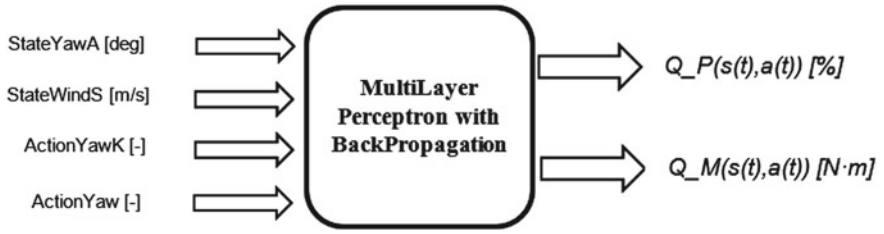


Fig. 6 Input and outputs of the MLP-BP designed for the ML based yaw control strategy

with one input layer with 4 neurons, two hidden layers with 75 neurons and 25 neurons respectively and one output layer with 2 neurons.

The learning rate for the training process of the MLP-BP has been set to $1 \cdot 10^{-50}$. The training ratio, validation ratio and test ration have been set to 90%, 5% and 5%, respectively. After the training process, correlation coefficients of 0.9999 and Mean Squared Error (MSE) of $1.62 \cdot 10^{-6}$ are obtained. The high value of the correlation coefficient and the low value of the MSE are indicators of a correct training process and that the MLP-BP is good enough to be used in the ML based yaw control strategy proposed in this chapter.

4.3 Yaw Control PSO and PoF

As it was stated in Sect. 2 of this chapter, the yaw actuation of a wind turbine allows alignment of the rotor of the wind turbine with the direction of the incoming wind and, thus, the power generated by the wind turbine can be maximized in some scenarios. Nevertheless, this power gain is achieved at cost of high mechanical loads in several components of the wind turbine, especially the yaw bearings, which could endanger the safe operation of the wind turbine or reduce its lifetime.

The objective of the PSO and PoF based optimization algorithm designed in this paper is to obtain a set of optimal yaw actions, ActionYawK [-/s] and ActionYaw [deg], that respond to the compromise necessity between RewardP [%] and RewardM [N m].

The output of the PSO and PoF optimization algorithm is a set of optimal solutions, known as PoF, that respond to the compromise necessity between the power gain and the mechanical loads due to the yaw rotation. To calculate this PoF the optimization algorithm makes use of the functions $Q\_P(s(t), a(t))$ and $Q\_M(s(t), a(t))$ as the fitness functions. The states of the system, StateYawA [deg] and StateWindS [m/s], are defined and the fitness value of different set of actions, ActionYawK [-/s] and ActionYaw [deg], is evaluated. The final optimal solutions are the solutions in which one of the fitness values cannot be increased without degrading the other one.

4.4 Yaw Control DM

The DM algorithm selects one of the optimal actions proposed as the result of the PSO-PoF optimization algorithm. The DM algorithm proposed in this chapter considers the mechanical loads as the limiting factor when selecting the yaw actuation and it could be summarized as follows:

- The solutions that suppose a value of the function $Q\_M(s(t), a(t))$ higher than a predefined threshold are not taken into consideration due to safety issues.
- From the set of solutions that are taken into consideration, the one with the highest value of the function $Q\_P(s(t), a(t))$ is selected.

Other different approaches for the selection of the yaw optimal actuation based on more complex principles could also be evaluated and implemented.

5 Conclusions

The design procedure of a ML based yaw control algorithm for a HAWT based on AI techniques has been presented in this chapter. The proposed yaw control strategy is aimed to improve the performance of classical yaw control strategies by means of the use of AI techniques, which emulate the performance of natural processes to provide digital systems with intelligence and self-learning capability. The self-learning capability is the main characteristic of the ML.

The proposed ML based yaw control strategy makes use of three different AI techniques for the development of the control strategy. The RL algorithm maps actions to states and thus allows the development of a policy in the wind turbine that selects the best actions in different wind turbine operation scenarios. The ANN provides a very important learning capability and allows a continuous learning process in the wind turbine, as well as, a simplified data management by storage of large amounts of data as continuous functions. Finally, the PSO and PoF based optimization algorithm allows to select the actions that maximize the power output of the wind turbine and minimize the mechanical loads generated as a result of the yaw rotation.

The most important capability of the proposed ML based yaw control strategy is the self-tuning. As a result of the self-learning capability of the ML system, there is no need for tuning a closed loop for the yaw angle control of the wind turbine. Therefore, the risk associated to a possible inadequate tuning of this control loop is erased. In fact, an inadequate control tuning could cause considerable power generation losses or high mechanical loads that could endanger the safe operation of the wind turbine.

Simulations of the proposed ML based yaw control strategy with the aeroelastic code FAST show promising results in comparison to other more simple controllers based on the classical control theory. The most visible improvements are increased generated power values and considerable mechanical load reductions in the yaw bearings of the wind turbine for different wind scenarios.

Acknowledgements The authors are grateful to the Government of the Basque Country and the University of the Basque Country UPV/EHU through the SAIOTEK (S-PE11UN112) and EHU12/26 research programs, respectively.

Funding: This research was partially funded by Fundation VITAL Fundazioa.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rosales-Asensio E, Borge-Diez D, Blanes-Peiro J, Perez-Hoyos A, Comenar-Santos A (2019) Review of wind energy technology and associated market and economic conditions in Spain. Renew Sustain Energy Rev 101:415–427
2. WindEurope (2019) Wind energy in Europe in 2018. Trends and Statistics
3. Nyanteh Y, Schneider N, Netter D, Wei B, Masson PJ (2015) Optimization of a 10 MW direct drive HTS generator for minimum levelized cost of energy. IEEE Trans Appl Supercond 25(3):1–4
4. Kim Y (2016) Robust data driven H-infinity control for wind turbine. J Franklin Inst 353(13):3104–3117
5. Saenz-Aguirre A, Zulueta E, Fernandez-Gamiz U, Lozano J, Lopez-Guede JM (2019) Artificial neural network based reinforcement learning for wind turbine Yaw control. Energies
6. Saenz-Aguirre A, Fernandez-Gamiz U, Zulueta E, Ulazia A, Martinez-Rico J (2019) Optimal wind turbine operation by artificial neural network-based active gurney flap flow control. Sustainability 11:2809
7. Aramendia I, Fernandez-Gamiz U, Zulueta E, Saenz-Aguirre A, Teso D (2019) Parametric study of a gurney flap implementation in a DU91W(2)250 airfoil. Energies
8. Saenz-Aguirre A, Zulueta E, Fernandez-Gamiz U, Ulazia A, Teso D (2019) Performance enhancement of the artificial neural network based reinforcement learning for wind turbine Yaw control. Wind Energy
9. Fernandez-Gamiz U, Zulueta E, Boyano A, Ansoategui I, Uriarte I (2017) Five megawatt wind turbine power output improvements by passive flow control devices. Energies 10(6):742
10. Zhang D, Han X, Deng C (2018) Review on the research and practice of deep learning and reinforcement learning in smart grids. CSEE J Power Energy Syst 4(3):362–370
11. Yang Z, Merrick K, Jin L, Abbass HA (2018) Hierarchical deep reinforcement learning for continuous action control. IEEE Trans Neural Netw Learn Syst 29(11):5174–5184
12. Bivona S, Bonanno G, Burlon R, Gurrera D, Leone C (2010) Stochastic models for wind speed forecasting. Stochas Models Wind Speed Forecast 52(2):1157–1165
13. Jonkman JM, Butterfield S, Musial W, Scott G (2009) Definition of a 5MW reference wind turbine for offshore system development. National Renewable Energy Laboratory (NREL)
14. Yücel M, Özder S, Design and efficiency of 5 kW wind turbine without gearbox, controlled by Yaw and pitch drivers. Çanakkale Onsekiz Mart Üniversitesi Fen Bilimleri Enstitüsü Dergisi 4 (1):74–87
15. Ahrens M, Kucera L, Larsonneur R (1996) Performance of a magnetically suspended flywheel energy storage device. IEEE Trans Control Syst Technol 4(5):495–502
16. Zheng S, Yang J, Song X, Ma C (2018) Tracking compensation control for nutation mode of high-speed rotors with strong gyroscopic effects. IEEE Trans Ind Electron 65(5):4156–4165
17. Shariatpanah H, Fadaeinedjad R, Rashidinejad M (2013) A new model for PMSG-based wind turbine with Yaw control. IEEE Trans Energy Convers 28(4):929–937
18. Wang X, Li X, Leung VCM (2015) Artificial intelligence-based techniques for emerging heterogeneous network: state of the arts, opportunities, and challenges. IEEE Access 3:1379–1391

19. Fadlullah ZM, Tang F, Mao B et al (2017) State-of-the-art deep learning: evolving machine intelligence toward tomorrow's intelligent network traffic control systems. IEEE Commun Surv Tutor 19(4):2432–2455
20. Jagodnik KM, Thomas PS, Bogert AJvd, Branicky MS, Kirsch RF (2017) Training an actor-critic reinforcement learning controller for arm movement using human-generated rewards. IEEE Trans Neural Syst Rehabilitat Eng 25(10):1892–1905
21. Mongillo G, Shteingart H, Loewenstein Y (2014) The misbehavior of reinforcement learning. Proc IEEE 102(4):528–541
22. Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. MIT Press, Cambridge, MA, USA
23. Kaelbling LP, Littman ML, Moore AW (1996) Reinforcement learning: a survey. J Artif Intell Res 4(1):237–285
24. Bertsekas DP (2013) Abstract dynamic programming. Athena Scientific, Belmont, MA, USA
25. Bertsekas DP (2012) Dynamic programming and optimal control: approximate dynamic programming, no 2. Athena Scientific, Belmont, MA, USA
26. Bertsekas DP (2017) Value and policy iterations in optimal control and adaptive dynamic programming. IEEE Trans Neural Netw Learn Syst 28(3):500–509
27. Kao K, Wu I, Yen S, Shan Y (2013) Incentive learning in Monte Carlo tree search. IEEE Trans Comput Intell AI Games 5(4):346–352
28. Coulom R (2006) Efficient selectivity and backup operators in Monte-Carlo tree search. In: Proceedings of 5th international conference computer games, pp 72–83
29. Watkins CJCH (1989) Learning from delayed rewards. PhD thesis, King's College, Cambridge, UK
30. Watkins CJCH, Dayan P (1992) Q-learning. Mach Learn 8(3):279–292
31. Adam S, Busoniu L, Babuska R (2012) Experience replay for real-time reinforcement learning control. IEEE Trans Syst Man Cybernet Part C (Appl Rev) 42(2):201–212
32. Chen SH, Jakeman AJ, Norton JP (2008) Artificial Intelligence techniques: an introduction to their use for modelling environmental systems. Mathemat Comput Simul 78:379–400
33. Yao X (1999) Evolving artificial neural networks. Proc IEEE 87(9):1423–1447
34. Oong TH, Isa NAM (2011) Adaptive evolutionary artificial neural networks for pattern classification. IEEE Trans Neural Netw 22(11):1823–1836
35. Yang XS (2013) Optimization and metaheuristic algorithms in engineering. In: Yang XS, Gandomi AH, Talatahari S, Alavi AH (eds) Metaheursitics in water, geotechnical and transport engineering. Elsevier, pp 1–23
36. Jain AK, Mao J, Mohiuddin KM (1996) Artificial neural networks: a tutorial. Computer 29(3):31–44
37. Wang L, Shen J (2017) A systematic review of bio-inspired service concretization. IEEE Trans Serv Comput 10(4):493–505
38. Khan B, Singh P (2017) Selecting a meta-heuristic technique for smart micro-grid optimization problem: a comprehensive analysis. IEEE Access 5:13951–13977
39. Bala A, Ismail I, Ibrahim R, Sait SM (2018) Applications of metaheuristics in reservoir computing techniques: a review. IEEE Access 6:58012–58029
40. Kennedy J, Eberhar RC (1995) Particle swarm optimization. In: Proceedings IEEE international conference on neural network. Perth, WA, Australia, pp 1942–1948
41. Ehrgott M, Gandibleux X (2000) A survey and annotated bibliography of multiobjective combinatorial optimization. OR-Spektrum 22(4):425–460
42. Durillo JJ, Nebro AJ, García-Nieto J, Alba E (2010) On the velocity update in multi-objective particle swarm optimizers. In: Coello CA, Dhaenens C, Jourdan L (eds) Advances in multi-objective nature inspired computing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 45–62
43. Coello Coello CA, Dhaenens C, Jourdan L (2010) Multi-objective combinatorial optimization: problematic and context. In: Coello CA, Dhaenens C, Jourdan L (eds) Advances in multi-objective nature inspired computing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–21
44. Pareto V (1896) Cours D'Economie Politique, F. Rouge, Lausanne, I(II)

45. Karakasis N, Mesemanolis A, Nalmpantis T, Mademlis C (2016) Active yaw control in a horizontal axis wind system without requiring wind direction measurement. IET Renew Power Gener 10(9):1441–1449
46. Song D, Fan X, Yang J, Liu A, Chen S, Joo YH (2018) Power extraction efficiency optimization of horizontal-axis wind turbines through optimizing control parameters of yaw control systems using an intelligent method. Appl Energy 224:267–279
47. Song D, Yang J, Fan X et al (2018) Maximum power extraction for wind turbines through a novel yaw control solution using predicted wind directions. Energy Convers Manage 157:587–599
48. Bharani R, Jayasankar KC (2015) Yaw control of wind turbine using fuzzy logic controller. Power Electron Renew Energy Syst 326:997–1006
49. NREL NWTC FAST version 7. Available online: https://nwtc.nrel.gov/FAST7/. Accessed 21 Oct 2018

Numerical Methods of Electric Power Flow in Interconnected Systems

Marian Gaiceanu, Vasile Solcanu, Theodora Gaiceanu, and Iulian Ghenea

Abstract In this chapter the power flow problem is treated as mandatory part of the Energy Management System. The electrical energy flow into power system cannot be stored. The Energy Management System balance the request between the generation and load demand. From this point of view, the steady state behaviour of the power system is essentially to take action in case of contingency events. The authors of this chapter takes into account the most used algorithms of power flow problem, and they are presented in context of interconnected power systems with different type of buses: with load, generators, and reference buses, taking into consideration the limitation of the generated reactive power. The theoretical aspects of the numerical methods are related to the electric power flow in interconnected systems, and they are proved through the delivered case studies: Gauss-Seidel, and Newton Raphson.

Keywords Power system · Load flow · Numerical methods · Gauss–seidel · Newton–raphson · Fast-decoupled load flow

M. Gaiceanu (✉)
Department of Automatic Control and Electrical Engineering, University Dunarea de Jos of Galati, Galati, Romania
e-mail: marian.gaiceanu@ugal.ro

V. Solcanu · I. Ghenea
Doctoral School of Fundamental and Engineering Sciences, University Dunarea de Jos of Galati, Galati, Romania
e-mail: vasilesolcanu@dedeman.ro

I. Ghenea
e-mail: iulian.ghenea@ugal.ro

T. Gaiceanu
Department of Automatic Control and Applied Informatics, Gheorghe Asachi Technical University of Iasi, Iasi, Romania
e-mail: gaiceanu.theodora@ac.tuiasi.ro

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_33

Abbreviations

FD	Fast decoupled
FDLF	Fast-decoupled load flow
EMS	Energy Management System
GA	Genetic algorithm
GS	Gauss-Seidel
NR	Newton-Raphson
PFP	Power flow problem
PSO	Particle Swarm Optimization
PV	Generator bus
PQ	Load bus
PS	Power system
SCADA	Supervisory Control and Data Acquisition
Slack (swing) bus	Reference bus
SVCs	Static var systems

Symbols

B	Susceptance
D1	First time derivative
G	Conductance
I	Current
J	Jacobian
y	Admittance
P	Active, real power
Q	Reactive, imaginary power
R	Resistance
$\underline{S}$	Complex (apparent) power
U	Voltage magnitude
$\underline{U}$	Voltage phasor
$\underline{x}$	Unknown vector
X	Reactance
ε	Tolerance
φ	Voltage angle

1 Introduction

Power flow problem (PFP) is one important issue in planning, operation and control of the power system (PS), in order to maintain a reliable PS operation within the imposed performances, as depicted in [1, 2]. In addition, the modern PS should face-up to natural disaster and deliberate attacks [3]. By solving the PFP, the voltage magnitudes, and the corresponding angles are computed to all network buses in chosen topology. The PFP takes part from the Energy Management System (EMS) of the power network [4].

The EMS takes corrective actions (by automatically choosing a standard network topology based on SCADA delivered data) in case of contingency occurrences in order to maintain the PS operation without load interruptions [5].

Analysis of steady-state load flow is performed in order to design power systems, to determine the voltage amplitudes and phase angles of the buses, planning energy production based on energy and economic efficiency, for extension of the power system with an efficient management [6, 7]. Methods for analysing the load flow [8] in power systems [9, 10] are based on algorithms for numerical computation: Gauss-Seidel (GS) or successive displacement [11], various algorithms of Newton-Raphson (NR) [12, 13] methods, decoupled [14] and fast-decoupled load flow (FDLF) [15], modern algorithms based on artificial intelligence [16, 17], algorithms rely on fuzzy logic [18], PSO [19] and GA optimization [20, 21]. In order to improve the performances of the load flow problems, the hybrid numerical solutions are in full development [22, 23].

The method of solving the problem of power flow with fast decoupled load is used in contingency calculations for electrical networks of any size and can be implemented efficiently on limited memory computing resources. The method can be applied also for the calculations of the normal operation of the networks, but especially for the studies of contingency evaluation on-line or off-line.

For the distribution systems of electricity and for the energy systems the analysis of the stationary regime is mandatory. These systems are characterized by the simultaneous production and consumption of electricity. Because the electricity cannot be stored in large quantities, and the control of the production must be permanently adapted to the variations of the consumption demand, the PS is sensitive to disturbances (any disturbance propagates instantaneously throughout the system).

The objective of the chapter is to solve PFP using specific numerical methods of power systems. In determining the solutions, direct (iterative) search methods and second-order methods from the existing numerical methods were used. Methods of the second order involve the determination of the Jacobian elements. In order to determine the general algorithm of NR methods, the authors start from the simple, monovariable case of determining the solution, where the inversion of the second order derivative is avoided. Thus, by generalization, the PFP solution is based on the Jacobian direct deduction, avoiding the Jacobian inversion calculation effort.

This chapter presents solving methods of PFP, and algorithms for implementing numerical solutions. In order to test the above mentioned algorithms, power flow

case studies for 4 buses power network solved by using GS (known as Liebmann method), NR, and FDLF numerical methods are presented, and the obtained results are shown in this chapter. The main conclusions of this chapter are presented in the last section of the chapter.

2 Type of Buses

In Fig. 1, one generic power network topology is represented [24]. The 9-bus power system [25] comprises generation buses (B1–B3), load buses (B5, 7, 9), neither

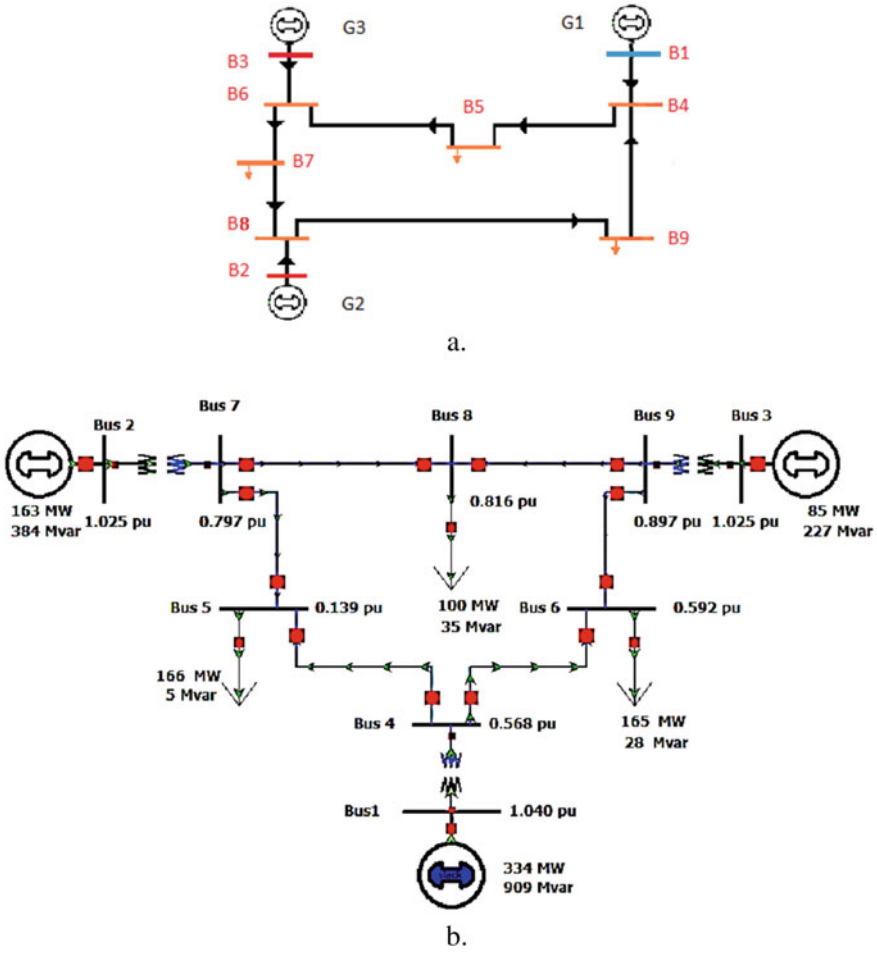


Fig. 1 **a** Single line diagram for 9 bus system: Slack, PV, PQ buses **b** IEEE 9-bus system [24, 25]

generation or load (B4, 6, 8). At the same time, one additional bus type may be found both generation and load (not represented).

Power injection (active or reactive) into a bus either from generator or load could be positive (power flow from external to the bus) or negative (power flows from the bus to the external connections). Generators are characterized by positive active power injection, or operating as motors being characterized by negative power injection. From the reactive power point of view, the positive (lagging operation mode of generators) if the generator delivers reactive power into the bus, and negative (leading operation mode of generators) if the generator absorbs the imaginary power from the bus. In case of unity power factor operation the reactive power is zero. From the loads point of view, the real and imaginary power could be either positive or negative. Conventionally, the allocated signs for the power (active and reactive) injection are as follows: into bus (positive defined): generator, out of bus (negative defined): load. For every bus four variables can be taken into account: active, P_k, and reactive, Q_k, power, voltage angle, φ_k, and voltage magnitude $|\underline{U}_k|$.

The positive net injection is considered if the power flows into the bus (with both generation and load elements): $P_k = P_{gk} - P_{lk}$, net active or real power injection; $Q_k = Q_{gk} - Q_{lk}$, net reactive or imaginary power; $S_k = S_{gk} - S_{lk}$, complex or apparent power, with $S_k = P_k + jQ_k$.

By following the above mentioned net power injection, there are three types of buses:

1. *PV (generator) Buses or voltagecontrolled buses*: the P_k and voltage magnitude $|V_k|$ values of the buses are known, Q_k or φ_k should be calculated. Special case of nongenerator bus (for example, a bus with load) can be considered as PV bus: containing the static var systems (SVCs) or switched shunt capacitors (B2, 3 from Fig. 1a).
2. *PQ(load) Buses*: where the P_k and Q_k are known, and $|\underline{U}_k|$ or φ_k should be calculated (specific for the load buses, B4-9 buses from Fig. 1.a).
3. *Slack* or *swing* Buses (*reference bus*): voltage magnitude $|\underline{U}|$ and voltage angle φ values are known. This is a generator bus, and it is unique. The generation supply both the load demands and the losses. All the PQ buses power injection values are known. The losses would be calculated by solving the *power flowproblem.*

Before solving the power flow problem, all injections at PQ buses (except of setting the active power injection for one generator are known. This generator swings or slack in order to compensate the losses). The swing bus voltage magnitude value is considered as generator voltage, and the voltage angle can be considered as reference value, i.e. 0°.

In Fig. 1b the IEEE 9-bus system [24, 25].

3 The Power Flow Problem PFP

One important element in managing the PFP is the capability of the imaginary power of the synchronous generators. By knowing the network topology, the power and the voltage magnitudes for studied buses, the voltage angles and corresponding magnitudes are determined.

The power flow solution is obtained through the numerical methods approaches and consists of determining all the voltages and angles to all buses. In this manner, the active and reactive power to all generation and load buses can be deducted.

3.1 The Operating Area for Unit Generator

The power limits of the generators are represented in Fig. 2 [26] (M-maximum value, m- minimum value) in PQ reference frame [27, 28]. The reactive power is absorbed when the generator is in leading operation mode, or can be produced in lagging power factor operation mode. The Q_M point is essential to maintain the field windings in the admissible temperature limits. The point P_M is the limitation for armature current. The Q_m point is an indicator of the power angle limitation to produce necessary real power [29].

3.2 The 4 Bus System Model

Current injection, I_i, into a bus i (Fig. 3) is nodal issue of one power system, instead of current flow which is related to the branch of one circuit [27]. Conventionally, when the current gets in into the bus, the positive current injection is considered. The relation between the current injections at a bus and the bus voltages is based on the admittance matrix [29].

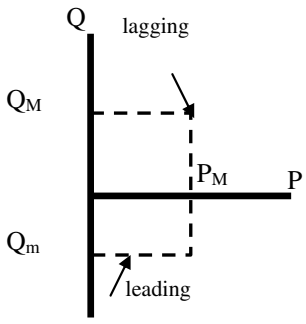


Fig. 2 The generator maximum capability (P_M) and the reactive power limits

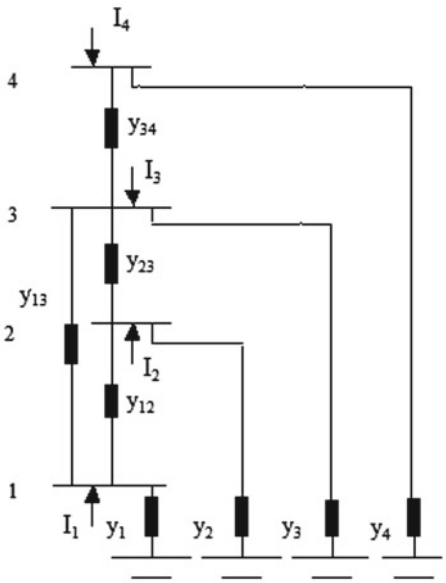


Fig. 3 One-line diagram of four-bus power system

In Fig. 3, in order to express the admittance matrix of a network the generic network model has been considered. The power system comprises 4 buses (1–4), considered as nodes of the network, branches to the ground (shunt capacitor or ends of line), branches between the buses (nodes). The branches are characterized by the admittance y_{ij} between *i*, *j* buses or by y_i for the branches connected between bus and ground (shunt).

According to the first Kirchhoff's Theorem, and Ohm's Law, for the *bus* 1, the injected current I_1 (knowing that $U_{ij} = U_i - U_j$) can be written as:

$$I_1 = U_{12}y_{12} + U_{13}y_{13} + U_1y_1 \tag{1}$$

In order to obtain a matrix representation, it can be considered that the bus 1 is linked to 4 through zero admittance (or infinite impedance) value, $y_{14} = 0$:

$$I_1 = U_{12}y_{12} + U_{13}y_{13} + U_{14}y_{14} + U_1y_1 \tag{2}$$

or in nodal voltages form:

$$I_1 = U_1(y_1 + y_{12} + y_{13} + y_{14}) - U_2y_{12} - U_3y_{13} - U_4y_{14} \tag{3}$$

By taking into account all the network buses, the matriceal representation of the system is obtained as follows:

$$\underline{I}_1 = \underline{\underline{Y}}\,\underline{U}_1 \tag{4}$$

The injected current from Eq. 3 can be extended to the other buses, and an adequate matriceal form:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} y_1 + \sum_{k=2}^{4} y_{1k} & -y_{12} & -y_{13} & -y_{14} \\ -y_{21} & y_2 + \sum_{\substack{k=1 \\ k\neq 2}}^{4} y_{2k} & -y_{23} & -y_{24} \\ -y_{31} & -y_{32} & y_3 + \sum_{\substack{k=1 \\ k\neq 3}}^{4} y_{3k} & -y_{34} \\ -y_{41} & -y_{42} & -y_{43} & y_4 + \sum_{k=1}^{3} y_{4k} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} \tag{5}$$

in which the symmetry of the admittance is used (from bus k to bus i, $y_{ki} = y_{ik}$).

The linear relationship between the nodal voltages and current injections is noted by Y-bus as follows:

$$\underline{Y} = \begin{bmatrix} y_1 + \sum_{k=2}^{4} y_{1k} & -y_{12} & -y_{13} & -y_{14} \\ -y_{21} & y_2 + \sum_{\substack{k=1 \\ k\neq 2}}^{4} y_{2k} & -y_{23} & -y_{24} \\ -y_{31} & -y_{32} & y_3 + \sum_{\substack{k=1 \\ k\neq 3}}^{4} y_{3k} & -y_{34} \\ -y_{41} & -y_{42} & -y_{43} & y_4 + \sum_{k=1}^{3} y_{4k} \end{bmatrix}, \tag{6}$$

or according to Fig. 3, the admittance matrix:

$$\underline{\mathbf{Y}} = \begin{bmatrix} G_{11} + jB_{11} & Y_{12} & Y_{13} & 0 \\ Y_{21} & G_{22} + jB_{22} & Y_{23} & Y_{24} \\ Y_{31} & Y_{32} & G_{33} + jB_{33} & Y_{34} \\ 0 & 0 & Y_{43} & G_{44} + jB_{44} \end{bmatrix} \tag{7}$$

If the voltage vector, $\underline{U}$, current vector, $\underline{I}$, are considered:

$$\underline{\mathbf{U}} = \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix}, \qquad \underline{\mathbf{I}} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} \tag{8}$$

The matriceal representation of the generic network (Fig. 3) is obtained as:

$$\underline{\mathbf{I}} = \underline{\mathbf{YU}} \tag{9}$$

or the k bus current injection component:

$$I_k = \sum_{j=\overline{1,N}} Y_{kj} U_j, \tag{10}$$

in which Y_{kj} does not has the admittances significance.

The elements of the main diagonal of the admittance matrix take into account both the branches connected between the buses and the branches connected to the ground by considering the following relationship:

$$Y_{ii} = y_i + \sum_{k=1,k\neq i}^{N} y_{ik} \tag{11}$$

3.3 The Problem Formulation of Power Flow

Taking into account the net complex apparent power definition

$$S_k = U_k I_k^*, \tag{12}$$

where $\underline{I}_k^*$ means the current conjugate vector, phasor $\underline{U}_k = |U_k| \angle \varphi_k$ is characterized both by magnitude and angle.

By inserting the current injection general equation from (10) into (12), the net complex apparent power function on the complex admittance can be obtained:

$$\underline{S}_k = U_k \sum_{j=\overline{1,N}} Y_{kj}^* U_j^* \tag{13}$$

It could be noted that per unit quantities and single-phase line representation have been taken into account:

By knowing the complex admittance defined as:

$$Y_{kj} = G_{kj} + j B_{kj}. \tag{14}$$

in which G_{kj} is the conductance (real part of the complex admittance) and B_{kj} is the susceptance (imaginary part of the complex admittance) and $\varphi_{kj} = \varphi_k - \varphi_j$, the net complex power becomes

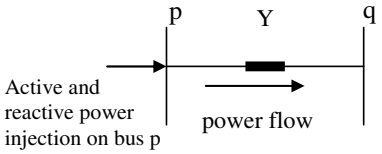


Fig. 4 Interconnected *p-q* buses

$$\begin{aligned}\underline{S}_k &= \left|\underline{U}_k\right|\angle\varphi_k \sum_{j=\overline{1,N}} Y_{kj}^* \left(\left|\underline{U}_j\right|\angle\varphi_j\right)^* \\ &= \sum_{j=\overline{1,N}} \left(\left|\underline{U}_k\right|\left|\underline{U}_j\right|\angle\varphi_{kj}\right) Y_{kj}^*\end{aligned} \tag{15}$$

By replacing voltage phasor $\underline{U}_j = \left|\underline{U}_j\right|\angle(\varphi_k - \varphi_j)$, with $\left|\underline{U}_j\right|\left(\cos\varphi_{kj} + j\sin\varphi_{kj}\right)$ into net complex apparent power $\underline{S}_k = P_k + jQ_k$, by separating the real side, and the imaginary side the power flow equations can be deducted, active power

$$P_k(U,\varphi) = \sum_{j=\overline{1,N}} |U_k|\left|U_j\right|\left(G_{kj}\cos\varphi_{kj} + B_{kj}\sin\varphi_{kj}\right) \tag{16}$$

and reactive power, respectively

$$Q_k(U,\varphi) = \sum_{j=\overline{1,N}} |U_k|\left|U_j\right|\left(G_{kj}\sin\varphi_{kj} - B_{kj}\cos\varphi_{kj}\right) \tag{17}$$

In particular case when the bus p is connected to one bus q (Fig. 4), inductive admittance $Y = (G - jB)$ (Fig. 4), without any shunt reactance, the power flow equations are obtained by components:

- Active (real) power

$$P_p = \left|U_p\right|^2 G - \left|U_p\right|\left|U_q\right| G\cos\varphi_{pq} + \left|U_p\right|\left|U_q\right| B\sin\varphi_{pq} \tag{18}$$

- Reactive (imaginary) power, respectively:

$$Q_p = \left|U_p\right|^2 B - \left|U_p\right|\left|U_q\right| G\sin\varphi_{pq} - \left|U_p\right|\left|U_q\right| B\cos\varphi_{pq} \tag{19}$$

3.4 Power Flow Problem

The following N buses network topology can be considered: 1 *swing* bus, N_G— voltage regulated buses, $N_G - 1$ PV buses, and PQ buses as $N - N_G$.

The knowing data is the input data for the power flow problem:

1. The complex admittance matrix is known;
2. $U_k, k = 1, \ldots, N_G$, voltage magnitudes of N_G generator buses are known,
3. In the remaining N – 1 buses, the known injected active power is denoted by P_k $(k = \overline{2, N})$;
4. For the k buses $k = \overline{N_G + 1, N}$, the known injected reactive power of PQ buses is denoted by Q_k.

Under these circumstances, the values of the active and reactive power injection of the system Eqs. 16, 17 are known, that means $2N - N_G - 1$ equations, i.e.:

$$\begin{aligned} P_k(\varphi, U) &= \sum_{j=\overline{1,N}} |U_k||U_j|\left(G_{kj}\cos\varphi_{kj} + B_{kj}\sin\varphi_{kj}\right), \quad k = \overline{2, N} \\ Q_k(\varphi, U) &= \sum_{j=\overline{1,N}} |U_k||U_j|\left(G_{kj}\sin\varphi_{kj} - B_{kj}\cos\varphi_{kj}\right), \quad k = \overline{N_G + 1, N} \end{aligned} \tag{20}$$

The output data of the power flow problem will be:

I. It could be mentioned that the angle voltage phasor of the swing bus is known (0°), but the other angles for the voltage phasors buses $\varphi_k, k = 2, .., N$ should be calculated;
II. The voltage magnitude to PQ buses, $|U_k|, k = N_G + 1, .., N$, should be determined.

The nonlinear system Eq. (20) can be solved through the numerical methods.

As a conclusion, the number of the necessary equations (the elements of the angle vector $\underline{\varphi}$, and the components of the voltage magnitude vector $|\underline{U}|$) should be equal with the unknown variables, $2N\text{-}N_G\text{-}1$, i.e. the dimension of the power flow solution problem or the $2N - N_G - 1$ components of the unknown vector $\underline{x}$:

$$\underline{\mathbf{x}}^T = \left[\varphi_2 \ldots \varphi_N \ |U_{N_G+1}| \ \left|U_{N_G+2}\right| \ldots |U_N\right]^T \tag{21}$$

Taking into consideration Eq. (21), the power flow Eqs. (20) can be viewed as $P_k(\underline{x}) - P_k = 0$, and $Q_k(\underline{x}) - Q_k = 0$.

By knowing the injected active and reactive power P_k $(k = \overline{2, N})$, Q_k $k = \overline{N_G + 1, N})$, the power flow Eqs. (20) can be expressed as dependent of $\underline{\mathbf{x}}$:

$$\underline{f}(\underline{\mathbf{x}}) = \underline{0} \tag{22}$$

where the vector-valued function size $\underline{f}(\underline{\mathbf{x}})$ is of $(2N - N_G - 1) \times 1$:

$$\begin{bmatrix} \Delta P_2(\varphi, U) \\ \vdots \\ \Delta P_N(\varphi, U) \\ - - - - - - - \\ \Delta Q_{N_G+1}(\varphi, U) \\ \vdots \\ \Delta Q_N(\varphi, U) \end{bmatrix} = \underline{0} \tag{23}$$

with $\Delta P_k\ (\varphi, U) P_k(\underline{x}) - P_k$, and $\Delta Q_k\ (\varphi, U) = Q_k(\underline{x}) - Q_k$.

The components of the power flow vector (22):

$$\underline{f}(\underline{x})^T = \left[f_1(\underline{x}) \ldots f_{N-1}(\underline{x}) \vdots f_N(\underline{x}) \ldots f_{2N-N_G-1}(\underline{x}) \right]^T. \tag{24}$$

From (23) and (24) the compact form equation is obtained as follows:

$$\underline{f}(\underline{x}^{(i)}) = \begin{bmatrix} \Delta \underline{P} \\ \Delta \underline{Q} \end{bmatrix}. \tag{25}$$

In order to solve the nonlinear vector-valued function (23) different iterative methods are introduced in this chapter: NR, and FDLF, and GN respectively.

4 Newton–Raphson Iterative Method

The problem is to find a solution to the nonlinear vector function $\underline{f}(\underline{x}) = \underline{0}$, denoted by $\underline{x}^*$.

The iterative procedure to find the series of the successive approximations to the solution is based on the following steps:

1. For the power systems the unity magnitude voltage bus value (we are supposing the per-unit representation) and zero degrees for the angles can be adopted as guess or initial solution. The estimated solution should be relatively close to the final solution.
2. The direction of finding new iterative solution guarantees the improvements of the new successive solution.
3. The new solution should be closer to the final solution, i.e. the convergence of the algorithm is assured.
4. The computing time should attain the time constraints.

For monovariablereal function $f(x) = 0$:

1. Initial estimate value of the solution $x^{(0)}$ is given;
2. Compute $f\left(x^{(0)}\right)$, which would be nonzero;

3. Calculate the first derivative around solution $x^{(0)}$: $D1f(0) = f'(x^{(0)})$, is a straight line approximation tangent at point $x^{(0)}$ to the nonlinear function f;
4. By using Taylor series expansion of the nonlinear f function around solution $x^{(0)}$, $f(x^{(0)} + \Delta x)$, and supposing $\Delta x^0 = \Delta x$, results $f(x^{(0)}) = -D1f(0) \cdot \Delta x^0$;
5. Extract from step 4 the ratio $\Delta x^0 = -[D1f(0)]^{-1} \cdot f\left(x^{(0)}\right)$;
6. Calculate the point $x^{(1)} = x^{(0)} - [D1f(0)]^{-1} \cdot f(x^{(0)})$, being an estimated solution of the f;
7. With deducted point $x^{(1)}$, compute the new function value $f(x^{(1)})$;
8. Calculate the new value of the first derivative around point $x^{(1)}$: $f'(x^{(1)}) = D1f(1)$;
9. Determine the new ratio $\Delta x^{(1)} = -f(x^{(1)})/D1f(1)$;
10. Calculate the second iterate value $x^{(2)} = x^{(1)} - [D1f(1)]^{-1} \cdot f(x^{(1)})$
11. If $\left|x^{(2)} - x^{(1)}\right| < \varepsilon,\ \varepsilon > 0$ and small, STOP.

Taking into account the above mentioned procedure, for the single variable case, the recurrent relation can be deducted:

$$x^{(j+1)} = x^{(j)} - [D1f(j)]^{-1} \cdot f(x^{(j)}) \tag{26}$$

The n-dimensionalcase.

The nonlinear n-dimensional equations system satisfies vector function $\underline{f}(\underline{\mathbf{x}}) = \underline{0}$ with n components of the solution vector $\underline{x}^*$.

By using the following notation:

$\underline{J} = \underline{D1f}(j) = \underline{f'}(\underline{x}^{(j)})$, $D1f1_1(j) = \frac{\partial f_1(\underline{x}^{(j)})}{\partial x_1}, D1f1_2(j) = \frac{\partial f_1(\underline{x}^{(j)})}{\partial x_2}$, … for n-dimensional case, the Jacobian matrix should be computed [27]:

$$\underline{D1f}(j) = \begin{bmatrix} D1f1_1(j) & D1f1_2(j) & \cdots & D1f1_n(j) \\ D1f2_1(j) & D1f2_2(j) & \cdots & D1f2_{n1}(j) \\ \vdots & \vdots & \vdots & \vdots \\ D1fn_1(j) & D1fn_2(j) & \cdots & D1fn_n(j) \end{bmatrix} \tag{27}$$

By analogy, with the single variable case, $x^{(i+1)} = x^{(i)} - [D1f(i)]^{-1} \cdot f\left(x^{(i)}\right)$ the recursive vector equation is denoted by:

$$\underline{\mathbf{x}}^{(i+1)} = \underline{\mathbf{x}}^{(i)} - \left[\underline{D1f}(i)\right]^{-1} f\left(\underline{\mathbf{x}}^{(i)}\right) \tag{28}$$

where $\left[\underline{D1f}(i)\right]^{-1} = \underline{J}^{-1}$.

The matrix inversion could be difficult and time consuming. Therefore, by using the vector equation:

$$\Delta \underline{x}^{(i)} = \underline{x}^{(i+1)} - \underline{x}^{(i)}, \tag{29}$$

the following relationship can be deducted:

$$-\underline{J}\Delta\underline{x}^{(i)} = \underline{f}(\underline{x}^{(i)}). \tag{30}$$

Taking into account that the Jacobian matrix is constant, and the vector $\underline{f}(\underline{x}^{(i)})$ is known, the above mentioned matrix equation is linear.

4.1 Power Flow Solution by Using Newton Raphson Method

By starting form the power flow Eq. (20), the pathway of finding solution is to solve numerically the power flow problem. The recursive solution can be find by using NR method (28).

As it can be observed, the Jacobian matrix (27) $N \times N$ differentiations should be determined.

The power flow consists of determining the solutions of the real power and imaginary power equations by finding the unknown vector variables (voltage angles and magnitudes of them). In this case, the Jacobian matrix contains four derivatives submatrices, with specific dimension: $\underline{\mathrm{J}}^{\mathrm{P}\theta}[(N-1)\times(N-1)]$, $\underline{\mathrm{J}}^{\mathrm{Q}\theta}[(N-N_G)\times(N-1)]$, $\underline{\mathrm{J}}^{\mathrm{PV}}[(N-1)\times(N-N_G)]$, $\underline{\mathrm{J}}^{\mathrm{QV}}[(N-N_G)\times(N-N_G)]$.

Therefore, the four derivatives submatrices are outlined below [24]:

$$\underline{J} = \begin{bmatrix} \underline{J}^{P\varphi} & \underline{J}^{PU} \\ \underline{J}^{Q\varphi} & \underline{J}^{QU} \end{bmatrix} \tag{31}$$

or in the form

$$\underline{J} = \begin{bmatrix} \frac{\partial P_j}{\partial \varphi_k} & \frac{\partial P_j}{\partial |U_k|} \\ \frac{\partial \underline{Q}_j}{\partial \varphi_k} & \frac{\partial \underline{Q}_j}{\partial |U_k|} \end{bmatrix} \tag{32}$$

where subscript j is the bus number of the appropriate active power, and subscript k is the bus number for the appropriate voltage angle.

The Eq. (30) is multiplied in both sides by -1, obtaining the following equation

$$\underline{J}\Delta\underline{x}^{(i)} = -\underline{f}(\underline{x}^{(i)}) \tag{33}$$

From Eqs. (24), (25), and (31) the Eq. (33) is transformed in the following relationship:

$$\left[\begin{array}{c|c} \underline{J}^{P\varphi} & \underline{J}^{PU} \\ \underline{J}^{Q\varphi} & \underline{J}^{QU} \end{array}\right] \cdot \begin{bmatrix} \Delta\underline{\varphi} \\ \Delta|\underline{U}| \end{bmatrix} = -\begin{bmatrix} \Delta\underline{P} \\ \Delta\underline{Q} \end{bmatrix} \tag{34}$$

Due to the power system specific conditions, the reminded four derivatives submatrices can be expressed by using the power flow Eq. (20) as follows:

$$\underline{\mathbf{J}}^{\mathbf{P}\varphi} = \begin{bmatrix} J_{jj}^{P\varphi} & J_{jk}^{P\varphi} \\ J_{kj}^{P\varphi} & J_{kk}^{P\varphi} \end{bmatrix}, J_{kk}^{P\varphi} = J_{jj}^{P\varphi} = \frac{\partial P_j(\underline{x})}{\partial \varphi_j}, \; J_{jk}^{P\varphi} = J_{kj}^{P\varphi} = \frac{\partial P_j(\underline{x})}{\partial \varphi_k} \tag{35}$$

Knowing that the derivative of the simple *cos* function is –*sin*, the derivative of the simple *sin* function is *cos* function, and $\cos\varphi_{jj} = 1, \; \sin\varphi_{jj} = 0, \; J_{jj}^{P\varphi} = \frac{\partial P_j(\underline{x})}{\partial \varphi_j}$ or in the deducted form $J_{jj}^{P\varphi} = -Q_j(\underline{x}) - B_{jj}\left|\underline{U}_j\right|^2$, by combining (35) with power flow (20) the following results are obtained:

$$\underline{\mathbf{J}}^{\mathbf{P}\varphi} = \begin{bmatrix} -Q_j(\underline{x}) - B_{jj}\left|\underline{U}_j\right|^2 & \frac{\partial P_j(\underline{x})}{\partial \varphi_k} \\ \left|\underline{U}_j\right|\left|\underline{U}_k\right|\left(G_{jk}\sin\varphi_{jk} - B_{jk}\cos\varphi_{jk}\right) & -Q_j(\underline{x}) - B_{jj}\left|\underline{U}_j\right|^2 \end{bmatrix} \tag{36}$$

In the same manner, the $J_{jj}^{Q\varphi} = \frac{\partial Q_j(\underline{x})}{\partial \varphi_j} = P_j(\underline{x}) - G_{jj}\left|\underline{U}_j\right|^2$, and $J_{jk}^{Q\varphi} = \frac{\partial Q_j(\underline{x})}{\partial \varphi_k} = -\left|\underline{U}_j\right|\left|\underline{U}_k\right|\left(G_{jk}\cos\varphi_{jk} + B_{jk}\sin\varphi_{jk}\right)$ components of the $\underline{\mathbf{J}}^{\mathbf{Q}\varphi}$ submatrix are determined:

$$\underline{\mathbf{J}}^{\mathbf{Q}\varphi} = \begin{bmatrix} P_j(\underline{x}) - G_{jj}\left|\underline{U}_j\right|^2 & J_{jk}^{Q\varphi} \\ J_{jk}^{Q\varphi} & P_j(\underline{x}) - G_{jj}\left|\underline{U}_j\right|^2 \end{bmatrix}. \tag{37}$$

Taking into account that $J_{jj}^{PU} = \frac{\partial P_j(\underline{x})}{\partial U_j} = \frac{P_j(\underline{x})}{\left|\underline{U}_j\right|} + G_{jj}\left|\underline{U}_j\right|$, and $J_{jk}^{PU} = \frac{\partial P_j(\underline{x})}{\partial U_k} = \left|\underline{U}_j\right|\left(G_{jk}\cos\varphi_{jk} + B_{jk}\sin\varphi_{jk}\right)$, the components of the $\underline{\mathbf{J}}^{\mathbf{PU}}$ submatrix can be found:

$$\underline{\mathbf{J}}^{\mathbf{PU}} = \begin{bmatrix} J_{jj}^{PU} & \frac{\partial P_j(\underline{x})}{\partial U_k} \\ \left|\underline{U}_j\right|\left(G_{jk}\cos\varphi_{jk} + B_{jk}\sin\varphi_{jk}\right) & \frac{\partial P_j(\underline{x})}{\partial U_j} \end{bmatrix} \tag{38}$$

Similarly, $J_{jj}^{QU} = \frac{\partial Q_j(\underline{x})}{\partial U_j} = \frac{Q_j(\underline{x})}{\left|\underline{U}_j\right|} - B_{jj}\left|\underline{U}_j\right|$, and $J_{jk}^{QU} = \frac{\partial Q_j(\underline{x})}{\partial U_k}$ or $J_{jk}^{QU} == \left|\underline{U}_j\right|\left(G_{jk}\sin\varphi_{jk} - B_{jk}\sin\varphi_{jk}\right)$, the components of the $\underline{\mathbf{J}}^{\mathbf{QU}}$ submatrix are determined as follows:

$$\underline{\mathbf{J}}^{\mathbf{QU}} = \begin{bmatrix} J_{jj}^{QU} & J_{jk}^{QU} \\ \frac{\partial Q_j(\underline{x})}{\partial U_k} & \frac{\partial Q_j(\underline{x})}{\partial U_j} \end{bmatrix} \tag{39}$$

The algorithm of the NR applied to power system is implemented in the following steps:

1. Input:
   - The admittance of the studied power system
   - Specify for all type of busses the real power demand (scheduled power), P_{d}, and imaginary power demand, Q_{d}
   - For all PV buses the generated active power, P_{g}, and the voltage magnitude $|\underline{U}|$ should be known;
   - Specify the data for the swing bus: voltage magnitude $|\underline{U}|$, and angle $\varphi = 0°$;
2. Initial estimate of the solution at the first step of iteration $j = 1$. There are two steps:

   2.1 Set, for all type of busses, the initial voltage magnitude to 1 and the angle to 0° ($\mathrm{U_k} = 1.0 \angle 0°$). Solve the PFP P_k^0, Q_k^0.
   2.2 Set the solution of the previously studied network at step 2.1 as initial starting point.
3. Compute the active and reactive power injection from the power flow equations, at each node, according to Eqs. (23), and (20). Compare, at each node, the obtained values with the specified power demand (scheduled), and determine the difference of the power: $\Delta P_{\mathrm{k}} = P_{\mathrm{k}}(\mathrm{x}) - P_{\mathrm{k}}{}^0$, $\Delta Q_{\mathrm{k}} = Q_{\mathrm{k}}(\mathrm{x}) - \mathrm{n}Q_{\mathrm{k}}{}^0$ (for bus k).

$$\begin{bmatrix} \sum\limits_{j=\overline{1,N}} |\underline{U}_2|\left|\underline{U}_j\right|\left(G_{2j}\cos\varphi_{2j} + B_{2j}\sin\varphi_{2j}\right) - P_2 \\ \vdots \\ \sum\limits_{j=\overline{1,N}} |\underline{U}_N|\left|\underline{U}_j\right|\left(G_{Nj}\cos\varphi_{Nj} + B_{Nj}\sin\varphi_{Nj}\right) - P_N \\ - - - - - - - \\ \sum\limits_{j=\overline{1,N}} |\underline{U}_{N_G+1}|\left|\underline{U}_j\right|\left(G_{N_G+1,j}\sin\varphi_{N_G+1,j} + B_{N_G+1,j}\sin\varphi_{N_G+1,j}\right) - P_{N_G+1} \\ \vdots \\ \sum\limits_{j=\overline{1,N}} |\underline{U}_N|\left|\underline{U}_j\right|\left(G_{N,j}\sin\varphi_{Nj} + B_{N,j}\sin\varphi_{Nj}\right) - P_N \end{bmatrix} = \underline{0} \tag{40}$$

4. Choose the tolerance (ε) value in order to establish the precision of the solution: determine the norm of the power error (between the demand or scheduled power and the obtained power through the power flow equations): $|\Delta P_k| = \max\limits_k |\Delta P_k|$ and $|\Delta Q_k| = \max\limits_k |\Delta Q_k|$.
   The stopping criterion should be checked: if the norm for PQ, and PV buses $|\Delta \mathrm{P_k}| < \varepsilon_{\mathrm{r}}$, and for PQ buses the norm $|\Delta \mathrm{Q_k}| < \varepsilon_{\mathrm{i}}$ are fulfilled, retain P_{k}^0 and Q_{k}^0 as the solutions. The algorithm can STOP.
5. If the norms are higher than the tolerance (ε) value, the initial voltage magnitude and the angle should be changed incrementally such that the new value can be calculated as: $U_i^1 = U_i^0 + \Delta U_i$, and $\varphi_i^1 = \varphi_i^0 + \Delta\varphi_i$.
   The incremented value, ΔU_i, $\Delta\varphi_i$ are chosen as the errors $\Delta \mathrm{P_k}$, $\Delta \mathrm{Q_k}$ are vanishes (see the next step).

6. The better solution can be determined as:
   - Calculate $\underline{J}[\underline{x}^{(i)}]$ (Jacobian) or simply note by $\underline{J}^{(i)}$
   - Find the incremented vector value $\Delta\underline{x}^{(i)}$ from:

$$\underline{J}^{(i)}\Delta\underline{x}^{(i)} = -\begin{bmatrix}\underline{\Delta p}\\ \underline{\Delta Q}\end{bmatrix} \tag{41}$$

   - The new incremented solution is determined as a new vector: $\underline{x}^{(i+1)} = \underline{x}^{(i)} + \Delta\underline{x}^{(i)}$.
   - Repeat for the next increment: $i = i + 1$.

7. Stop.
   In real case, the reactive power should be maintained within the admissible limits. Therefore, at each iteration step, the reactive generation power limits of the PV bus should be checked. The above mentioned algorithm is augmented with two more steps:
8. The reactive generation power limits are verified
   (a) For PV buses, the maximum reactive or imaginary power value is set to $Q_{gk,max}$. If the limit is reached, the PV bus is changed to PQ bus.
   (b) For PQ buses: if both the reactive power and the magnitude voltage reaches the limits, change the bus into PV bus. If the reactive power attained the minimum value and the magnitude voltage is lower than the specified (set) value, the generator bus should be changed into the PV bus.
9. If the limits are not attained then the algorithm STOP. If the limits are violated, then changes the type of buses into adequate ones. The Jacobian matrix should be modified. The stop condition should be checked (step 4).

Taking into account the conjugate of the vector current flow between *j-k* buses, the apparent complex power flow is determined based on the relationship [27], and considering the numerical method for on-line power analysis [29, 30]:

$$\underline{S}_{jk} = \underline{U}_j \cdot \left[\left(\underline{U}_j - \underline{U}_k\right) y_{jk}\right]^* \tag{42}$$

4.2 NR Power Flow. Case Study

In order to solve the PFP through NR method, the 4 bus power system (Fig. 3) is taken into consideration, and the topology data are introduced as follows. The loads are connected to entire PS buses: B1–B4 [27]. The energy is produced by the generators

connected to B1 and B4 [27]. The known data consist of: I. *specified flow data*, and II. *data for the individually bus.*

I. The explanation for the *specified flow data* is as following:

- the impedance values (R, and X), filled in columns 2–3, as in specified flow data shown below;
- the calculated admittance value for each bus to bus data, as inverse of the corresponding side of the impedance values. The adequate results are filled in the columns 4, 5;
- shunt data (the end of each line), connection between the bus and ground is capacitive: corresponding charging reactive power and the equivalent per-unit admittance magnitude: values filled in columns (4, 5).

II. The data for the individual bus are considered as:

- the real and reactive power for generator buses (columns 2, 3), and load buses (columns 4, 5). The constant value of the generated imaginary power at the 1, 4 buses is not inserted in the table because is maintained at the constant value;
- voltage magnitude is unitary for a flat start (the initialization of the first iteration with 1 value). For a hot start, the first iteration value is the previously obtained solution;
- for a flat start the voltage angles are null, i.e. supposes the initialization of the first iteration with 0 value;

Remarks:

- For a PV bus, the generated real power and the magnitude voltage are specified (bus 4);
- The setting values of the generator buses are known and they are maintained at the specified constant values (the magnitude and phase of the voltage vector at the reference bus, i.e. $|\underline{U}| = 1$, $\varphi = 0°$);
- The generated reactive power is set to 0, magnitude $|\underline{U}|=1$ and maintained at this constant value;
- The reactive or imaginary load power is deducted from active load power with a 0,85 inductive power factor.

The above mentioned data can be introduced as inputs in Matlab programing language [31], or equivalent open source software Scilab [32], Octave [33].

I. The specified flow data:

(1)→(2)	0.01008	0.0504	3.8156	-19.078	10.25	0.05125
(1)→(3)	0.00744	0.0372	5.1696	-25.848	7.75	0.03875
(2)→(4)	0.00744	0.0372	5.1696	-25.848	7.75	0.03875
(3)→(4)	0.01272	0.0636	3.0237	-15.119	12.75	0.06375

II. The data for each bus consists of

1	0	0	50	30.99	1	0	1
2	0	0	170	105.35	1	0	2
3	0	0	200	123.94	1	0	2
4	318	0	80	49.58	1.02	0	3

Bus type: 1=slack(S); 2-Load(L)-inductive; 3-Voltage Controlled(VC).

The base values of the considered PS are the magnitude of the complex power $S_b = 100$ MVA, the voltage magnitude $U_b = 230$ kV.

Based on the specified input data above (I, II), the components of the admittance matrix of the network is properly determined, having the [4 × 4] size:

$Y_{bus} =$

Columns 1 through 3

1	8.9852 - 44.836i	-3.8156 + 19.078i	-5.1696 + 25.848i
2	-3.8156 + 19.078i	8.9852 - 44.836i	0 + 0i
3	-5.1696 + 25.848i	0 + 0i	8.1933 - 40.864i
4	0 + 0i	-5.1696 + 25.848i	-3.0237 + 15.119i

Column 4

0 + 0i
-5.1696 + 25.848i
-3.0237 + 15.119i
8.1933 - 40.864i

By applying the NR method, at the beginning NR algorithm the adequate tolerance should be chosen.

The calculated Jacobian at the first iteration, J(1), is:

J (1) =

0	0	0	-123.81	0
0	0	0	0	-113.24
0	0	0	73.821	43.179
-127.24	0	73.821	0	0
0	-115.55	43.179	0	0

with:

J11 =

```
    0    0    0
    0    0    0
    0    0    0
```

J12 =

```
  -123.81         0
       0   -113.24
   73.821    43.179
```

J21 =

```
  -127.24         0    73.821
       0   -115.55    43.179
```

J22 =

```
    0    0
    0    0
```

% difference of the active power at each bus:
$\Delta P_1 = P_1(\mathrm{x}) - P_1^0$,

deltap =

```
    0
-3.4175
-3.1541
 4.4001
```

% difference of the reactive power at each bus: $\Delta Q_1 = Q_1(\mathrm{x}) - Q_1^0$.

deltaq =

```
    0
-1.0535
-1.2394
```

The solution at the first iteration:
(a) Voltage Magnitude (per unit) at each bus

U =

```
       1
       1
       1
    1.02
```

(b) Phase angle (degrees) at each bus:

angles =

-0.97612
-1.8722
1.5231

By following the NR algorithm described in this chapter, after few iterations (N = 112), the numerical results of the power flow problem are obtained:

(a) The final numerical values of the Jacobian elements are obtained as:

J (N) =

44.327	0	-25.651	6.9721	0
0	39.609	-14.74	0	5.6932
-26.103	-15.094	41.196	-4.0459	-2.0984
-10.372	0	6.3048	42.22	0
0	-9.6932	3.8683	0	37.131

with the following subcomponent values at the final iteration:

J11 =

44.327	0	-25.651
0	39.609	-14.74
-26.103	-15.094	41.196

J12 =

6.9721	0
0	5.6932
-4.0459	-2.0984

J21 =

-10.372	0	6.3048
0	-9.6932	3.8683

J22 =

42.22	0
0	37.131

% difference of the active power at each bus: $\Delta P_N = P_N(x) - P_N^0$,

deltap(N) =

0
6.4393e-15
8.8818e-15
-1.1546e-14

% difference of the reactive power at each bus: $\Delta Q_{\mathrm{N}} = Q_{\mathrm{N}}(\mathrm{x}) - Q_{\mathrm{N}}^{0}$.

deltaq =

0
-2.2204e-16
-3.7748e-15

The solution of the PFP at the final iteration (with obtained 0.001 tolerance):
(b) the voltage magnitude at each bus (pu)

U =

1
0.98242
0.969
1.02

(c) the voltage angle at each bus (degree)

angle(N) =

0
-0.97612
-1.8722
1.5231

The compact form of the obtained power flow solution by using NR method is described below:

1	1	1	0	0	0	50	30.99
2	2	0.98	-0.97	0	0	170	105.35
3	2	0.969	-1.87	0	0	200	123.94
4	3	1.02	1.52	318	0	80	49.58

Bus type: 1=slack(S); 2-Load(L)-inductive; 3-Voltage Controlled(VC).

Conclusion:

The first six columns of the last resulted table contain the information of the individual bus from the considered four buses topology. In the columns (3) and (4) the solution of the load power flow is presented.

5 Fast Power Flow Method

The NR algorithm is a good choice for large power systems, being a robust algorithm; although the convergence speed depends on the initial estimate. For example, for an initial guess value far from the solution would slowly the path to the solution.

The fast decoupled (FD) methods for the power flow solutions increase the speed of finding the solution, and are accurate. The disadvantage of the FD methods is the decrease of the solution robustness.

The fast decoupled Newton Raphson power flow

The NR FDPF method allows the reduction the size of the linear equation systems at each iteration. Therefore, the memory allocation space is diminished.

In high voltage line transmission network the reactance is greater than resistance, therefore the R/X ratio is small. Usually, the phase angles of the line transmission do not exceed 25°.

The method of solving the FDPF is used in contingency calculations for electrical networks of any size and can be implemented efficiently on limited memory computing resources. The method can be applied also for the calculations of the normal operation of the networks, but especially for the studies of contingency evaluation on-line or off-line.

The method offers a particularly attractive combination of advantages over GN or NR methods, in terms of speed, reliability, simplicity and storage, for conventional load flow solutions.

The Jacobian matrix (31) can be inserted in the deducted incremental term from NR method:

$$\underline{J}\Delta\underline{x}^{(i)} = -\begin{bmatrix} \Delta\underline{P} \\ \Delta\underline{Q} \end{bmatrix}. \tag{43}$$

The active power flow depends mostly to the voltage phase angles, than of the magnitude voltage.

$$\frac{\partial P_j}{\partial |U_k|} << \frac{\partial P_j}{\partial \varphi_k}. \tag{44}$$

In other words, the sensitivity of active power to voltage magnitude is small

$$\frac{\partial P_j}{\partial |U_k|} \approx \underline{0}. \tag{45}$$

Opposite, the reactive power flow depends mostly to the magnitude voltage, than of the voltage phase angles:

$$\frac{\partial Q_j}{\partial \varphi_k} << \frac{\partial Q_j}{\partial |U_k|}, \tag{46}$$

i.e. the sensitivity of reactive power to voltage phase angle is small:

$$\frac{\partial \underline{Q}_j}{\partial \varphi_k} \approx \underline{0} \tag{47}$$

As a conclusion, the secondary diagonal of the Jacobian has null vectors, i.e. the matrix equation system is decoupled:

$$\begin{bmatrix} \underline{J}^{P\varphi} & | & \underline{0} \\ \underline{0} & | & \underline{J}^{QU} \end{bmatrix} \cdot \begin{bmatrix} \Delta\underline{\varphi} \\ \Delta|\underline{U}| \end{bmatrix} = -\begin{bmatrix} \Delta\underline{P} \\ \Delta\underline{Q} \end{bmatrix}. \tag{48}$$

The matrix multiplication conduct to the following matriceal equations:

$$\begin{aligned} \underline{\mathbf{J}}^{\mathbf{P}\varphi}\boldsymbol{\Delta}\underline{\varphi} &= -\Delta\underline{P} \\ \underline{\mathbf{J}}^{\mathbf{QU}}\Delta|\underline{U}| &= -\Delta\underline{Q} \end{aligned} \tag{49}$$

The significance of the above deducted matriceal equations is that both active (from the voltage magnitudes) and reactive power (from the voltage angles) are decoupled. Therefore, the independent power solutions can be obtained.

Moreover, for transmission the conductance is small and the angular differences are small, i.e.

$$cos(j_{ij}) \approx 1,\ sin(j_{ij}) \approx 0. \tag{50}$$

By considering the above mentioned practical remarks, result a constant Jacobian matrix. The solution of the system (49) is find by matrix triangularization (by avoiding the matrix inversion at each iteration).

The FD NR method reduces both the used computational memory and the calculus time.

The FD algorithm is based on the NR method, the only difference consists of modifying the 6th step of NR implementation as:

6. The better solution can be determined as:

– Calculate $\underline{\mathrm{J}}[\underline{\mathrm{x}}^{(\mathrm{i})}]$ (Jacobian) or simply note by $\underline{\mathrm{J}}^{(\mathrm{i})}$
– Find the incremented vector value $\Delta\underline{\mathrm{x}}^{(\mathrm{i})}$ by using LU decomposition:

$$\begin{aligned} \underline{\mathbf{J}}^{\mathbf{P}\varphi(i)}\boldsymbol{\Delta}\underline{\varphi}^{(i)} &= -\Delta\underline{P}^{(i)} \\ \underline{\mathbf{J}}^{\mathbf{QU}(i)}\Delta|\underline{U}|^{(i)} &= -\Delta\underline{Q}^{(i)} \end{aligned} \tag{51}$$

– The new incremented solution is determined as a new vector: $\underline{\mathrm{x}}^{(\mathrm{i}+1)} = \underline{\mathrm{x}}^{(\mathrm{i})} + \Delta\underline{\mathrm{x}}^{(\mathrm{i})}$.
– Repeat for the next increment: $i = i + 1$.

7. Stop.

The matrix triangularization increase the speed of finding the solution. The direction of finding the solution is performed by the Jacobian, thus the increased number

of iterations is necessary. Taking into account the above mentioned antagonism, the FDLF method is twice per iteration faster than with NR method.

From the point of view of accuracy, the FDLF has the same accuracy with the NR method.

6 The Gauss Seidel Method Applied to a Network

In order to solve a nonlinear equations system in iterative case, the GS is one of the method. This method suppose that there is a vector as solution to the system. This solution can be guess from the power load flow practical experience. The first iterate result is based on this first solution. The convergence rate is very sensitive to the adopted guess values. The procedure is going on up to fulfilling the prescribed accuracy condition [13, 27].

By assuming a network with only load buses, in which the active (P_p) and reactive power (Q_p) are known. The problem is to determine the magnitude voltage (U_p) and phase angle (φ) for each bus.

In rectangular form, the voltage phasor on p bus is described by:

$$\underline{U}_p = \left|\underline{U}_p\right| e^{j\varphi_p}. \tag{52}$$

The injected current into the bus p flowing into admittance is as follows Y_{pq} to the bus q is obtained by the following equation:

$$\underline{I}_p = \sum_{q=\overline{1,n}} Y_{pq} \left|\underline{U}_q\right| e^{j\varphi_q} \tag{53}$$

The conjugate of the complex apparent power can be expressed as:

$$P_p - jQ_p = \left|\underline{U}_p\right|^* \underline{I}_p \tag{54}$$

In order to solve the problem, the GS method is used.

The algorithm is as follows [27]:

I. **Data input:**

For each bus p, it is assumed the known value for the magnitude voltage and phase angle: $\left|\underline{U}_p\right|^0$ and φ_p^0 with $p = \overline{1,n}$.

II. **Relaxation:**

Starting from the initial data, the current injection on the first calculus iteration at each bus can be computed as

$$\underline{I}_1^{(1)} = (P_1 - jQ_1)/\left|\underline{U}_1^{(0)}\right|^* \tag{55}$$

On the other hand, the current injection into bus 1 at first iteration can be computed as:

$$I_1^1 = Y_{11}.\left|\underline{U}_1^{(1)}\right| + \sum_{q\neq 1} Y_{1q}.\left|\underline{U}_q^{(0)}\right| \tag{56}$$

From the last two equations, the magnitude of the voltage at bus 1 is computed by using:

$$\left|\underline{U}_1^{(1)}\right| = \left[\frac{P_1 - jQ_1}{\left|\underline{U}_1^{(0)}\right|^*} - \sum_{q=2}^{n} Y_{1q}\left|\underline{U}_q^{(0)}\right|\right] \cdot [Y_{11}]^{-1} \tag{57}$$

By proceeding as in the first step case, the second step follows the way:

$$\left|\underline{U}_2^{(1)}\right| = \left[\frac{P_2 - jQ_2}{\left|\underline{U}_2^{(0)}\right|^*} - Y_{21}\left|\underline{U}_1^{(1)}\right| - \sum_{q=3}^{n} Y_{2q}\left|\underline{U}_q^{(0)}\right|\right] \cdot [Y_{22}]^{-1} \tag{58}$$

The general form of the magnitude voltage at bus p, at the first iteration, is:

$$\left|\underline{U}_p^{(1)}\right| = \frac{P_p - jQ_p}{\left|\underline{U}_p^{(0)}\right|^* Y_{pp}} - \frac{1}{Y_{pp}}\sum_{q=1}^{p-1} Y_{pq}\left|\underline{U}_q^{(1)}\right| - \frac{1}{Y_{pp}}\sum_{q=p+1}^{n} Y_{pq}\left|\underline{U}_q^{(0)}\right| \tag{59}$$

By knowing the voltage magnitude at the first iterate on each bus, $\left|\underline{U}_p^{(1)}\right|$, the next iteration voltage magnitude $\left|\underline{U}_p^{(2)}\right|$ can be computed based on the second iteration currents I_p^2 and on the $\left|\underline{U}_p^{(1)}\right|$. This procedure continue until the voltages find the solution. This process is called *relaxation.*

III. **The stop condition (convergence issue):**

By choosing an admissible tolerance value, ε_p, the difference between the magnitude voltage vector values (of all buses) of two successive iterations (step $k+1$, and step k) is calculated, the program stops when the difference is less than the prescribed tolerance. In order to implement the stop condition the maximum difference value between applied to all vector components should satisfy the condition:

$$\max_p\left[\left|\left|\underline{U}_p^{(k+1)}\right| - \left|\underline{U}_p^{(k)}\right|\right|\right] < \varepsilon_p \tag{60}$$

The speed of the convergence is increased by using the *relaxation factor* $\mu \in (1, 2)$ when calculating the next iterate value:

$$\left|\underline{U}_p^{(k+1)}\right| = (1+\mu)\left|\underline{U}_p^{(k-1)}\right| + \mu\left|\underline{U}_p^{(k)}\right| \tag{61}$$

The benefit of the relaxation method is the low number of iterations. Therefore, a low memory capacity is requested. However, the main disadvantage of using Gauss Seidel method is the low speed of convergence. From this point of view, the NR method is more appropriate to be used in large power systems [33].

By considering *g*- generator bus, in the PV buses (generator), the active power injected into the studied network and the magnitude of the voltage are known. The problem consists of finding the reactive power (injected to the network or absorbed by the generator from the network) and the appropriate voltage phase angles. The final solution is provided in case of reactive power restrictions ($Q_g \in [Q_{\min}, Q_{\max}]$) if the following three steps are implemented [27]:

a. The magnitudes of the voltage at the generator bus at iteration k are calculated according to:

$$\underline{U}_g^{(k)} = \left|\underline{U}_g^{(k)}\right| e^{\varphi_g^k} \tag{62}$$

b. The reactive power at the generator bus can be determined as:

$$Q_g^k = \Im m\left\{\underline{S}_g^*\right\} \tag{63}$$

or more explicitly:

$$Q_g^k = \Im m\left\{\underline{U}_g\left(\sum_{q=1}^{n} Y_{gq}\underline{U}_q^{(k)}\right)^*\right\} \tag{64}$$

c. The reactive power limits should be checked. If this condition is fulfilled, the searching iterative process will goes on. Otherwise, the PV bus is considered as PQ bus or load bus. The process of finding the solution in iterative manner will continue up to attains the convergence condition.

7 GS Power Flow. Case Study

Taken into account the same power system as in NR method (4 buses power system), the known data consist of: I. *specified flow data*, and II. *data for the individual bus* [27].

I. *specified flow data*

(1) → (2)	0.01008	0.0504	3.8156	-19.078	10.25	0.05125
(1) → (3)	0.00744	0.0372	5.1696	-25.848	7.75	0.03875
(2) → (4)	0.00744	0.0372	5.1696	-25.848	7.75	0.03875
(3) → (4)	0.01272	0.0636	3.0237	-15.119	12.75	0.06375

II. *data for the individual bus*
The data for each bus consists of.

1	0	0	50	30.99	1	0	1
2	0	0	170	105.35	1	0	2
3	0	0	200	123.94	1	0	2
4	318	0	80	49.58	1.02	0	3

*Bus type: 1=slack; 2-load (inductive); 3-voltage controlled

Taking into consideration the above mentioned values of the line data, the admittance of the power network can be computed.

$Y_{bus} =$

Columns 1 through 3

	(1)	(2)	(3)
1	8.9852 - 44.836i	-3.8156 + 19.078i	-5.1696 + 25.848i
2	-3.8156 + 19.078i	8.9852 - 44.836i	**0 + 0i**
3	-5.1696 + 25.848i	**0 + 0i**	8.1933 - 40.864i
4	**0 + 0i**	-5.1696 + 25.848i	-3.0237 + 15.119i

Column 4

	(4)
1	**0 + 0i**
2	-5.1696 + 25.848i
3	-3.0237 + 15.119i
4	8.1933 - 40.864i

By considering the computed admittance, relaxation coefficient as $\mu = 1.6$ (acting as acceleration factor), and the considered precision as $\varepsilon = 0.1$, the GS power flow solution at the first iterate is:

(a) the reactive power (pu) —VoltageControlled bus—at the first step

$Q(1) =$

-0.3099
-1.0535
-1.2394
1.6541

(b) the voltage magnitude (pu) for each bus at the first step:

U(1) =

```
     1 +        0i
 0.9737 -   0.051706i
0.95395 -   0.066708i
 1.0199 -   0.010625i
```

(c) the voltage angle (degree) at each bus on the first iteration:

angle(1) =

```
      0
 -3.0397
 -4.0001
-0.59686
```

(d) %the active power for each bus (pu)

P(1) =

```
-0.5
-1.7
  -2
2.38
```

(e) the tolerance at the first step

tol(11) =

0.043721

After 11 iterations, the imposed tolerance of 0.001 was fulfilled (0.00084229). The solution of the PFP by using GS method is as follows:

(a) the reactive power (pu) -VoltageControlled bus -at the final step (11)

Q(11) =

```
-0.3099
-1.0535
-1.2394
 1.6541
```

(b) the voltage magnitude (pu) for each bus at the final step:

U (11) =

Bus no.	
1	1 + 0i
2	0.97507
3	0.95628
4	**1.02**

(c) the voltage angle (degree) at each bus on the first iteration:

angle(11) =

0
-3.0397
-4.0001
-0.59686

(d) the active power for each bus (pu) at the final step:

P(11) =

-0.5
-1.7
-2
2.38

(e) the tolerance at the final step

tol(11) =

0.00084229

The compact form of the results are depicted below:

1	1	1	0
2	2	0.97507	-3.0397
3	2	0.95628	-4.0001
4	3	1.02	-0.59686

For voltage controlled buses, by knowing the matrix admittance components and the specified voltage magnitudes, the reactive power is found.

Q(11) =

-0.3099
-1.0535
-1.2394
1.6541

The complex voltage vector is calculated at each iteration based on the delivered GN algorithm.

The components of the complex voltage vector are computed for each bus, starting from 1 to 4:

Bus no.	U =
1	1 + 0i
2	0.97507- 0.051706i
3	0.95628- 0.066708i
4	1.0199 - 0.010625i

8 Conclusion

The purpose of stationary power analysis is to obtain total information about the voltage vector (voltage amplitude and argument, voltage phase) at each bus of the power system depending on the load demand, the active power of the generator and the voltage conditions.

The PFP has as outputs the voltage vectors components (magnitudes and phase angles) at the existing buses into a power network. Based on these components, the active and reactive power flows can be computed in any branch of power network. Moreover, the total losses of the power network can be determined.

In this chapter numerical methods of the load flow in the interconnected systems are presented. The GS, NR, and FDLF methods are largely presented. The GS is an iterative method dedicated to solve the linear equations systems. The method derives from the Jacobi iterative method.

The GS method proved to be more advantageous than the Jacobi method, in terms of convergence speed. The Newton method, by its rapid convergence, is far superior to the GS method and nowadays is the main method for analysing the permanent operating regimes of the electricity distribution systems, or electric energy systems. However, the GS method is still used today, especially as a starting point when applying the NR algorithm or to determine the voltage vector components of the small disturbed power system found on the stability limit [34, 35].

The convergence rate of the Gauss-Seidel method is linear, so slow, the number of iterations is directly proportional to the number of buses in the system. Therefore, the method is used in small topology of the power systems. Due to the sparsity of the network matrix (diagonal dominant, symmetric positive definite), the number of operations in one iteration are fewest than in other power flow methods, being a simple method. The consequence is the reduction of time per iteration. In the GS method the convergence depends on the choice of the weak bus. The NR method (rectangular or polar coordinates) is superior to the GS method, the convergence is quadratic, very fast and the number of iterations is constant (does not depend on the size of the system), the accuracy of the solution being high. The NR method is used on large power systems. The computation time per iteration in the NR method is much higher than in the GS method (the Jacobian elements must be calculated). In the NR method the convergence is not sensitive to the choice of the weak bus and an approximate solution is obtained in a few iterations [1].

Although a large number of load flow methods are available in the literature, it has been observed that they are only the most popular Newton-Raphson and Fast Decoupled methods. The fast decoupled load flow is definitely superior to the Newton–Raphson method in terms of speed and storage (the computation time per iteration is 5 times smaller than in the case of the classical NR method) [15].

The sensitive point of the NR method is the starting value of the algorithm. Because the convergence is insensitive to certain parameters (network size, slack bus choice, etc.), the estimated starting solution value can lead to oscillations, the convergence to the solution fails. To avoid this problem, the load flow solution is often initialized with the GS method.

The specific solutions are provided through the case studies sections. The NR method is more robust than GS method. In order to increase the speed of finding the solution, the FS NR method is provided. The NR methods are more useful to the large power system. The numerical simulation results are provided, the methods being implemented in Matlab platform [31].

References

1. Kothari DP, Nagrath IJ (2006) Modern power system analysis, 3rd edn. McGraw Hill, New York
2. Andrei H, Gaiceanu M, Stanculescu M, Arama IN, Andrei PC (2019) Power systems connectivity and resiliency. In: Mahdavi Tabatabaei N, Najafi Ravadanegh S, Bizon N (eds) Power systems resilience. Power systems. Springer, Cham
3. Andrei H, Andrei PC, Gaiceanu M, Stanculescu M, Arama IN, Marinescu I (2019) Power systems recovery and restoration encounter with natural disaster and deliberate attacks. In: Mahdavi Tabatabaei N, Najafi Ravadanegh S, Bizon N (eds) Power systems resilience. Power systems. Springer, Cham
4. Andrei H, Gaiceanu M, Stanculescu M, Andrei PC, Buhosu R, Badea CA (2020) Energy storage systems in microgrid. In: Mahdavi Tabatabaei N, Kabalci E, Bizon N (eds) Microgrid architectures, control and protection methods. Power systems. Springer, Cham
5. Wongdet P, Leeton U, Marungsri B (2018) Line loss reduction by optimal location of battery energy storage system for the daily operation in microgrid with distributed generations. Int J Energy Convers (IRECON)
6. Saadat H (1999) Power system analysis. McGraw Hill
7. Rajicic D (2003) A power flow method suitable for solving OPF problems using genetic algorithms. In: The IEEE Region 8 EUROCON 2003 computer as a tool EUROCON-03
8. Kuehn C (2013) Warning signs for wave speed transitions of noisy Fisher–KPP invasion fronts. Theor Ecol
9. Karami A (2008) Radial basis function neural network for power system load-flow. Int J Electr Power Energy Syst
10. Tang M, Ma L, Zhang B, Wu C, Wang L, Li E, Chu X (2017) Optimal placement of dynamic reactive power compensation devices for improving immunity to commutation failure in multi-infeed HVDC systems. In: 2017 4th international conference on systems and informatics (ICSAI)
11. Gilbert GM, Bouchard DE, Chikhani AY (n.d.) A comparison of load flow analysis using DistFlow, Gauss-Seidel, and optimal load flow algorithms. In: Conference proceedings of IEEE Canadian conference on electrical and computer engineering (Cat. No.98TH8341). https://doi.org/10.1109/ccece.1998.685631

12. Le Nguyen H (1997) Newton-Raphson method in complex form [power system load flow analysis]. IEEE Trans Power Syst 12(3):1355–1359. https://doi.org/10.1109/59.630481
13. Stott B (1974) Review of load-flow calculation methods. Proc IEEE 62(7):916–929. https://doi.org/10.1109/proc.1974.9544
14. Stott B (1972) Decoupled newton load flow. IEEE Trans Power Appar Syst PAS-91(5), 1955–1959. https://doi.org/10.1109/tpas.1972.293524
15. Stott B, Alsac O (1974) Fast decoupled load flow. IEEE Trans Power Appar Syst PAS-93(3):859–869. https://doi.org/10.1109/tpas.1974.293985
16. Tiwari S, Ansari MA, Kumar K, Chaturvedi S, Singh M, Kumar S (2018) Load flow analysis of IEEE 14 bus system using ANN technique. In: 2018 international conference on sustainable energy, electronics, and computing systems (SEEMS). https://doi.org/10.1109/seems.2018.8687353
17. Alsulami WA, Kumar RS(2017) Artificial neural network based load flow solution of Saudi national grid. In: 2017 Saudi Arabia Smart Grid (SASG). https://doi.org/10.1109/sasg.2017.8356516
18. Dixit S, Srivastava L, Agnihotri (2006) Power flow analysis using fuzzy logic. In: 2006 IEEE power India conference
19. Zhou Y, Li Z, Zhou H, Li R (2016) The application of PSO in the power grid: A review. In: 35th Chinese control conference (CCC). https://doi.org/10.1109/chicc.2016.7554948
20. Yin X (n.d.) Application of genetic algorithms to multiple load flow solution problem in electrical power systems. In: Proceedings of 32nd IEEE conference on decision and control. https://doi.org/10.1109/cdc.1993.325915
21. Chakraborty D, Sharma CP, Das B, Abhishek K, Malakar T (2009) Distribution system load flow solution using Genetic Algorithm. In: 2009 international conference on power systems. https://doi.org/10.1109/icpws.2009.5442749
22. Zhang W (2008) Optimal multi-agent coordination under tree formation constraints. IEEE Trans Autom Control 4/2008
23. Lin H, Kalbarczyk ZT, Iyer RK (2019) RAINCOAT: randomization of network communication in power grid cyber infrastructure to mislead attackers. IEEE Trans Smart Grid
24. Afolabi O, Ali W, Cofie P, Fuller J, Obiomon P, Kolawole E (2015) Analysis of the load flow problem in power system planning studies. Energy and Power Eng 7:509–523. https://doi.org/10.4236/epe.2015.710048
25. https://electricgrids.engr.tamu.edu/electric-grid-test-cases/wscc-9-bus-system//home.engineering.iastate.edu/
26. Fernandez LM, Garcia CA, Jurado F (2010) Operating capability as a PQ/PV node of a direct-drive wind turbine based on a permanent magnet synchronous generator. Renew Energy 35(6):1308–1318. https://doi.org/10.1016/j.renene.2009.11.046
27. Grainger JJ, Stevenson WD (1994) Power system analysis. McGraw-Hill
28. Bompard E, Vaccaro A, Xie N, Torelli F (2013) Dynamic computing paradigm for comprehensive power flow analysis. IET Gener Trans Distrib.
29. Murty PSR (2010) Power systems analysis, butterworth-heinemann, 2017. In: Khaitan SK, McCalley JD, Raju M (2010) Numerical methods for on-line power system load flow analysis. Energy Syst 1(3):273–289.https://doi.org/10.1007/s12667-010-0013-6
30. Kyriakopoulos K (2008)Non-linear symbolic analysis for advanced program parallelization. IEEE Trans Parallel Distrib Syst
31. www.mathworks.com. Accessed 2019
32. www.scilab.org. Accessed 2019
33. https://www.gnu.org/software/octave/. Accessed 2019
34. https://www.artemis.cslab.ntua.gr
35. O'Kelly P (2013). Computer simulation of thermal plant operations. https://doi.org/10.1007/978-1-4614-4256-1

Numerical Methods in Selecting Location of Distributed Generation in Energy Network

Reza Effatnejad, Mahdi Hedayati, Keyvan Choopani, and Milad Chanddel

Abstract Integration of Distributed Generation (DG) units provide benefits to distribution systems. The presence of DG can affect various parameters of the distribution system such as reducing power losses and improving voltage profiles. But to maximize the profit from DG, the optimal location and the best amount of DG must be determined. Optimal placement and sizing of DG in the distribution network is a Complicated optimization problem. This paper presents a simple method for optimal sizing and optimal placement of distributed generators. This chapter of the book is about DG placement using numerical and innovative methods to solve the DG placement problem and comparing the two methods with each other. These methods are performed in a radial distribution system to minimize the total real power loss and improve the voltage profile. The proposed methods are tested on the standard IEEE 33-bus test system and the results are presented and compared with different approaches.

Keywords DG placement · Numerical methods · Heuristic methods · DG sizing · Power losses reduction · Voltage stability factor · Renewable energy

The original version of this chapter was revised: The incorrect tagging of the organisation name have been updated. The correction to this chapter is available at https://doi.org/10.1007/978-3-030-62191-9_36

R. Effatnejad (✉) · M. Hedayati · K. Choopani · M. Chanddel
Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran
e-mail: reza.efatnejad@kiau.ac.ir

M. Hedayati
e-mail: m.hedayati@kiau.ac.ir

K. Choopani
e-mail: chopani@jku.ac.ir

M. Chanddel
e-mail: m.chanddel@kiau.ac.ir

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_34

Abbreviations

CF	The total operating costs
$P_B^+(t)$	Battery discharge rate
EF	The total weight of polluting gases
$P_B^-(t)$	The amount of battery charge predicted
CG(t)	The energy purchased cost from the network
$P_i^{\min}$	Minimum allowable amount of micro-source production
Ppp	Power purchased from the network
$P_i^{\max}$	The maximum allowable amount of micro-source production ith
Ta_{pp}	Electricity purchase price from the network
$P_i^{RU\_MAX}$	Ramp up of micro-resources
RG(t)	The revenue generated through the main network
$P_i^{RD\_MAX}$	Ramp down of micro-resources
P_{SP}	Power sold to the network
SOC	State of charge
Ta_{SP}	The selling electricity price to the network
SOC_{max}	State of charge maximum
MC(i,t)	The cost of repairing and maintaining a small-scale resource
SOC_{min}	State of charge minimum
$K_{MC}(i)$	The cost of repairing and maintaining the power source
P_{BESS}^+	The amount of discharged power
P(i,t)	The power output of a small-scale source
P_{BESS}^-	The amount of charged power
SU(i,t)	The cost of restoration a small-scale resource
$R_{max(i,t)}$	Minimum required storage
Scost(i)	The cost of restoration a small iM source
R(i,t)	The amount of storage provided by the ith micro-source
u(i,t)	Binary variable to determine the on and off state of small scale sources
$Ploss$	power losses
FC_t^{DEG}	Diesel generator fuel cost
X_n	Decision in step n
FC_t^{MT}	MTs fuel costs
G_{best}	The best global position of the particle i in dimension j in t
$EM_{g,t}^{DEG}$	Exhaust gases (NO_2, CO_2, SO_2 and CO) proportional with fuel consumption (diesel generator)
v_{ij}^{t+1}	Particle velocity i in dimension j at time t
$EM_{g,t}^{MT}$	Exhaust gases (NO_2, CO_2, SO_2 and CO) proportional with fuel consumption (MT)
x_{ij}^t	The position of the particle i in the dimension j at time t
P(i,t)	The power output of a small-scale source
$p_{best,i}^t$	The best unique position of the particle i in the dimension j in the time t
VSF	Voltage stability index
ILq^k	Reactive power losses

ATC	Available Transfer Capacity
U	Controller vector
IVD^k	Voltage profile
L	Lagrangian shape
ϕ	a, b, and c phases
λ	the vector of the Lagrange
$\bar{V}_{\phi 0}$	Voltages in base node
S_n	n-state modes
$\bar{V}\phi_i^k$	Voltages at node i for Kth low voltage network arrangement
ILp^k	Active power losses
NN	Number of nodes
d	A direction vector to search (n × 1)
IC^k	The current capacity of the conductors
$ISC1^k$	Single phase short circuit
NL	Number of lines
$ISC3^k$	Three-phase short circuit

1 Introduction

The reliability and security of power systems increase power quality and voltage delivered to the load. The use of DGs, to achieve these goals is considered a lot of attention. In the past decades, the power required is generated centrally. With the beginning of the 20th century, power networks spread widely throughout a country. The regional power companies had a lot of power over the area they covered [1].

These companies monopolized the production, transmission, and distribution of power. With the advancement of technology, the electricity industry has become more competitive in different countries. Also, with the expansion of cities in different parts of a country and the construction of large industrial factories, the amount of power required to supply network loads increased. Transmission lines are highly capable of transmitting power. Therefore, creating such facilities will cost a lot of money for the relevant companies. All of the above factors have led to a wide-ranging change in how power networks are powered. DGs can be classified into two categories, renewable and non-renewable, according to Fig. 1.

DGs are not new, but they are economical, and their widespread use is a new concept. DGs refer to sources that have a production capacity from a few kW to about 30 MW. These units are placed in substations and distribution feeders near the loads. The use of distribution generators transforms a network from a centralized mode of production to a radically generated network. Figure 2 shows the structure of a network with centralized production and a network with DG.

Figure 3 shows the required power of the loads that are produced from DGs relative to the total network power in Europe and Denmark. As can be seen, the production of DG power in Europe has increased from 13.4% from 1998 to 2008 to 16.7%. Also,

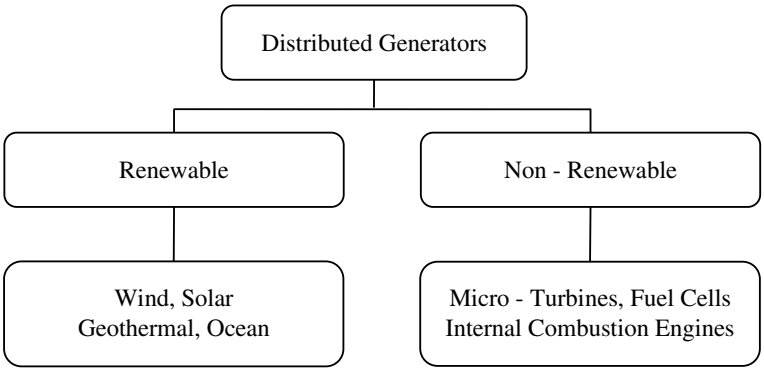


Fig. 1 Types of DGs

the capacity of DGs during these ten years in Denmark has increased from 11.7 to 28.7%.

In addition to the use of DG products in distribution networks, the development of simultaneous power and heat generation systems due to technological advances and reduced investment costs, as well as the economic benefits of this energy conversion process, led to a widespread approach by major energy consumers to use this technology is becoming. The high energy efficiency of these systems on a large scale, while reducing fuel consumption and environmental pollution by 50%, is a step towards the development of the electricity industry and the benefit of DGs. In these systems, the chemical energy of the fuel is released by a primary actuator (motor or turbine) and converted to mechanical power in the output shaft. Then the actuator shaft is coupled with a generator and electric power is generated. Due to the efficiency of the primary actuator, which is less than 50%, more than half of the fuel energy is dissipated in the form of heat, which can be converted into high-temperature and usable heat by placing suitable heat exchangers. DGs are one of the new trends in power systems to support increased demand. There is no clear definition of these resources. Different countries use different symbols, such as "embedded generation", "DG" and "decentralized generation". However, different definitions of DGs are provided by different organizations (IEEE, CIGRE, etc.), each of which has a specific aspect. This book uses the following reference definition [2]:

DGs are the sources of electrical energy connected to the power system that these sources are very close to the costumers and their production is much less than the centralized energy production sources.

To clarify the concept of DGs, Table 1 provides the dimensions of DGs [2].

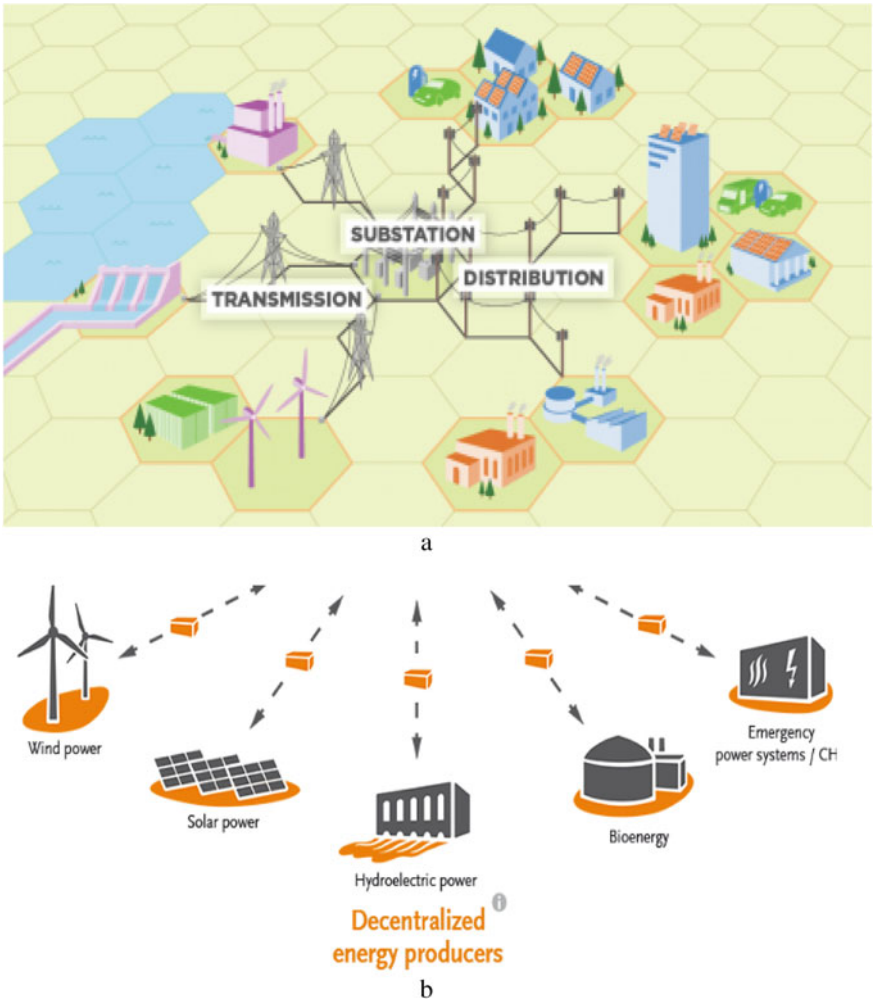


Fig. 2 **a** Centralized power system **b** Power system with DG

2 Types of DG Resources

In general, DGs can be categorized into two groups: DGs based on converters and the other DGs based on rotating machines. Converters are commonly used in DGs after DC processing to convert voltage to DC or AC. However, since the output value of the voltage and frequency must reach the nominal value, this voltage is first converted to DC and then to AC. In this section, we introduce the different types of DGs used in today's power systems (photovoltaic (PV) system, wind turbine (WT), fuel cell (FC), micro turbine(MT), as well as synchronous and asynchronous generators).

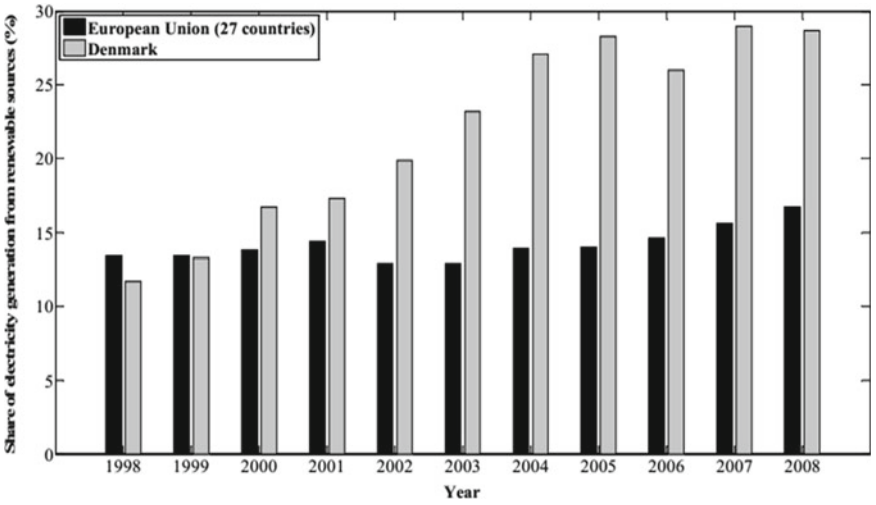


Fig. 3 The ratio of energy produced by DGs to total power

Table 1 Size of DGs [2]

Types of DGs	Size
Micro DGs	1 W to 5 kW
Small DGs	5 kW to 50 kW
Medium DGs	50 kW to 5 MW
Large DGs	5 MW to 30 MW

2.1 PV System

A PV system converts the amount of light absorbed by the sun into electrical energy. In this system, semiconductor materials are used in the structure of solar cells. When these semiconductors are exposed to sunlight, they convert the energy absorbed by the photon into electrical energy. Cells are fixed or varied in a specific arrangement so that they absorb the maximum amount of sunlight and subsequently produce the maximum power [3]. Environmentally, these systems have no harmful pollutants. These systems are very easy to use and require no fuel other than sunlight. The disadvantages of these systems include the need for a large space to be exposed to sunlight and the high initial cost. Figure 4 shows a schematic of a PV system with a brief detail [3].

The output voltage of PV systems is a DC voltage which is then converted to AC voltage using a converter.

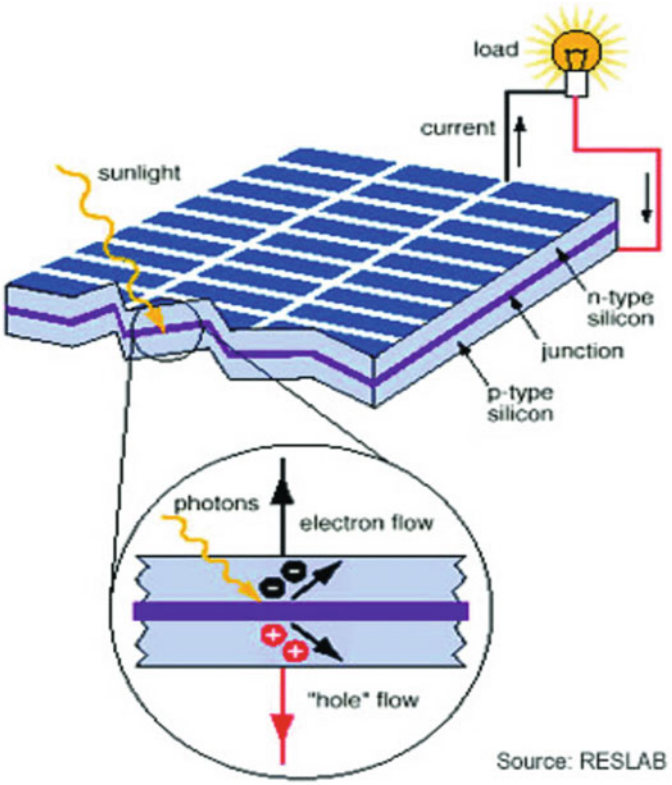


Fig. 4 Schematic of a PV system [3]

2.2 WTs

WTs convert wind energy into electrical energy. Since wind speed is not constant during the day and night, the amount of power generated by wind generators also varies. A general schematic of wind generators is given in Fig. 5 so that its main components are shown in this figure [3].

The performance of a wind generator can be explained in two steps. Initially, the rotor converts the amount of kinetic energy absorbed by the wind into mechanical torque on the turbine shaft. Then, the production system converts this torque into electrical energy. In today's wind generators, the output voltage is AC. The value of this voltage depends on the wind speed. Due to the variable wind speed, the generated AC voltage is first converted to DC voltage and then converted to the nominal AC voltage of the network using the converter. However, fixed-speed wind turbines are connected directly to the grid [3].

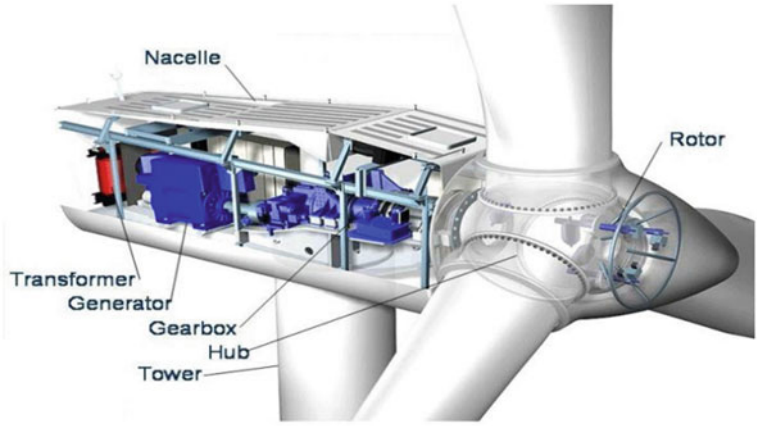


Fig. 5 Schematic of a wind generator [3]

2.3 FCs

FCs act like a battery and are charged by gas with high hydrogen. Oxygen is also taken from the air to cause a chemical reaction [3]. The FC produces a DC voltage from the reaction of hydrogen and oxygen with the help of electrolytic ions. Then, the DC voltage generated using a converter is converted to AC voltage and delivered to the network. Figure 6 shows a schematic of a FC.

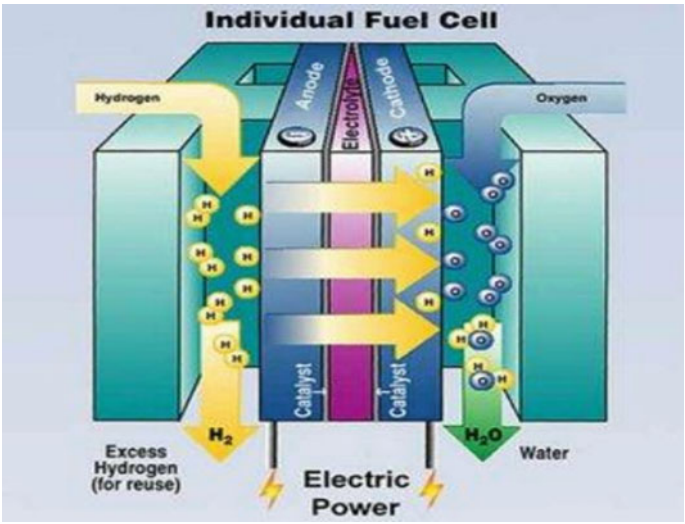


Fig. 6 Schematic of a FC [7]

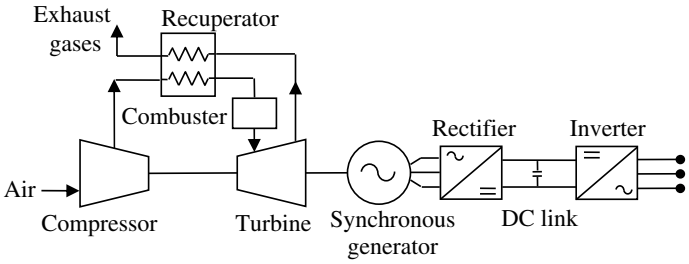


Fig. 7 Schematic of a MT

FCs also produce water along with electricity. The most important advantage of a FC is that it does not have a motive system, which increases reliability and does not produce noise. Also, these systems are able to operation with a wide range of FCs. On the other hand, its disadvantages include its impact on environmental pollution, its characteristic electrolytic aging, and the limited lifetime of FCs.

2.4 MTs

A MT uses the flow of a gas to convert thermal energy into mechanical energy. The combustible material (usually gas) is mixed with the air in the combustion chamber (air is pumped by a compressor). The result of this combination is the rotation of the rotor and power generation using a generator. Figure 7 shows a schematic of an MT [3]. The AC output voltage of the MTs cannot be connected directly to the main network. Therefore, it is first converted to DC voltage by a converter and then to AC voltage. One of the advantages of MTs is its clean performance (very low production of environmental pollutants) as well as its high efficiency. On the other hand, its disadvantages include its high maintenance costs and the lack of sufficient experience in this field [3].

2.5 Synchronous and Asynchronous Generators

The synchronous and asynchronous generators are electrical machines that convert mechanical energy into electrical energy. Asynchronous generators generate electricity when the shaft rotates faster than the synchronous frequency. The shaft of Asynchronous generators is rotated using a turbine or an engine. The power factor of asynchronous generators depends on the load and its value varies. The performance of such generators is much more important than that of synchronous generators.

Asynchronous generators require reactive power to create a magnetic field. At present, reactive power generation sources such as capacitive banks are used to

generate reactive power generators and loads at the relevant location. Therefore, asynchronous generators cannot be used as a backup generator when the system becomes an island [3].

Synchronous generators are used at a certain and synchronous speed. Unlike asynchronous generators, synchronous generators have a lag and variable power factor. Therefore, these generators can be used in power factor correction applications. A generator connected to a large network (connected to an infinite bus) has a small effect on the frequency and voltage of the network. Therefore, the rotor speed of these generators as well as the voltage value of the terminal is managed by the network.

In synchronous generators, a change in the excitation field changes the power factor. Also, a change in the amount of mechanical input power will change the output power of this generator. Therefore, when the synchronous generator is connected to an infinite bus, the over-excitation mode causes its power factor to be lagging, and also in the state the under excitation power factor is lead [3]. Therefore, excitation generators can be either reactive power generator or reactive power consumer. Today, synchronous generators are used in distribution systems as wind, hydro, and thermal generators. In general, such generators act as a power generator when placed at a low voltage level and have no role in controlling the network frequency. These generators are available in sizes from a few kW to a few MW [4].

3 The Effect of DGs on Power Networks

The sources of DGs used in distribution networks play a key role in the quality of power delivered to customers. However, the impact of such resources on distribution networks can have a positive or negative impact. In this section, the impact of such generators is evaluated and their behavior on an electrical system is examined. The difference in results between the presence of DGs and their absence in the network provides useful information to electricity companies and costumers. We now evaluate the impact of such resources on the different characteristics of a distribution network.

3.1 The Effect of DGs on Voltage Profile

Radial distribution systems set the grid voltage to an allowed level using load shift transformers (LTCs). In addition, line regulators are used on the distribution feeder as well as parallel capacitors in the feeders or along the lines to regulate the distribution network voltage. Installing DGs improves the voltage profile. Because the amount of the active and reactive power required is often generated at the load site, less power passes through the distribution network lines. Therefore, the amount of losses is reduced. Also, voltage drop decrease throughout the distribution network. However, in some cases, the installation of DGs on the network may also cause the voltage

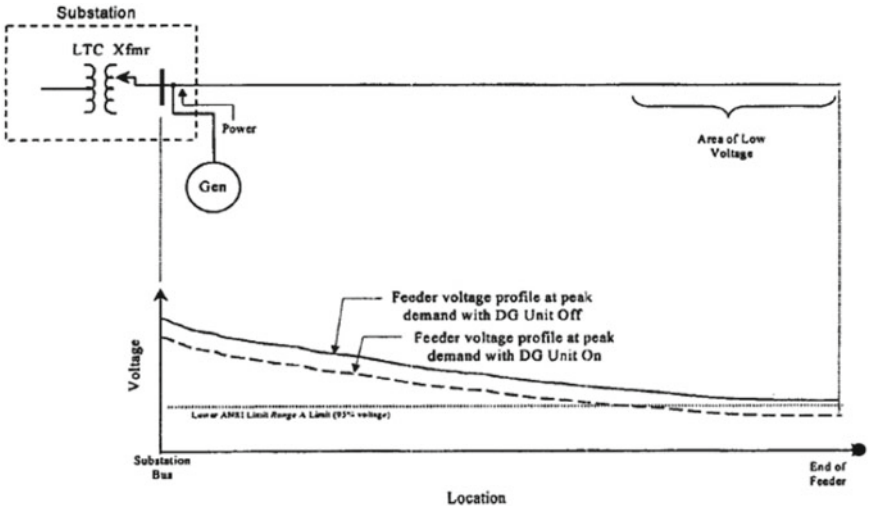


Fig. 8 Voltage profile of a feeder with and without DG sources [4]

profile to deteriorate, depending on the characteristics of the DG resources as well as the location of their installation. Figure 8 shows the voltage profile of a feeder with and without installation of DG. This figure shows DGs alongside an LTC equipped with a voltage reduction compensator. According to the figure, the voltage profile of the feeders in presence of DGs is higher when the DG resources are not installed in the distribution network. This is because the voltage regulator inserts less tap into the circuit and the voltage profile decreases. Since the DG sources are located after the voltage regulator, the amount of load seen by the voltage regulator decreases with the presence of these sources, and the reason is that the regulator regulates less voltage at the end of the feeder [4]. To solve this problem, DG can be transferred to before the voltage regulator.

Installing DG resources across a network may cause the voltage to rise too high due to too much active and reactive power to the network. For example, installing a DG source at a location several loads powered by a distribution transformer increases the secondary voltage of the transformer and therefore increases the voltage at the location several loads. This state occurs when the distribution transformer is installed at a point in the network where the initial voltage value is set near or above the range (for example, according to the US National Standards Institution, the high setting value for the base voltage 120 V is 126 V). Under normal system conditions, in the absence of DGs, the voltage on the load side is less than the initial voltage of the distribution transformer. The installation of DG can reverse the transmission power and thus increase the secondary voltage of the transformer on the costumers' side. When the DG size is less than 10 MW, its effect on the initial feeder voltage or the transformer's primary side is negated. However, when the DG size is greater than

10 MW, then a voltage regulation analysis is necessary to ensure that the customer voltage is within its allowable range [4].

3.2 The Effect of DG Resources on Network Power Losses

One of the most important effects of installing DGs is its impact on distribution network power losses. In general, the installation of DGs reduces network power losses. The reason for this is that part of the power consumption is often produced in their place by DGs and the amount of active and reactive power passing through the distribution lines is reduced. According to the reference [5], the installation of DGs in the network to reduce power losses is similar to the use of capacitive banks to reduce power losses. The difference between the two is that DGs inject the amount of active power as well as reactive power into the network, while the capacitive bank injects only reactive power into the network. In general, transmission network generators are used with a power factor of about 0.85 phases. Therefore, by installing DGs in the network and producing reactive power, network generators can be used with a higher power factor [6].

The optimal location of DGs to minimize network power losses can be calculated using load flow analysis software. For example, if the feeders have a large amount of power losses, installing one or more DG sources in different parts of the network will have a positive effect on the total feeder power losses. On the other hand, if a large-sized distribution generator is installed in some parts of the network, it may cause an increase in the number of adjacent lines. In this case, the thermal limit of these lines exceeds its allowable value and causes problems for the distribution network. The owner most of the DGs are customers. Therefore, network operators are not able to decide on the location of these resources. Therefore, it is assumed that the amount of power losses is reduced by installing DG at the load site.

3.3 The Effect of DGs on Harmonics

In general, a network waveform (for example, voltage and current) is not a pure sine wave and has some harmonics. This fact is shown in Figs. 9, 10 and 11.

In general, harmonics exist to somewhat extent in all power systems. DGs can be one of the harmonic producers in the network. Harmonics can be generated from the generator itself (eg, synchronous generators) or through electronic power equipment such as converters. In the converter section, one of the most important parts of harmonic production is the power converters of the controlled silicon rectifier (SCR) type, which produce a high level of harmonic currents. Today, converters use IGBT technology based on pulse width modulation to reduce harmonics. This new technology causes the output waveform closer to a pure sine waveform and has a small amount of harmonics, which is acceptable for the IEEE 1547-2003 standard.

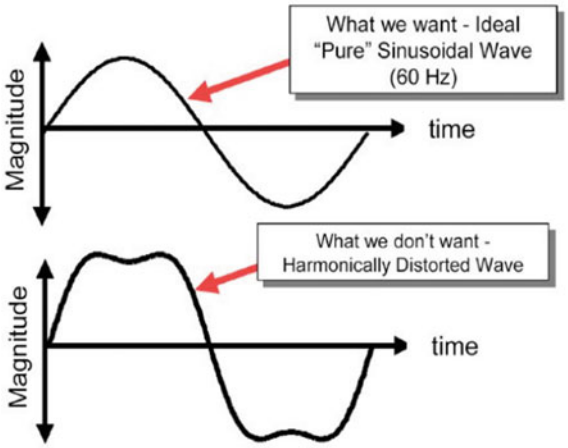


Fig. 9 Comparison between a pure sinusoidal waveform and a harmonic waveform destroyed by harmonics

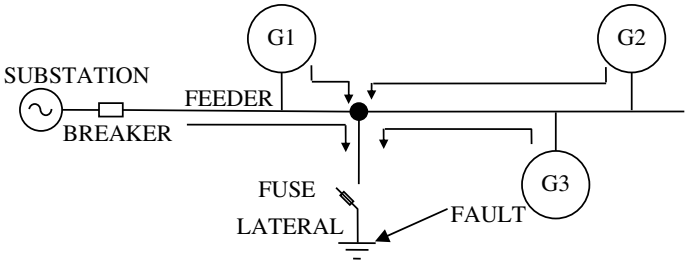


Fig. 10 The contribution of fault currents drawn from DG sources and the loss of coordination between the main feeder relay and the side fuse

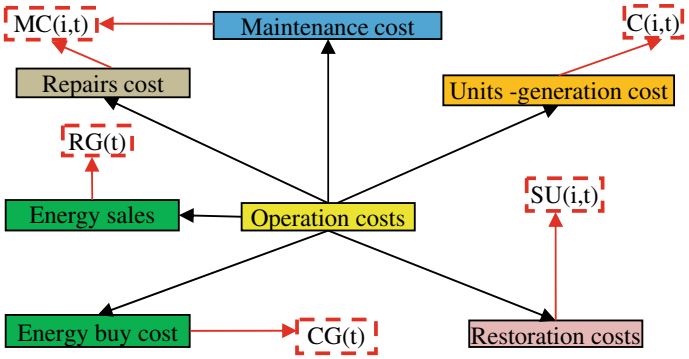


Fig. 11 Operation costs

Generator rotation is another factor in harmonic production. This harmonic value depends on many factors, including: the design of the generator winding (coil step), the nonlinearity of the core, the grounding of the generator, and other factors that cause harmonic production. The coil step is the most important factor in harmonic production. Comparing the types of winding steps in synchronous generators, the 2/3 winding step is the best step in terms of harmonics. Step 2/3 produces at least the third harmonic value. The third harmonic is the largest harmonic in the waveform of the system. On the other hand, the step 2/3 in generators has less impedance than the rest of the steps and therefore allows more harmonic current to pass through other parallel sources. Therefore, grounding the generator as well as using incremental transformers are good ways to limit production harmonies and prevent them from entering the network. The generator grounding scheme is a good choice for eliminating or reducing injectable harmonics into the network. In general, if we want to have a general comparison between harmonic production by DGs and other positive effects on the network, it can be seen that the harmonics produced by DGs do not cause much trouble for the network [5]. However, in some cases, the level of harmonic value generated by distribution generators may exceed its allowable value by the IEEE 519-1992 standard. These standard values are given in Table 2. These problems usually occur when they create resonance with capacitive banks or have equipment that is sensitive to harmonics. In the worst case, the additional heat from the harmonic generated by the generators of the DGs may cause damage to the network equipment. In this case, the DGs must be removed from the circuit.

Therefore, to design and install DGs in the network, special attention should be paid to the amount of harmonics produced. This harmonic value is evaluated according to the IEEE-519 standard. Therefore, the total harmonic value of the output should not exceed 5% and the value of each harmonic should not exceed 3%.

Table 2 Permitted Harmonic Rates for Distribution Generators by IEEE 519-1992 Standard

Harmonic order	Permissible value relative to the main harmonic (odd harmonics) (%)
Less than the eleventh harmonic	4
Harmonics between eleventh and seventeenth	2
Harmonics between seventeenth and twenty third	1.5
Harmonics between the twenty-third and thirty-fifth	0.6
Harmonics more thirty-fifth	0.3
Total harmonics	5

3.4 The Effect of DGs on the Short Circuit Level in the Network

The presence of DGs in the network affects the short circuit level of the network. When these resources are present in the network, the amount of fault current increases compared to the state without it [5]. The involvement of one of these sources in the occurrence of an error is not very high, but if a fault occurs, it will increase the fault current. If the number of small DGs in the network is high (or there are several large DGs in the network), the short circuit current level can be variable enough to eliminate the coordination between the relays and fuses of the network. The contribution of DGs in the faults that occur in the distribution network depends on various factors. These factors include the size of the DGs and its distance from the fault position. This contribution of the source of DGs in network faults can affect the reliability and security of distribution networks.

Therefore, DG sources have a significant impact on the fault current in the distribution network. In general, the contribution of harmonic generating sources of synchronous generators in the fault current is more than other types. After few cycles of fault, the contribution of fault currents from the DG sources related to the asynchronous and the synchronous generators is low, while after this time, the effect of DG resources on the fault current increases significantly.

3.5 The Effect of DGs on the Self-healing of the Distribution System

Self-healing refers to the specific ability of the smart grid that takes precautionary actions before the fault. It also performs fault location, isolation, and service restoration to minimize network damage after the fault. This process is performed with the help of communications infrastructure and remote control. Permanent fault in the distribution network, in addition to customer dissatisfaction, can cause heavy losses to power companies. Self-healing is the most important feature of the smart grid ensures continuity of electricity for the costumers. Also, with the increasing penetration of DG resources, the operating status of the network is changed in both normal and emergencies. In normal network operation mode, the goal is to minimize the costs of generation, operation, and pollution in the network. But in the self-healing mode, the goal is to feed maximum load in the faulty part. There are several ways to implement a self-healing scheme, including: Distribution system dividing into a number of MGs, Distributed generation resources, Multi-agent systems, Coordination storage and renewable resource, and Electric vehicles. In [56], the effect of DG on self-healing and reliability of the distribution system is investigated. In [57], the effect of storage and DG coordination on system self-healing is investigated. In [58, 59], a model is proposed to carry out Self-Healing in distribution level with consideration of combined gas and electrical network and with consideration of DGs types

include Micro-Turbines (MTs), and Combined Heat and Power (CHP) units Energy Storages (ESs).

4 The Economic Aspects of Low Voltage Networks, Especially Micro-grids (MGs)

Power and heat supply to areas not connected to the main network have led power system operators and engineers to move toward MGs. The supply of electricity and heat through low-voltage networks such as MGs, which are composed of DG sources such as energy storage systems and wind turbine systems, has led to a reduction in costs. Construction of a new line will also reduce pollution from conventional power plants. The use of CHP system in low voltage networks will lead to optimal use of heat loss. Smart grids, especially MGs, do not have transmission and distribution power losses compared to conventional systems because power generation takes place near the load [8]. The economic dimension of DG resources in low-voltage networks is so important that it can also challenge the economies of conventional systems. One very important point about MGs is, firstly, the high cost of construction and, secondly, the cost of operating such systems that use new resources and energy storage systems is not clear accurately. The economic aspects of low voltage networks with the presence of renewable resources and small-scale fossil resources can be mentioned as follows:

1. Economic dispatch
2. Minimize costs
3. Buying and selling energy.

Use and combine a variety of DG resources optimal economic conditions.

The differences in the economic dimensions of renewable resources and small-scale fossil resources with traditional systems, the following can be mentioned:

1. The main goal in traditional systems is to generate electricity, and heat generation is not the main goal. Therefore, they have low efficiency. While on smart grids, especially MGs, the CHP is a major goal to increase overall power efficiency. To further explain a small-scale gas turbine, the efficiency of a gas turbine in the form of a simple cycle is approximately 40%, but if the same gas turbine is equipped with a CHP system, the efficiency will be about 70 to 80%.

2. Environmental issues are not considered in traditional systems in the study of production and demand optimization. However, in smart grids, low voltage is not present with the presence of DG sources, and in studies, the discussion of pollutants in the objective function of planning is considered [8]. Minimizing the costs and pollutants caused by fossil fuels play a major role in the optimal use of power systems. To achieve these goals, the researchers moved toward renewable energy and small-scale DGs, and considered limitations and limitations for resources using fossil fuels to reduce environmental concerns caused by fossil fuels (Tables 3, 4, 5). In other words, minimizing costs and pollutants in operation studies are considered as a function of the short-term production planning target, which is stated in Eq. 1 [9, 10].

Table 3 Objective functions and constraints to economic programs [9–12]

Cost (energy sales and buy, restoration, repairs, maintenance, and fuel costs) and the amount of exhaust gases commensurate with fuel consumption (diesel generator and MT)		Description
Energy costs and the revenue generated through the main network	$CG(t) = Ta_{pp} \times P_{pp}$	$P_{pp:}$ Power purchased from the network. $Ta_{pp:}$ Electricity purchase price from the network
	$RG(t) = Ta_{SP} \times P_{SP}$	$P_{SP:}$ Power sold to the network. $Ta_{SP:}$ The selling electricity price to the network
The cost of repairing and maintaining a small-scale resource	$MC(i,t) = P(i,t), K_{MC}(i)$	$K_{MC}(i)$: The cost of repairing and maintaining the power source. P(i,t): The power output of a small-scale source
The cost of restoration a small-scale resource	$SU(i,t) \geq S\cos t(i).(u(i,t) - u(i,t-1))$ $SU(i,t) \geq 0$	Scost(i): The cost of restoration a small iM source u(i,t): Binary variable to determine the on and off state of small scale sources
Diesel generator fuel cost	$FC_t^{DEG} = \sum_{d=1}^{ND} (f1_d + f2_d P_{d,t}^{DEG} + f3_d (P_{d,t}^{DEG})^2)$	$f1_d$, $f2_d$, and $f3_d$ are determined by fuel consumption and prices characteristics
MTs fuel costs	$FC_t^{MT} = \sum_{d=1}^{ND} (f1_d + f2_d P_{d,t}^{MT} + f3_d (P_{d,t}^{MT})^2)$	$f1_d$, $f2_d$, and $f3_d$ are determined by fuel consumption and prices characteristics
Exhaust gases (NO2, CO2, SO2 and CO) proportional with fuel consumption (diesel generator and MT)	$EM_{g,t}^{DEG} = \sum_{d=1}^{ND} (\alpha_d^g + \beta_d^g P_{d,t}^{DEG} + \gamma_d^g (P_{d,t}^{DEG})^2)$	α_d^g, β_d^g, and γ_d^g according to the exhaust gases (NO_2, CO_2, SO_2 and CO) measured in different capacities are determined
	$EM_{g,t}^{MT} = \sum_{d=1}^{ND} (\alpha_d^g + \beta_d^g P_{d,t}^{MT} + \gamma_d^g (P_{d,t}^{MT})^2)$	
Constraints		Description
Power balance constraint	$(\sum_{i=1}^{I} P(i,t)) + P_B^+(t) - P_B^-(t) \geq Demand(t)$	P(i,t): The power output of a small-scale source $P_B^+(t)$: Battery discharge rate $P_B^-(t)$: The amount of battery charge predicted

(continued)

Table 3 (continued)

Constraints on the production of micro-resources	$P_i^{\min} \leq P(i,t) \leq P_i^{\max}$	$P_i^{\min}$: Minimum allowable amount of micro-source production $P_i^{\max}$: The maximum allowable amount of micro-source production ith
Ramps up and down of micro-resources	$P(i,t) - P(i,t-1) \leq P_i^{RU\_MAX}$	$P_i^{RU\_MAX}$: Ramp up of micro-resources
	$P(i,t-1) - P(i,t) \leq P_i^{RD\_MAX}$	$P_i^{RU\_MAX}$: Ramp down of micro-resources
Constraints of charging and discharging states energy storage system (battery)	$SOC(t) = SOC(t-1) + P_{BESS}^{-}(t) - P_{BESS}^{+}(t)$	SOC: State of charge SOC_{max}: State of charge maximum SOC_{min}: State of charge minimum P_{BESS}^{+}: The amount of discharged power P_{BESS}^{-}: The amount of charged power
	$SOC_{min} \leq SOC(t) \leq SOC\max$	
	$P_{BESS}^{-}(t) \leq P_{BESS}^{-}\_\max$	
	$P_{BESS}^{+}(t) \leq P_{BESS}^{+}\_\max$	
	$X(t) + Y(t) \leq 1; \quad X, Y \in \{0,1\}$	
Constraint of system storage	$\sum_{i=1}^{I} R(i,t) \geq R\max(t)$	$R_{max(i,t)}$:Minimum required storage R(i,t): The amount of storage provided by the ith micro-source

$$Minimize\, F = \{CF, EF\} \tag{1}$$

In the above relation, CF and EF express the total operating costs and the total weight of polluting gases, respectively. Operating costs are shown in Fig. 2. Equation 2 represents the cost function, which consists of the cost of repairs, maintenance costs, the cost of selling electricity, the income from energy, the cost of start-up energy production sources such as WTs and PVs, and the cost of producing units. Table 1 is given to further explain the Eq. 1 and express the constraint of the problem [9, 10].

$$CF = \sum_{t=1}^{T} \sum_{i=1}^{I} [MC(i,t) + C(i,t) + SU(i,t)] + \sum_{t=1}^{T} [-RG(t) + CG(t)] \tag{2}$$

Table 4 Effective indicators in low voltage networks to accommodate and determine the size of DGs [15]

Indicator (conductor current capacity, three-phase short circuit, single-phase short circuit and voltage profile)		Description
Voltage profile	$IVD^k = 1 - \max\left(\frac{\lvert\bar{V}_{\phi 0}\rvert - \lvert\bar{V}\phi_i^k\rvert}{\lvert\bar{V}_{\phi 0}\rvert}\right)_{i=1}^{NN-1}$	ϕ: a, b, and c phases $\bar{V}_{\phi 0}$: Voltages in base node $\bar{V}\phi_i^k$: Voltages at node i for Kth low voltage network arrangement NN: Number of nodes
	$IVR^k = 1 - \frac{\sum_{i=1}^{NN-1} \max\left(\left\lVert \frac{\bar{V}_{\phi_i^k} - \bar{V}_{\phi_i^k \min}}{\bar{V}_{\phi_i^k \min}} \right\rVert\right)}{NN-1}$	
The current capacity of the conductors	$IC^k = 1 - \max\left(\frac{\lvert\bar{J}_{\phi_m^k}\rvert}{CC_{\phi_m}}, \frac{\lvert\bar{J}_{n_m^k}\rvert}{CC_{n_m}}\right)_{m=1}^{NL}$	NL: Number of lines $\bar{J}_{\phi_m^k}$ and $\bar{J}_{n_m^k}$: Indicates the currents of the m branch in kth of the low-voltage network arrangement $CC_{\emptyset m}$ and CC_{nm} Indicates the current capacity of the conductors
Three-phase short circuit	$ISC3^k = 1 - \frac{\max\left(\frac{I_{SC}abc_i^k}{I_{SC}abc_i^0}\right)}{\frac{I_{SC}abc_*^k}{I_{SC}abc_*^0}}$	$I_{SC}abc_i^k$: The short-circuit current of three phases in node i for Kth low voltage network arrangement $I_{SC}abc_i^0$: Indicates the amount of short-circuit current of three phases in node i in the network without the presence of MTs $I_{SC}abc_*^k$: Indicates the maximum amount of short-circuit current of three phases in the system for kth low voltage network arrangement $I_{SC}abc_*^0$: Indicates the maximum amount of short-circuit three-phase current in the system without the presence of MTs
Single phase short circuit	$ISC1^k = 1 - \frac{\max\left(\frac{I_{SC}\varphi_i^k}{I_{SC}\varphi_i^0}\right)}{\frac{I_{SC*}^k}{I_{SC*}^0}}$	$I_{SC}^{\phi_i^k}$: Indicates the short-circuit single-phase at node i for kth of low voltage network arrangement $I_{SC}^{\phi_i^0}$: Indicates the short-circuit single-phase current in node i without the presence of MTs I_{SC*}^k: Indicates the maximum amount of single-phase short-circuit current in the system for kth low voltage network arrangement I_{SC*}^0: Indicates the maximum amount of single-phase short-circuit current in the system without the presence of MTs

(continued)

Table 4 (continued)

Indicator (conductor current capacity, three-phase short circuit, single-phase short circuit and voltage profile)		Description
Active and reactive power losses	$ILp^k = 1 - \frac{\mathrm{Re}\langle Losses^k \rangle}{\mathrm{Re}\langle Losses^0 \rangle}$	Losses^k: Indicates the total power losses for the kth low voltage network arrangement. Losses^0 : Indicates the apparent power losses without the presence of MTs
	$ILq^k = 1 - \frac{\mathrm{Im}\langle Losses^k \rangle}{\mathrm{Im}\langle Losses^0 \rangle}$	

Table 5 Results for the studied methods

Methods	Bus number selected for placement	Optimal amount of active power (MW)	Power losses (kW)	Exection time (sec)
PSO	13 24 30	1.36 1.4 1.45	58.4	3.156
GAPSO	13 24 30	1.22 1.4 1.45	55.2	4.6
MINLP	13 24 30	1.12 1.4 1.47	53.2	0.9

5 Placement of DG Resources

5.1 Introduction

In recent decades, changes in the structure of the electricity industry, as well as the privatization of the industry, have taken place in some countries. During this time, the electricity industry has undergone fundamental changes in terms of management and ownership due to increased efficiency and encouragement of investors, so that different parts of it, including production, transmission, and distribution, have been separated to create a suitable competitive environment. These changes, on the one hand, and factors such as environmental pollution, problems with the construction of new transmission lines and other problems, as well as technological advances in economics, and the construction of small-scale production units, on the other, have led to a desire to increase the use of small power units. Local power at the distribution voltage level is by renewable energy sources such as WT, PV, FC, MT, as well as a combination of these in distributed networks. With the presence of a combination of DGs in the distribution networks, they act as active networks, and therefore such networks are called active distribution networks.

Due to the increasing use of DG resources in distribution networks, determining their proper location is a basic need and the proper use of these resources in the network will have many benefits, including reducing losses and improving voltage profiles. Also, despite the energy resources in the distribution networks, in the event of an outage of the main source, the power supply will be partially possible in the form of island performance. Obviously, with the use of optimal islands, the number of blackouts will decrease. On the other hand, determining the boundaries of the islands will depend on the location and capacity of the DG units in the network. If the location of the DGs is not determined correctly, not only will there be no benefit. It worsens network else parameters. Therefore, the most important challenge in DG projects is to determine the location of these resources in distribution networks. To increase efficiency and reduce the negative effects of installing DG resources, location can be done with the aim of improving one parameter or more than one parameter. Single-target deployment greatly improves the desired parameter, but does not have much effect on other network parameters and sometimes destroys them. In contrast, multi-purpose locating helps to improve all parameters at the same time.

The most important step in using DG in distribution networks is their optimal location and capacity. This is important because it reduces costs associated with losses, reliability, and the cost of building production units. Also, the placement of DG units in distribution networks, along with other possible options for the development of the distribution network, such as the construction or strengthening of lines and substations, has been considered. On the other hand, uncertainty is one of the most important factors in increasing risk in the planning of power systems so that not consider this parameter leads to significant economic and technical losses.

5.2 Problem Modeling

In this section, we introduce the objective functions and the limitations that must be considered in the DG placement problem. Among the objective functions that can be considered in the problem of placement are.

Minimizing active power losses, minimizing energy losses, minimizing SAIDI, minimizing costs, minimizing voltage deviations, maximize DG capacity, maximize profits, maximize rate profits to costs, maximize loading limit voltage and improve power quality index. In the following, we will deal with some of the important objective functions that have been used in the literature on this issue.

5.2.1 Power Losses Index

Power Losses are always a concern for the electricity industry due to the waste of energy and resources, and its reduction is a priority for the industry's programs and research activities. The distribution of DG in the distribution network has a great impact on power losses. This effect varies depending on the size and location of the

DG installation. With proper placement of DG resources, network power losses can be significantly reduced. Since distribution networks with DG are active networks, the performance and control of the network are very interesting. Although there are several issues with the distribution network's performance, there are also major technical implications due to the influence of renewable DGs, such as network power losses, voltage stability, and network security discussed here.

Distribution networks are usually radially structured to reduce the complexity of protection. Assessing and reducing network power losses is essential to increase system efficiency. The average power losses per year can be calculated as follows [13]:

$$Ploss = \sum_{i=1}^{N_b} \frac{r_i \left(\left(P_{D,i+1} \right)^2 + \left(Q_{D,i+1} \right)^2 \right)}{\left(V_{i+1} \right)^2} \tag{3}$$

Here $Q_{D,i+1}$ and $P_{D,i+1}$ require the active and reactive power in the end bus receiving the $-\mathrm{i} + 1$. V_{i+1} is the voltage range at the end bus of the receiver $-\mathrm{i} + 1$.The resistance of the completed transmission line r_i to bus $-\mathrm{i} + 1$, $\mathrm{N_b}$ is the sum of the number of lines in the system.

5.2.2 Network Development Time Delay Index

Due to the development of population centers and the increase in electricity consumption, as well as a reliable supply of energy to customers, the development and strengthening of electricity distribution networks are inevitable. Distribution companies make significant investments in this section annually. If the load demand in the network exceeds the maximum allowable load, network reinforcement must be performed. In existing networks at the time of reaching some technical constraints such as maximum network transmission capacity or maximum voltage drop (due to increased load), network reinforcement is performed. Where DG is connected to the power supply, the demand for feeder pure peak (demand less than the capacity assigned to DG) reduces and releases the transmission capacity in its upstream branches to meet the demand for a new peak. The time required for development postpones. Therefore, the release of power transmission capacity in the distribution network can be considered as an indicator to delay the time of network development. This indicator is called the available transmission capacity and is indicated by the symbol (ATC) and is defined as the following relation for each branch i of the network. The margin of certainty of the transmission and storage capacity of the lines has been omitted [14].

$$ATC_i = P_{i,\max} - P_i \tag{4}$$

In the above relationship, P_i is the power passing through branch i (branch with end node i) and $P_{i,\max}$ is the maximum allowable power passing through branch i.

5.2.3 Voltage Stability Index

Voltage stability indicators are used to assess the level of voltage stability of buses in the transmission or distribution network. These tools are very quick and effective for calculating the offline stability of bus voltages. The voltage stability index (VSF) for each bus -i + 1 in the time interval t can be expressed as [13].

$$VSF_{i+1} = (2V_{i+1} - V_i) \tag{5}$$

The average voltage stability for the entire distribution network can be calculated as follows.

$$VSF = \frac{\sum_{i=2}^{N_b} VSF_{i+1}}{(N_b - 1)} \tag{6}$$

5.2.4 Constraints

(A) The power passing through each of the branches should not exceed the allowable power.

$$P_i \leq P_{i,\max} \tag{7}$$

(B) The voltage must remain within the allowable range.

$$V_{i,\min} \leq V_i \leq V_{i,\max h} \tag{8}$$

(C) The one-way current flow constraint is as follows.

$$(ATC)_{i,n} = \frac{P_{i,\max} - P_i}{P_i} \leq 1 \tag{9}$$

(D) The voltage of each bus shall not exceed its permissible limits.

$$V^{\min} \leq \sqrt{\sum h \left(V_i^{(h)}\right)^2} \leq V^{\max} \tag{10}$$

The general indicators used in various articles to discuss DG placement and measurement are summarized in the table below.

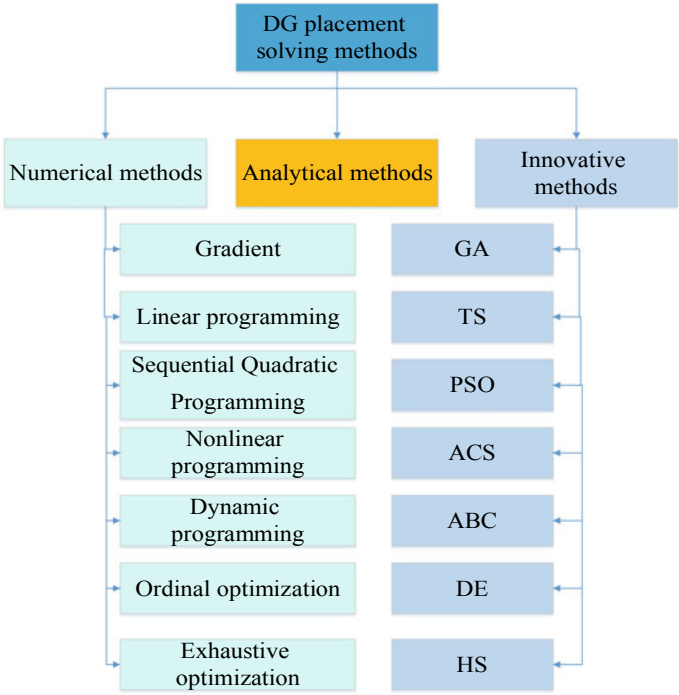


Fig. 12 Types of problem solving methods

5.3 Solution Methods

In terms of problem solving method, DG placement is divided into three categories, which are shown in Fig. 12 [16].

5.3.1 Analytical Methods

In this method, DG is installed as much as 2/3 input power to the feeder. Its installation location is 2/3 Feeder. In this method, it is assumed that the load is evenly distributed throughout the feeder. In [17–21], this method has been used to solve the placement problem.

5.3.2 Numerical Methods

In numerical methods, mathematical modeling is used to solve the DG placement problem. Among the methods used to solve these problems can be divided into the following categories. In [16], the linear planning method is also introduced for DG

placement, if this method cannot be done and only the optimal size of DG can be determined.

Gradient Search Method

This method is mostly used in mesh networks. In [22, 23], this method has been used to solve the DG placement problem. Suppose buses i… k are a set of whole n buses. A set of c resource units in these buses can use any technology. Our goal is to determine the amount of injection in each of these nodes for maximum profit. Therefore, the u controller vector can be expressed as follows [22].

$$U = \{P_1 \dots P_k\}, \quad i, k \in n \tag{11}$$

And the total injections should be expressed as an equality limit as follows:

$$P_1 + \cdots + P_k = C \tag{12}$$

To minimize power losses, the f index is obtained as follows:

$$f = \sum_{j=1}^{n} P_i \tag{13}$$

The formulation of this issue can be summarized as follows:

$$Minimize \ f = \sum_{j=1}^{n} P_i \tag{14}$$

and

$$g = 0 \tag{15}$$

The mode vector x is as follows:

$$x = \delta_2 \dots \delta_3 \dots \delta_n \ V_2, V_3 \dots P_1 \tag{16}$$

where $\delta_2 \dots \delta_3 \dots \delta_n$ are the angles of the buses 2 to n (δ_1 is the angle in the reference bus and its value is zero.), V_2, V_3 indicate the voltage in the uncontrolled buses of the voltage. Note that the state vector of equation and the set of equations g must include the power injection bus P_1, since f is containing P_1. The reduced gradient method can be used to solve this problem. In this method, the Lagrangian shape is as follows:

$$L = f + \lambda^T g \tag{17}$$

Where λ is the vector of the Lagrange and g coefficients including the grid equations. To minimize:

$$\frac{\partial L}{\partial X} = \frac{\partial f}{\partial X} + \lambda^T \frac{\partial g}{\partial X} \tag{18}$$

In the reduced gradient method, an initial assumption is made for the vector U and used to calculate λ. The Lagrangian coefficients for the specified values of X and U in the repetition of k are determined as follows:

$$\lambda^k = -\left[\frac{\partial g}{\partial X}\right]^{T-1} \left[\frac{\partial f}{\partial X}\right]_{X^{k\cup K}} \tag{19}$$

The Lagrangian gradient is determined by U as shown below [22]:

$$\frac{\partial L}{\partial U} = \frac{\partial f}{\partial U} + \left[\frac{\partial g}{\partial X}\right] \lambda^k \tag{20}$$

The control vector in k + 1 repetition is updated using the following equation [22]:

$$U^{k+1} = U^k - A\left[\frac{\partial L}{\partial U}\right]_{X^k\, U^k\, \lambda^k} \tag{21}$$

A is a constant. But instead of the reduced gradient method, the second method is usually used because the reduced gradient method does not converge for positive injections. In the second method, the vector z is expressed as follows [22]:

$$z = [X, U, \lambda]^T \tag{22}$$

Starting from the initial assumptions for z in the repetition of k + 1 is updated as follows [22].

$$z^{k+1} = z^k - \left[H\left(z^k\right)\right]^{-1} g\left(z^k\right) \tag{23}$$

The Hsian H (z) matrix is determined as follows [22]:

$$H_{1j} = \frac{\partial^2 L}{\partial_{z_1}\, \partial_{z_j}} \tag{24}$$

Second-Order Sequential Planning

In [25, 26], this method has been used to solve the DG placement problem without using fault level constraints. Conventional SQP is a numerical optimization technique

that solves a finite minimization problem by solving a second-order planning sub problem (QP) in each main iteration to obtain a vector to search for a new d. For search, along with a suitable step size of the α scalar, it forms the next approximate solution point that can be used to start repeating another original SQP. Under the QP sub problem, a first-order objective function with real value is minimized that is subject to linear equality and inequality constraints. The QP sub-problem in repetition k is formulated using the second-order Taylor extension in approximating the SQP objective function. The Taylor expansion is formulated in the linearization of the SQP equality and inequality constraints at the asymmetrical point. The QP sub problem is formulated as follows [26]:

$$\begin{aligned} &Minimize\ X \in R^n\ \ P_{Losses}\left(X^{(k)}\right) + \nabla P_{Losses}\left(X^{(k)}\right)^t d + \frac{1}{2} d^T H^{(k)} d \\ &\tilde{h}^k(d) : h\left(X^{(k)}\right) + \nabla h^t\left(X^{(k)}\right) d = 0 \\ &\tilde{g}^k(d) : g\left(X^{(k)}\right) + \nabla g^t\left(X^{(k)}\right) d \le 0 \\ &X^l \le X \\ &\le X^u \end{aligned} \tag{25}$$

where

$\nabla P_{Losses}\left(X^{(k)}\right)^t$:The gradient of the objective function at the point $X^{(k)}$.

d: A direction vector to search (n × 1).

$H^{(k)}$ A symmetric matrix of Hessian (n × n) at the point $X^{(k)}$.

$\tilde{h}^k$: First order Taylor expansion of equality constraints.

$\nabla h^t\left(X^{(k)}\right)$: A Jacobine matrix (n × m) of equality constraints.

$\tilde{g}^k$: Second order Taylor expansion of nonequality constraints.

$\nabla g^t\left(X^{(k)}\right)$: Jacobin matrix (n × p) of inequality constraints.

n,m and p: The number of equallity and non-equallity equations, respectively.

The SQP method uses the Lagrangian coefficient method for the general optimization problem constrained in Eq. 15. First defining the Lagrangian function problem at an approximate solution point $X^{(k)}$. Then apply the optimization conditions of the first order Karush-Kan-Tucker (KKT) to the Lagrange function; Finally, Newton's method is used in Lagrangian function gradients to solve unknown variables. The KKT is the first order that the optimal conditions for the nonlinear set of Lagrangian gradient equations are shown in Eq. 26 [26].

$$\begin{bmatrix} \nabla_X L(X, \lambda, \beta) \\ h(X) \\ g_A(X) \end{bmatrix}^{(k)} = 0 \tag{26}$$

The λ and β column vectors are of the Lagrangian equivalence and inequality coefficients. Newton's method is used to solve the above Eq. 16. Using Taylor's first-order extension at a hypothetical solution point $(X^{(k)}, \lambda^{(k)}, \beta^{(k)})$ to estimate

$(X^*, \lambda^*, \beta^*)$ Newton's KKT solution can be expressed as follows:

$$\begin{bmatrix} \nabla^2_{XX} L(X,\lambda,\beta) & \nabla h(x) & \nabla g_A(X) \\ \nabla h^t(X) & 0 & 0 \\ \nabla g^t_A(X) & 0 & 0 \end{bmatrix}^{(k)} \begin{bmatrix} d^{(k)} \\ \lambda^{(k+1)} \\ \beta^{(k+1)} \end{bmatrix} = - \begin{bmatrix} \nabla f_{RPL}(X) \\ h(x) \\ g_A(X) \end{bmatrix}^{(k)} \tag{27}$$

By solving the KKT Newton system, the solution $\begin{bmatrix} d^{(k)} \\ \lambda^{(k+1)} \\ \beta^{(k+1)} \end{bmatrix}$ is of the QP subset. The new vector obtained for the search provides new values for the Lagrangian coefficients to be used in the next iteration. The Hessian matrix of the Lagrangian function does not need to be calculated in each subsequent iteration. Instead, using the method (Go-Fletcher-Goldfrab-Shano) is approximated and updated in almost every major iteration [24]. Due to the resilience and effectiveness of SQP in controlling constraints on nonlinear optimization issues, it is used in many commercial optimization software [25].

Nonlinear Planning

The most widely used method in DG placement literature is the nonlinear planning method [27–30]. In [31], the proposed method for locating and sizing DGs is a multi-objective NLP problem with three objective functions, which include the number of DG units, power losses, and Voltage Stability Margin (VSM). The details of the functions and limitations of the proposed method are as follows.

Less DGs are preferred than many small DGs. Placement more DGs causes the distribution system in many buses to have problems with short circuit levels and requires a rescheduling of the protection coordinates. Therefore, minimizing the number of DGs is considered as the first objective function. The number of DGs is an integer variable, and this can create the proposed MINLP problem, which requires more calculations and also makes it more difficult to converge than an NLP problem. To prevent the proposed method from becoming a MINLP problem, a new method is used to calculate the number of DGs [31].

$$ndg = \sum_{i \in SB} \frac{P_{DGi}}{P_{DGi} + \varepsilon} \tag{28}$$

If a DG is assigned to bus i while repeating the optimization problem, P_{DGi} is non-volatile, and in this case the effect of ε is ignored and the fraction in (34.28) is very close to the unit. On the other hand, if no DG is assigned to bus i, the fraction will be zero. As a result, the ndg result in (34.28) shows the number of DGs in the network. Of course, although ndg is not the exact number of DG units as an integer, the minimum number of DGs is obtained by minimizing ndg. The second objective function is the active power losses, which is usually significant in distribution

systems. Using an AC load flow model, branch power losses are different from total output and total loads [31]:

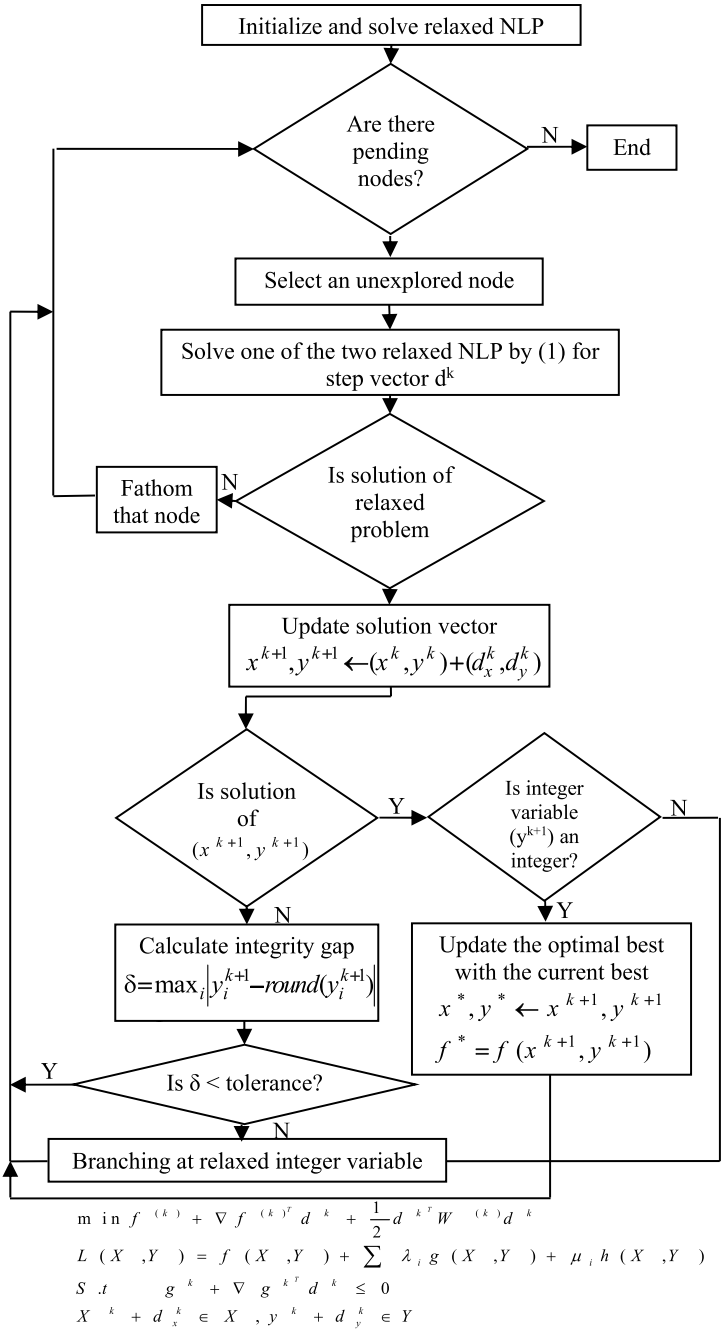
$$P_{Loss} = \sum_{i \in SB} PG_i - PD_{total} \tag{29}$$

Voltage stability has also been used as a objective function after placement. Fuzzy methods have been used in this article to determine the weight of the target functions. Also, all the limitations of the system (load flow, voltage, power passing through lines, etc.) are considered in this article [31].

Mixed Integer Nonlinear Programming

In [32], the MINLP model is used for optimal placement of different types of DG units. This study presents a Mixed Integer Nonlinear Programming (MINLP) approach to determine the optimal position and number of DGs in the hybrid market. For the optimal location of DGs, the most suitable area is first identified based on the actual node price and the sensitivity index to the loss of active power as an economic and operational criterion. After identifying the appropriate area, the MINLP method has been used to determine the location and optimal number of DGs in the area. The nonlinear optimization approach includes minimizing total fuel cost from conventional sources and DG, as well as minimizing line losses in the network. The pattern of active and reactive prices has been achieved, reducing line losses and saving fuel costs. Optimal DG penetration can provide both economic and operational benefits.

In a large regional power grid, identifying the most suitable area for DG penetration is important. Therefore, to find the most suitable area, economic and operational aspects must be considered. Node cost is an important indicator of the MW unit injection price per node and congestion in the transmission network. Sensitivity to line power losses provides information on the pattern of power losses in the transmission network with the power injection unit in each node. To get the most suitable area, the change in the price of the node in each bus and the sensitivity of line power losses have been used as economic and operational criteria. After identifying the most suitable area based on the price of the node and the sensitivity of the line power losses, the MINLP method can be used to find the optimal location and the number of DG in suitable area. The proposed MINLP model in [32] has been solved by the simultaneous use of GAMS and MATLAB software. The load flow equations have been solved in MATLAB software and its answers are transferred to GAMS software so that the rest of the modeling can be done by GAMS software. In the proposed work, an approach based on integrated second-order programming integration (SQP) and branch and border (BAB) techniques has been adopted to solve the MINLP formulation. This algorithm has better computational performance and strong convergence property [33]. The steps of the algorithm are as follows and the flowchart is shown in Fig. 13.



$$\min f^{(k)} + \nabla f^{(k)T} d^k + \frac{1}{2} d^{kT} W^{(k)} d^k$$

$$L(X, Y) = f(X, Y) + \sum \lambda_i g(X, Y) + \mu_i h(X, Y) \quad (1)$$

$$S.t \quad g^k + \nabla g^{kT} d^k \leq 0$$

$$X^k + d_x^k \in X, \; y^k + d_y^k \in Y$$

Fig. 13 Flowchart for MINLP-based algorithm [32]

Dynamic Planning

The optimization problem in this method is modeled in several ways. One of the advantages of this method is that in different modes, the load can be modeled more accurately [34]. In most practical problems, successive programs must be performed at different times to solve a problem. Problems that are solved by successive decisions are called sequential decision problems. Dynamic planning is a type of multi-step decision-making problem that is an efficient mathematical method for studying and optimizing multi-step sequential decision problems [35]. The main steps of this algorithm are:

- The problem is divided into stages. For each step, a decision policy is needed. On the other hand, each step represents a part of the problem that needs to be decided. The number of steps in DG allocation is equal to the number of candidate locations for DG installation. Decision-making procedures at each stage include reducing power losses and improving reliability.
- Each step involves related situations. In the present study, the number of DGs with a certain capacity that can be allocated in the mentioned stages will be considered as an issue. At each stage, the current state will be transferred to the next state by decision.
- Independent policies for the remaining steps can be pursued by understanding the current situation. In general, to optimize dynamic planning, current status information transfers all the information needed from previous behaviors that are needed to identify optimized policies from the current state to another.
- The problem will be solved by finding the optimized policy for each case of the last case, which is named as the return solution. The answer to this step is obvious because the process is followed by the destination.
- Optimized policy for all states of the "n" stage can be determined by a return function, assuming that the optimized policy is defined for all "n + 1" stages.
- The solution is applied using the return function from one step to the previous step that runs from the end. At each stage, the optimized policy will be specified for all states of that stage. Finally, the optimized policy for the first stage will be specified.

$$
\begin{aligned}
f_n(S_n, X_n) &= Z_n(X_n) + f^*_{n+1}(S_n - X_n) \\
f_n(S_n, X_n) &= Z_n(X_n) + Max \sum_{L=n+1}^{Nloc} Z_L X_l \; f^*_{n+1}(S_n - X_n) \\
f_n(S_n, X_n) &= Max\, f_n(S_n, X_n) \; and \sum_{L=n+1}^{Nloc} X_l \\
&= S_n
\end{aligned}
\tag{30}
$$

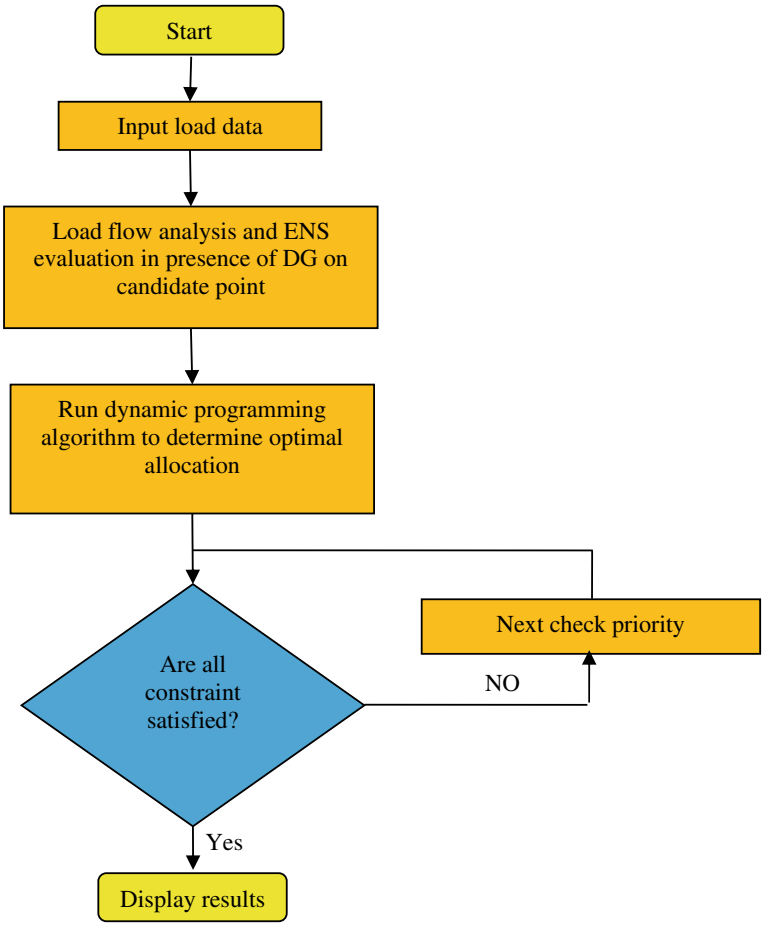


Fig. 14 Optimal DG placement method based on dynamic planning method [34]

where are n-state modes;
f^*_{n+1}Optimized function in step n + 1.
X_nDecision in step n.

The proposed algorithm is based on dynamic planning for DG placement and can be extended to different distribution networks. The general diagram of the algorithm followed for DG placement is shown in Fig. 14.

Sequential Optimization

In [36], this method has been used to solve the DG placement problem. This paper discusses the compromise between reducing power losses and maximizing DG as a

target function. Searching for an accurate global solution to the problem of DG placement and sizing requires a great deal of computational time. For example, selecting 9 DG locations from 68 potential candidates requires the following evaluation:

$$\binom{68}{9} = \frac{68!}{9!(68-9)!} \cong 4.93 \times 10^{10} \tag{31}$$

with about 0.15 seconds per OPF solution on a 1.73 GHz Core-Duo processor, the evaluation of the above combinations takes more than 200 years in a row. OO theory provides a possible framework for reducing search space and computational effort in ranking different options. The OO algorithm for finding the predetermined number of DGs can be summarized as follows:

- Select N design (i.e. combinations of DG locations) from the search space so that at least one of them is at the top of the α percentage of all solutions with a probability level of P. The normal values of α and P are 0.1 and 99%, respectively.
- Evaluate and order N designs using a raw and easy-to-calculate model.
- Calculate the number of best examples in N designs that have at least one real good alternative, such as a case above 50, with a probability level equal to AP (K = 1). 0.95.
- Evaluate each sample using an accurate OPF model and determine the appropriate solution.

Comprehensive Optimization

In [37–39], a comprehensive optimization method has been used to solve the DG placement problem by considering the load change. In power losses analysis, different types of customers are modeled, where a load of different types of customers varies during the day. Therefore, system power losses vary as a function of time. The issue of non-linear load flow is that the variable load over time makes it more complex. Therefore, the optimization problem involves a nonlinear model with voltage, current, and time-variable load limitations. The time variable load is interpreted as the change in the hourly load so that 24 loading conditions are modeled to show a daily load pattern. A comprehensive search of existing parts of the system is used to assess the reduction in power losses that can be achieved by adding DGs. The location of the DG is considered only in the separate devices of the sections.

In a search, DG locations are evaluated only at the beginning or at the separate input device of each section. The method for determining the optimal location for DG placement in the system is as follows.

(1) Perform load search statistics on the system and draw the load curve.

(2) Divide the load curve in different load windows so that the load conditions in each window are relatively constant.

(3) Adjust the DG output for different load windows to prevent the feedback from returning to the post.

(4) Do a comprehensive search for the load status of each window with the corresponding DG output and get the optimal locations based on two criteria - minimum power losses and minimum SAIDI for comparison. There are several reasons why a comprehensive search is used instead of the optimization algorithm. First, comprehensive search programming is easy. Second, a comprehensive search is done quickly. Third, this method ensures that global optimization is achieved. Due to the nonlinearity of the constraint problem, some optimization approaches may be trapped at a local optimal point.

5.3.3 Innovative Methods

Particle Swarm Optimization

Particle Swarm Optimization (PSO) has been used in [40–42] to DG placement. In [43], the Global Best PSO (gbest) algorithm is a method in which the position of each particle is affected by the best particle in all population particles. In this method, each particle has a specific position X_i and a certain speed V_i in the research space. The best unique position is called $P_{best,i}$ in the research space. It is proportional to the position in the research space that the position of this particle is the best value for minimizing (maximizing) the objective function compared to the previous values of this particle. Also, the best position of the particles among all the particles is called the optimal global position and is indicated by G_{best}. The following relationships show how the values of each particle and the total global value are updated. Considering a minimization problem, the unique position $P_{best,i}$ of each particle and the time of the next time step, $t + 1$, are calculated as follows:

$$P_{best,i}^{t} = \begin{cases} P_{best,i}^{t} \; if \; f(x_i^{t+1}) > P_{best,i}^{t} \\ x_i^{t+1} \; if \; f(x_i^{t+1}) \leq P_{best,i}^{t} \end{cases} \tag{32}$$

In this equation, f is a function of compatibility. The value of the best global position G_{best} in the time step t is calculated as follows:

$$G_{best} = \min\{P_{best,i}^{t}\}, where \; i \in [1, \ldots, n] \text{ and } n > 1 \tag{33}$$

Therefore, the value of the best possible value is obtained by comparing all the particles. Since the value has a direct relationship with the values obtained for the best location of each particle, so the initial values for selecting the position of each particle at the beginning of the process are very important. For the gbest PSO method, calculate the particle velocity i as follows:

$$v_{ij}^{t+1} = v_{ij}^{t} + c_1 r_{1j}^{t}[p_{best,i}^{t} - x_{ij}^{t}] + c_2 r_{2j}^{t}[G_{best} - x_{ij}^{t}] \tag{34}$$

In this equation:
v_{ij}^{t+1}: Particle velocity i in dimension j at time t.
x_{ij}^{t}: The position of the particle i in the dimension j at time t.
$p_{best,i}^{t}$: The best unique position of the particle i in the dimension j in the time t.
G_{best}: The best global position of the particle i in dimension j in t.
$c_1$andc_2: As cognitive and social parameters, respectively.
r_{1j}^{t} and r_{2j}^{t}: Random numbers of uniform distribution of U (0.1) at time t.

In this section, DG placement is examined using the PSO algorithm. According to the explanations given for the two methods of gbest PSO and lbest PSO, the gbest PSO method is used in the replacement. In addition to using the PSO algorithm, the forward-backward sweep or Newton Raphson methods are also used to load flow on the network in each iteration. If the purpose of the placement is to sizing and placement in a distribution network simultaneously, the particles used in the algorithm take the size and position simultaneously. It should be noted that in the PSO algorithm, each particle is not a single value but contains a matrix with the required number of dimensions. In the following, the proposed algorithm will be explained in more detail.

Step 1: Enter the distribution network information along with the PSO algorithm parameters. In the first step, the network information is examined, including network load information, line intersection matrix, lines resistance and reactance, etc. Also, PSO algorithm parameters such as the number of repetitions, number of particles, and values of c_1, c_2,$r_1$andr_2 are determined.

Step 2: Consider an initial random value for the velocity and position of all the particles. Set the counter to value 1. In this step, the proposed algorithm particles are randomly selected in a certain range. Since the particle value indicates the size and value of DG, therefore, the initial values are assigned to an acceptable value. Also in this step, the value of the counter that shows the same number of repetitions will be set to the value $k = 1$ or the same number of repetitions.

Step 3: Compare the value of k with the maximum value of the setting. In this step, the value of k, which represents the same number of repetitions, is compared to its maximum value. Processing will continue if the value of k is less than or equal to the maximum value, but when the value of k is greater than the maximum value of the setting, the processing is completed and the optimal position and size of the DGs in the distribution network are delivered to output.

Step 4: Implement load flow, obtain voltage, voltage deviation, power losses, etc. In this step, using the load flow and other available information, the amount of network bus voltage, voltage deviation, the amount of power passing through the lines, the amount of line power losses and other network information will be obtained.

Step 5: Evaluate the defined objective function, calculate and save Pbest and Gbest. In this step, the value of the defined objective function is obtained in each iteration, and the Pbest and Gbest values are obtained for each particle based on the value obtained for the objective function. In this research, the objective function is defined as follows:

$$Minimize \sum_{i=1}^{nline} P_{Loss,i} \tag{35}$$

where i, the line index, nline, the number of network lines, and the loss $P_{Loss,i}$ of line i is shown. Given the above relationship, the main goal is to minimize total network losses.

Step 6: Update the position and speed of the particles. In this section, according to the values obtained in each iteration, the value of the velocity and position of the particles for the next iteration is displayed. The update of these values is based on the relationships expressed for the velocity and position of the particle. After the update, the number of counters will increase by one and the processing will be repeated. The figure shows the proposed algorithm for placement using the gbest PSO algorithm. The above steps are described in the Fig. 15.

Hybrid GAPSO Optimization Technique

In [55], a new methodology, hybrid GAPSO (HGAPSO), is developed to optimization of an off-grid hybrid energy system (HES). Since standard particle swarm optimization (PSO) algorithm suffers from premature convergence due to low diversity, and genetic algorithm (GA) suffers from a low convergence speed, in this study modification strategies are used in GAs and PSO algorithms to achieve the properties of higher capacity of global convergence and the faster efficiency of searching. The superiority of HGAPSO algorithm over GAs and PSO algorithms is in terms of convergence generations and computation time. The HGAPSO algorithm combines the advantages of swarm intelligence of the PSO algorithm and the natural selection mechanism of the genetic algorithm to increase the number of highly evaluated agents at each iteration step. We see how the proposed algorithm works in Fig. 16.

DG placement is investigated by other intelligent algorithms such as GA [44–48], Tabu Search [49, 50], Ant Clooney [51], Artificial Bee Colony [52], Differential Evaluation [53] and Harmony Search [54] algorithms.

Example In the 33 bus distribution network shown in Figs. 17 and 18, do the DG placement problem with PSO, GAPSO, and MINLP methods. For better analysis, also specify the DG sizing system. The objective function for optimal placement and sizing is to reduce power losses. Also, after solving the problem, compare the voltages of each bus in both ways. Each DG can have a capacity of 0.5 to 1.5 MW. After determining the DG placement and sizing, check the system specifications in terms of power losses, execution time and voltage.

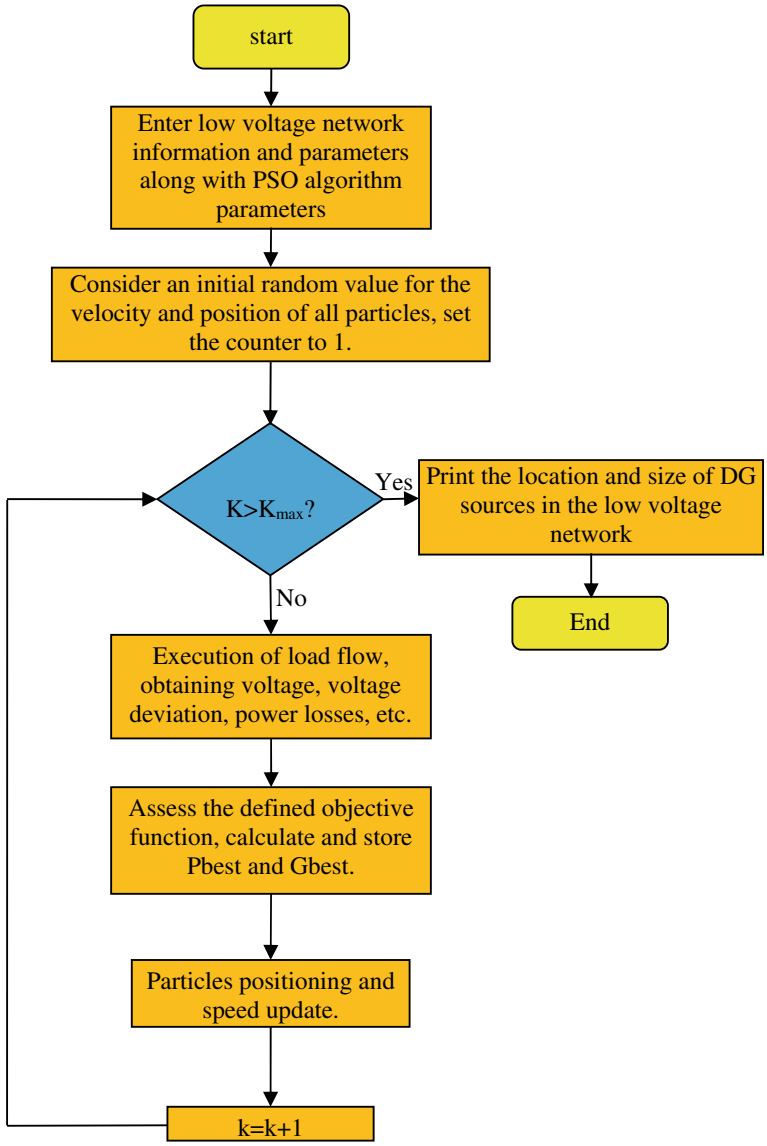


Fig. 15 PSO algorithm for placement and sizing DG

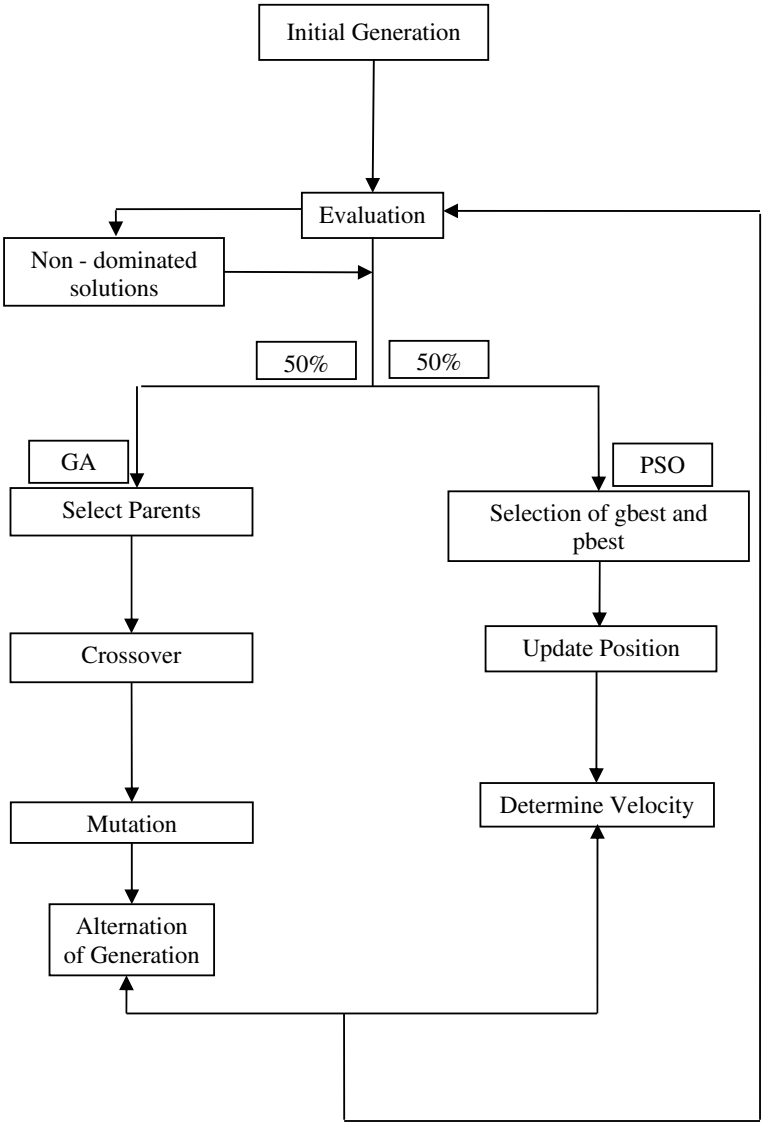


Fig. 16 Flowchart of HGAPSO algorithm

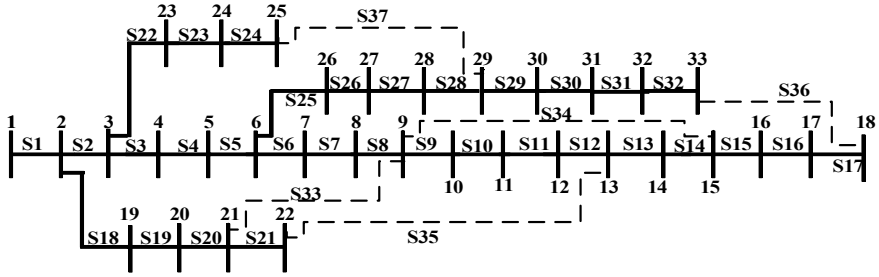


Fig. 17 33 bus distribution system

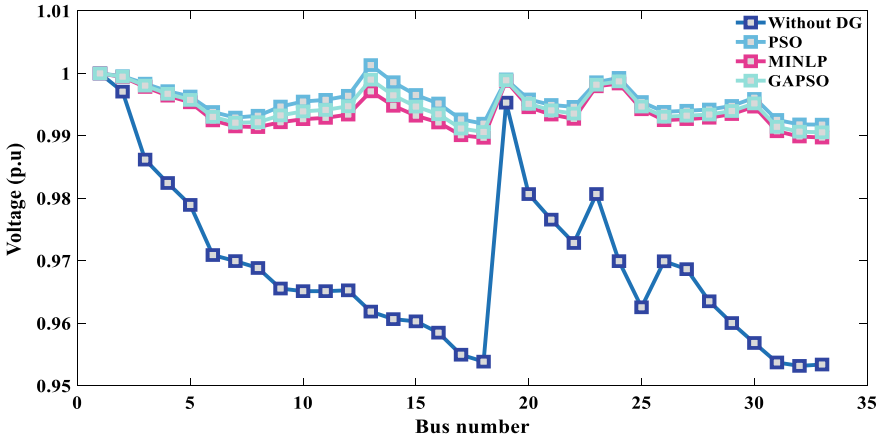


Fig. 18 Output voltage results for the studied methods

This has been done using the MINLP method of GAMS software and Knitro solver. The formulation for DG placement and sizing in the MINLP method is in accordance with the reference [32]. The PSO and GAPSO methods use the Newton Raphson load flow method to obtain power losses and voltage in each network arrangement and explains how this method works in Sections 3-4-5. The placement of all methods shows the same result, and according to the results, DGs should be in the 13th, 24th, and 30th buses. The power losses of active power and execution time in the MINLP method are less than the PSO and GAPSO. The reason for the lower execution time is the use of the Newton Raphson load flow method to solve power losses and voltage. In terms of voltage, the PSO method has a higher voltage profile level due to the selection of DGs with higher capacities. Because the allowable value set for this example is 0.95 to 1.05 p.u, voltages in the all methods are within the allowable range in terms of voltage. Although the MINLP method, like the PSO, is not guaranteed to achieve the optimal global response due to its nonlinear and non-convex problem, the MINLP method performs better performance due to lower power losses

and execution time. The PSO and GAPSO algorithm for high-dimensional networks is not used due to the long time required to solve the problem.

6 Conclusions

This chapter first explains the advantages and disadvantages of using DGs in a distribution network. DG placement methods are divided into two categories: the use of intelligent algorithms and numerical optimization methods, and each of these methods is described. Finally, to compare these methods, simulation for the DG placement problem is performed with one of the numerical optimization methods and two heuristic methods (PSO and GAPSO). The numerical optimization method works better in terms of execution time and reducing power losses, and also allocates smaller values for each DG. PSO and GAPSO methods also have higher voltage profile levels due to the allocation of larger values for each DG. Therefore, the use of numerical optimization methods in high dimensional networks is recommended.

References

1. Puttgen HB, Macgregor PR, Lambert FC (2003) Distributed generation: Semantic hype or the dawn of a new era? IEEE Power Energ Mag 1(1):22–29
2. Zhang X, Karady GG, Ariaratnam ST (2013) Optimal allocation of CHP-based distributed generation on urban energy distribution networks. IEEE Trans Sustain Energy 5(1):246–253
3. Kumar K, Kumar M (2020) Impacts of distributed generations on power system: transmission, distribution, power quality, and power stability. Handbook of research on new solutions and technologies in electrical distribution networks. IGI Global, pp 171–190
4. Freitas W et al. (2006) Comparative analysis between synchronous and induction machines for distributed generation applications. IEEE Trans Power Syst 21(1):301–311
5. Barker PP, De Mello RW (2000) Determining the impact of distributed generation on power systems. I. Radial distribution systems. 2000 Power engineering society summer meeting (Cat. No. 00CH37134), vol 3. IEEE
6. Khan UN (2008) Impact of distributed generation on electrical power network. Wroclav University of Technology, Wroclav, Poland 1
7. Photovoltaics, Dispersed Generation, and Energy Storage (2009) IEEE application guide for IEEE Std 1547™. IEEE standard for interconnecting distributed resources with electric power systems
8. Series IRE (2009) Microgrids and active distribution networks. The institution of Engineering and Technology
9. Zakariazadeh A, Jadid S (2014) Smart microgrid operational planning considering multiple demand response programs. J Renew Sustain Energy 6(1):013134
10. Zakariazadeh A, Jadid S, Siano P (2014) Smart microgrid energy and reserve scheduling with demand response using stochastic optimization. Int J Electr Power Energy Syst 63:523–533
11. Mizani S, Yazdani A (2009) Optimal design and operation of a grid-connected microgrid. In: 2009 IEEE electrical power & energy conference (EPEC). IEEE
12. Mohamed FA, Koivo HN (2010) System modelling and online optimal management of microgrid using mesh adaptive direct search. Int J Electr Power Energy Syst 32(5):398–407

13. Kayal P, Chanda CK (2015) Optimal mix of solar and wind distributed generations considering performance improvement of electrical distribution network. Renew energy 75:173–186
14. Hedayati H, Nabaviniaki A, Akbarimajd A (2008) A method for placement of DG units in distribution networks. IEEE Trans Power Delivery 23(3):1620–1628
15. Ochoa LF, Padilha-Feltrin A, Harrison GP (2006) Evaluating distributed generation impacts with a multiobjective index. IEEE Trans Power Delivery 21(3):1452–1458
16. Georgilakis PS, Hatziargyriou ND (2013) Optimal distributed generation placement in power distribution networks: models, methods, and future research. IEEE Trans Power Syst 28(3):3420–3428
17. Costa PM, Matos MA (2009) Avoided losses on LV networks as a result of microgeneration. Electr Power Syst Res 79(4):629–634
18. Gözel T, Hocaoglu MH (2009) An analytical method for the sizing and siting of distributed generators in radial systems. Electr Power Syst Res 79(6): 912–918
19. Lee SH, Park JW (2009) Selection of optimal location and size of multiple distributed generations by using Kalman filter algorithm. IEEE Trans Power Syst 24(3):1393–1400
20. Acharya N, Mahat P, Mithulananthan N (2006) An analytical approach for DG allocation in primary distribution network. Int J Electr Power Energy Syst 28(10):669–678
21. Hung DQ, Mithulananthan N (2011) Multiple distributed generator placement in primary distribution networks for loss reduction. IEEE Trans Indus Electron 60(4):1700–1708
22. Rau NS, Wan Y (1994) Optimum location of resources in distributed planning. IEEE Trans Power Syst 9(4):2014–2020
23. Vovos PN, Bialek JW (2005) Direct incorporation of fault level constraints in optimal power flow as a tool for network capacity analysis. IEEE Trans Power Syst 20(4):2125–2134
24. Rao SS (1996) Engineering optimization: theory and practice. John, New York, p 903
25. NEOS wiki (2007) Available at: http://www-fp.mcs.anl.gov/otc/Guide/
26. AlHajri MF, AlRashidi MR, El-Hawary ME (2010) Improved sequential quadratic programming approach for optimal distribution generation deployments via stability and sensitivity analyses. Electr Power Comp Syst 38(14):1595–1614
27. Ochoa LF, Dent CJ, Harrison GP (2009) Distribution network capacity assessment: variable DG and active networks. IEEE Trans Power Syst 25(1):87–95
28. Dent CJ, Ochoa LF, Harrison GP (2010) Network distributed generation capacity analysis using OPF with voltage step constraints. IEEE Trans Power Syst 25(1):296–304
29. Atwa YM et al (2009) Optimal renewable resources mix for distribution system energy loss minimization. IEEE Trans Power Syst 25(1):360–370
30. Kumar A, Gao W (2010) Optimal distributed generation location using mixed integer non-linear programming in hybrid electricity markets. IET Gener Transm Distrib 4(2):281–298
31. Esmaili M (2013) Placement of minimum distributed generation units observing power losses and voltage stability with network constraints. IET Gener Transm Dist 7(8):813–821
32. Kaur S, Kumbhar G, Sharma J (2014) A MINLP technique for optimal placement of multiple DG units in distribution systems. Int J Electr Power Energy Syst 63:609–617
33. Leyffer S (2001) Integrating SQP and branch-and-bound for mixed integer nonlinear programming. Comput Optim Appl 18(3):295–309
34. Khalesi N, Rezaei N, Haghifam M-R (2011) DG allocation with application of dynamic programming for loss reduction and reliability improvement. Int J Electr Power Energy Syst 33(2):288–295
35. Bellman RE, Dreyfus SE (1962) Applied dynamic programming. Princeton UniversityPress, Princeton, New Jersey, 1962. Bellman Applied Dynamic Programming
36. Jabr RA, Pal BC (2009) Ordinal optimisation approach for locating and sizing of distributed generation. IET Gener Transm Distrib 3(8):713–723
37. Singh D, Misra RK, Singh D (2007) Effect of load models in distributed generation planning. IEEE Trans Power Syst 22(4):2204–2212
38. Kotamarty S, Khushalani S, Schulz N (2008) Impact of distributed generation on distribution contingency analysis. Electr Power Syst Res 78(9):1537–1545

39. Khan H, Choudhry MA (2010) Implementation of Distributed Generation (IDG) algorithm for performance enhancement of distribution feeder under extreme load growth. Int J Electr Power Energy Syst 32(9):985–997
40. Moradi MH, Abedini M (2012) A combination of genetic algorithm and particle swarm optimization for optimal DG location and sizing in distribution systems. Int J Electr Power Energy Syst 34(1):66–74
41. Gomez-Gonzalez M, López A, Jurado F (2012) Optimization of distributed generation systems using a new discrete PSO and OPF. Electr Power Syst Res 84(1):174–180
42. Pandi VR, Zeineldin HH, Xiao W (2012) Determining optimal location and size of distributed generation resources considering harmonic and protection coordination limits. IEEE Trans Power Syst 28(2):1245–1254
43. Taşgetiren MF, Liang Y-C (2003) A binary particle swarm optimization algorithm for lot sizing problem. J Econ Soc Res 5(2):1–20
44. Singh RK, Goswami SK (2009) Optimum siting and sizing of distributed generations in radial and networked systems. Electr Power Comp Syst 37(2):127–145
45. Singh D, Singh D, Verma KS (2009) Multiobjective optimization for DG planning with load models. IEEE Trans Power Syst 24(1):427–436
46. Shukla TN et al (2010) Optimal sizing of distributed generation placed on radial distribution systems. Electr Power Comp Syst 38(3):260–274
47. Singh RK, Goswami SK (2010) Optimum allocation of distributed generations based on nodal pricing for profit, loss reduction, and voltage improvement including voltage rise issue. Int J Electr Power Energy Syst 32(6):637–644
48. Shaaban MF, Atwa YM, El-Saadany EF (2012) DG allocation for benefit maximization in distribution networks. IEEE Trans Power Syst 28(2):639–649
49. Golshan MEH, Arefifar SA (2007) Optimal allocation of distributed generation and reactive sources considering tap positions of voltage regulators as control variables. Eur Trans Electr Power 17(3):219–239
50. Novoa C, Jin T (2011) Reliability centered planning for distributed generation considering wind power volatility. Electr Power Syst Res 81(8):1654–1661
51. Wang L, Singh C (2008) Reliability-constrained optimum placement of reclosers and distributed generators in distribution networks using an ant colony system algorithm. IEEE Trans Syst Man Cybern C (Appl Rev) 38(6):757–764
52. Abu-Mouti FS, El-Hawary ME (2011) Optimal distributed generation allocation and sizing in distribution systems via artificial bee colony algorithm. IEEE Trans Power Delivery 26(4):2090–2101
53. Arya LD, Koshti A, Choube SC (2012) Distributed generation planning using differential evolution accounting voltage stability consideration. Int J Electr Power Energy Syst 42(1):196–207
54. Rao RS et al. (2012) Power loss minimization in distribution system using network reconfiguration in the presence of distributed generation. IEEE Trans Power Syst 28(1):317–325
55. Sharma D, Gaur P, Mittal AP (2014) Comparative analysis of hybrid GAPSO optimization technique with GA and PSO methods for cost optimization of an off-grid hybrid energy system. Energy Technol Policy 1(1):106–114
56. Choopani K, Hedayati M, Effatnejad R (2020) Self-healing optimization in active distribution network to improve reliability, and reduction losses, switching cost and load shedding. Int Trans Electr Energy Syst 30(5):e12348
57. Choopani K, Effatnejad R, Hedayati M (2020) Coordination of energy storage and wind power plant considering energy and reserve market for a resilience smart grid. J Energy Storage 30:101542
58. Vazinram F et al. (2020) Self-healing model for gas-electricity distribution network with consideration of various types of generation units and demand response capability. Energy Conver Manage 206:112487
59. Vazinram F et al. (2019) Decentralised self-healing model for gas and electricity distribution network. IET Gener Transm Distrib 13(19):4451–4463

Numerical Methods for Power System Analysis with FACTS Devices Applications

Mahdi Hedayati, Reza Effatnejad, Keyvan Choopani, and Milad Chanddel

Abstract Flexible AC Transmission Systems (FACTS) is created higher controllability in power systems. Several FACTS-devices are introduced for various applications. The basic applications of FACTS-devices are including power flow control, voltage control, reactive power compensation, stability improvement, power quality improvement, flicker mitigation, interconnection of renewable and distributed generation and storages. But to achieve the maximum profit from FACTS devices, the optimal size and location of these devices must be determined. Numerical optimization method is one of the methods to achieve the optimal global solution in optimization problems. This chapter of the book is about FACTS devices placement using numerical method. This method is performed in a radial Transmission system to minimize the power losses and improve the voltage profile. The proposed method is tested on the standard IEEE 14-bus test system. The results of case studies demonstrate the effectiveness of the proposed methodology.

Keywords SVC placement · Voltage stability factor · Numerical methods · Newton Raphson · FACTS · DAE · Voltage profile

The original version of this chapter was revised: The incorrect tagging of the organisation name have been updated. The correction to this chapter is available at https://doi.org/10.1007/978-3-030-62191-9_36

M. Hedayati (✉) · R. Effatnejad · K. Choopani · M. Chanddel
Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran
e-mail: m.hedayati@kiau.ac.ir

R. Effatnejad
e-mail: reza.efatnejad@kiau.ac.ir

K. Choopani
e-mail: chopani@jku.ac.ir

M. Chanddel
e-mail: m.chanddel@kiau.ac.ir

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_35

Abbreviations

V_i	The magnitude of voltage at bus i.
V_j	The magnitude of voltage at bus j.
V_m	The magnitude of voltage at bus m.
V_k	The magnitude of voltage at bus k.
V	The load voltage.
V_{ser}	Controllable voltage source in series branch.
V_{sh}	Controllable voltage source in parallel branch.
P_m	Mechanical power.
P_{ik}	Active power flow between the two buses i and k.
P_k	Active power flow between the two buses k and m.
P_m	Active power flow between the two buses m and k.
P_i	Active power of UPFC in node i.
P_j	Active power of UPFC in node j.
P_{sh}	The shunt converter active power.
P_{ser}	The series converter active power.
P_{Gi}	Active power generation.
P_{Di}	Active power consumption.
Q_{ik}	Reactive power flow between the two buses i and k.
Q_k	Reactive power flow between the two buses k and m.
Q_m	Reactive power flow between the two buses m and k.
Q_i	Reactive power of UPFC in node i.
Q_j	Reactive power of UPFC in node j.
Q_{sh}	The shunt converter reactive power.
Q_{ser}	The series converter reactive power.
Q_{Gi}	Reactive power generation.
Q_{Di}	Reactive power consumption.
X_{ik}	Reactor transmission line.
δ_{ik}	Power transfer angle.
C	The constant capacitor.
ω_m	The rotor speed.
δ	The phase voltage of the load voltage.
M	The inertia of the generator
x	The decision variable.
$[\Delta X]$	The solution vector.
[J]	The Jacobin matrix.
Ω	The solution range.
C_j (x)	The equal constraint.
H_k (x)	The unequal constraint.
F_v (x)	The objective function of the voltage deviation.
F_s (x)	The objective function of the system overload.
F_{PL} (x)	The objective function of active power losses.
Lmn	Line stability index.
VSF	Voltage stability factor.
α	Fire angle.

σ Conductivity angle.
k TCSC ratio.
B Susceptance.
Y Admittance.

1 Introduction

In the past decades, the increase in demand for electricity has led to an integrated power system with thermal and insulating limitations and issues such as the passage of power over unwanted routes and stability constraints. Experts and designers of power systems have proposed solutions to address these issues, such as the construction of a new line, Flexible Alternating Current Transmission Systems (FACTS) in transmission and distribution, and High Voltage Direct Current (HVDC). One of the disadvantages of building a new transmission line is environmental issues, and the disadvantages of HVDC are the high cost of this system. Therefore, to increase the capacity of the transmission line and improve the stability of the system, the appropriate solution is to use FACTS devices. The best solution is to use FACTS controllers to increase the capacity of distribution and transmission networks. FACTS controllers control important parameters of a transmission line (angle, voltage, current, active power, and reactive power). Applications of FACTS include improved transient stability, small signal stability improvement, voltage profile improvement, power losses reduction, and reduction of inrush currents caused by transformers. The important advantages of FACTS are shown in the Fig. 1 [1, 2].

The issue of FACTS devices is not new, as it has been built over the past 20 years using large-scale power technology and semiconductors. Today, the most important issue in FACTS controllers is the economic issue [1]. In addition to the applications mentioned above, these devices are also used to manage congestion, and reduce power losses. FACTS controllers are based on power electronics and are used in low voltage and transmission networks as series, parallel and series–parallel compensators to increase system controllability. For example, Static Synchronous Series

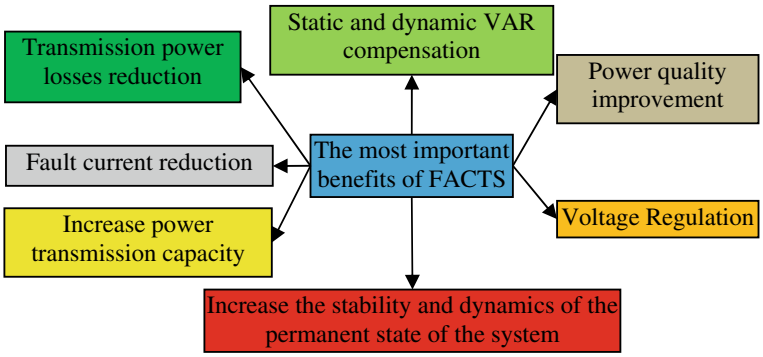


Fig. 1 Advantages of FACTS [1, 2]

Compensator (SSSC) and Thyristor Controlled Series Capacitor (TCSC) controllers are used as series compensators and Static Compensator (STATCOM) and Static Var Compensator (SVC) controllers are used as parallel compensators in the system. FACTS devices category will be presented later in the chapter [3].

At the beginning of the chapter, we will study the numerical techniques and solutions of MATLAB software, and in this section, we will explain the interval solution and open methods. We will then discuss the FACTS devices, the classification of FACTS controllers, the modeling, and the formulation of the optimal placement of FACTS devices. In the last part of this chapter, we will present a test system to analyze and simulate the optimal placement of SVC and UPFC controllers using Newton Raphson's numerical technique using MATLAB software.

2 Numerical Technique

Numerical analysis is widely used in the analysis of nonlinear systems in the fields of electrical and mechanical engineering, so that the MATLAB software has designed its solvers based on numerical analysis. Choosing the appropriate numerical technique to design and analyze power systems has become an important research challenge. Therefore, it is important to know numerical methods to solve nonlinear problems. In this chapter, we examine the advantages and disadvantages of widely used numerical techniques in power systems and their application in power systems. Applications of these techniques include load flow studies, maximum power point tracking (MPPT), optimal DG placement, and optimal FACTS devices placement. The first step in better understanding of these methods is to classify them. Therefore, we categorize numerical techniques into four subdivisions, taking into account their application in power systems:

1. Interval solution methods
2. Open methods
3. Introspection
4. Numerical Integration.

Table 1 Cardiovascular parameters [4]

Methods	Accuracy	Problem type	Solvers
Runge–Kutta	Intermediate	Non-hard	Ode45
Runge–Kutta	Low	Non-hard	Ode23
Adams	Low to high	Non-hard	Ode113
NDFs	Low to intermediate	Hard	Ode15s
Rosenbrock	Low	Hard	Ode23s
Trapezoidal rule	Low	Relatively hard	Ode23t
TR-BDF2	Low	Hard	Ode23tb

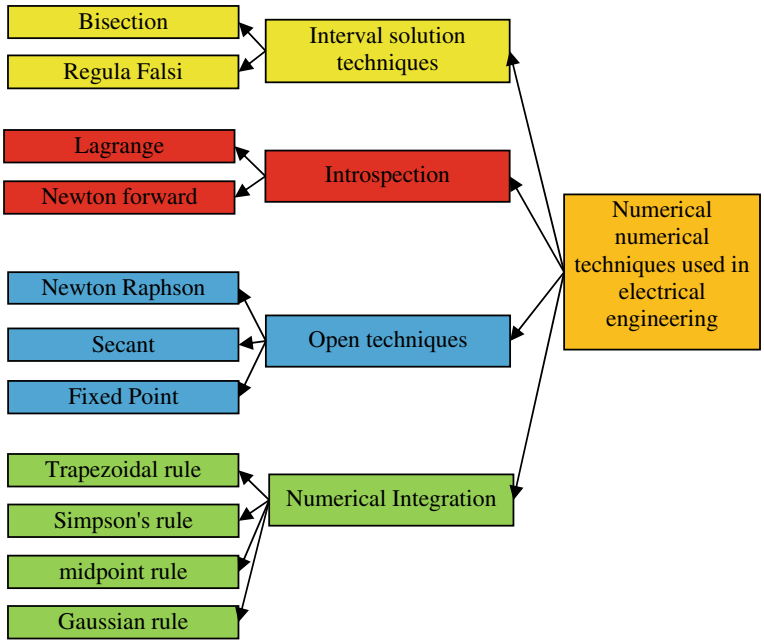


Fig. 2 Numerical techniques used in electrical engineering

In addition to the above classification, MATLAB software has also solvers based on numerical techniques to solve difficult and non-difficult problems so that the user can choose the appropriate solver according to the type of system. Table 1 shows the types of MATLAB solvers. Figure 2 shows the most widely used numerical techniques in the electrical industry. It is worth noting that in the studies of researchers, interval solution methods and open methods have received more attention. Therefore, the theoretical explanations of the methods of solving the interval and open methods will be discussed.

2.1 Bisection Method [5–9]

In this method, we assume that the numbers a and b are available in the following ways:

(a) The function f in interval [a, b] is continuous
(b) $f(a) \cdot f(b) < 0$
(c) $f(x) = 0$ has only one root (α) in the range (a, b).

In this method, we make the sequence $\{x_n\}$ as follows:

$$\lim_{n\to\infty} X_n = \alpha \tag{1}$$

To build this sequence, we do the following:

We first X_1 consider (a + b)/2, which divides [a, b] into two parts [a, X_1] and [X_1, b]. Now, assuming case C, α is necessarily in one of these two parts. To continue, we consider the part where α is located, where it is possible to select the desired range through the Bolzano theorem. After selecting the desired range, we take the X_2 method in the same way and continue the same process until we reach the condition of stopping the problem. Note that in the bisection method, the sequence is constructed to converge to the root of the function, and it is easy to show that:

$$|x_n - \alpha| \leq \frac{b-a}{2^n} \tag{2}$$

Disadvantages include the following:

1. This numerical technique is not suitable for finding double roots.
2. Slow dynamically answer.
3. Very slow performance when finding multiple roots.

Application of this numerical method in smart networks include operation studies in low voltage networks.

In the reference [10], a distributed bisection technique with the presence of the small-scale resources has proposed for the economic dispatch of load on the smart networks.

Another application of this technique in power systems is the discussion of the MPPT in renewable sources such as wind turbines and photovoltaic systems.

In other words, this method can be used to increase the extraction capacity of the mentioned resources [11]. In addition to electrical engineering, this technique has many applications in other branches of engineering such as mechanics. The bisection algorithm is shown in Fig. 3.

2.2 Regula Falsi Method [5–9]

Assume that the function f in [a, b] is continuous and that f(a)f(b) < 0 and the equation f(x) = 0 have only one root in (a, b). Now, to find an approximation of this root, which we call α, we action as follows. Connect the two points (a, f(a)) and (b, f(b)) with a straight line, we call the intersection of this line with the axis of the x¬ x_1 which approximate of the desired root is α, now to find the value of x_1, we place point (x_1, 0) in the equation of the line connecting the two points (a, f(a)) and (b, f(b)), that is, in Eq. 3:

$$\frac{y - f(a)}{x - a} = \frac{f(b) - f(a)}{b - a} \tag{3}$$

We place the point (x_1, 0), so we have:

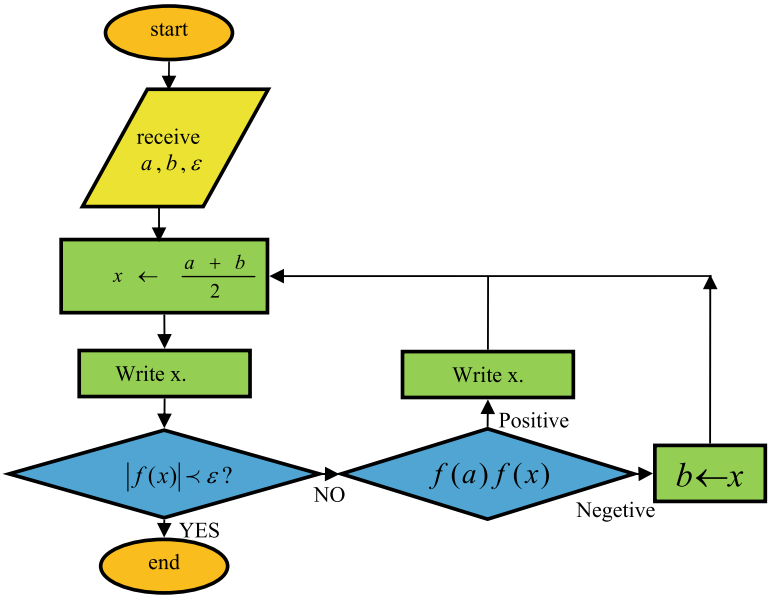


Fig. 3 Bisection algorithm [8]

$$\frac{0 - f(a)}{x_1 - a} = \frac{f(b) - f(a)}{b - a} \tag{4}$$

After simplifying, we get to the following equation:

$$x_1 = \frac{af(b) - bf(a)}{f(b) - f(a)} \tag{5}$$

Now to determine x_2, do the following:

1- If $f(a)f(x_1) < 0$, then according to Bolzano's theorem the root α is in (a, x_1), so in formula 5 instead of b, we put x_1 and we have:

$$x_2 = \frac{af(x_1) - x_1 f(a)}{f(x_1) - f(a)} \tag{6}$$

2- If $f(a)f(x_1) > 0$, then the root is in (x_1, b). So in formula 5, instead of a, we put x_1 and we have:

$$x_2 = \frac{x_1 f(b) - bf(x_1)}{f(b) - f(x_1)} \tag{7}$$

3- If $f(a)f(x_1) = 0$ then it is clear that x_1 is the desired root and the problem is solved.

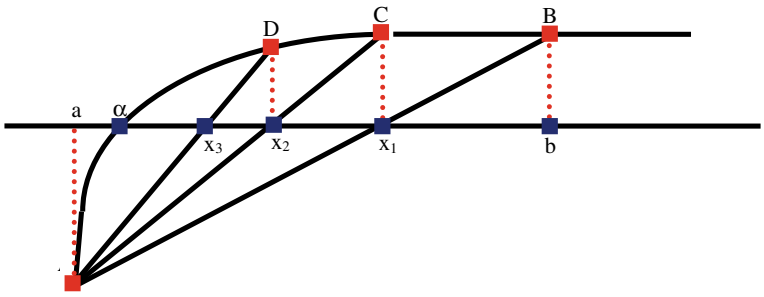


Fig. 4 Details of Regula Falsi method [8]

Thus, a sequence of numbers is obtained that converges to the root α. As you can see, the Regula Falsi method is convergent similar to the Bisection method. Sometimes it works even faster than the Bisection method. Of course, the operation of this method is more than the two-part method. See Fig. 4 for more details.

2.3 Simple Repetition or Fixed Point Method [5–9]

In this method, by applying changes in the equation f(x) = 0 to x = g(x), if α is the root of f(x), then g(α) = α. In fact, the root of f will be a fixed point g.

Fixed point theorem:

Suppose that $g : [a, b] \to [a, b]$ function is a continuous function ($g \in c[a, b]$) in this case:

(A) g has at least one fixed point in [a, b]. In addition, if g' is present on the interval (a, b) and the number $0 < L < 1$ is present so that we have $|g'(x)| \leq L < 1$ for each x in (a, b), then the constant point g in [a, b] is unique.

(B) The sequence produced by $x_{n+1} = g(x_n)$ that $n = 0, 1, 2, 3, \ldots$ is convergent for each $x_0 \in [a, b]$ to the constant point g (root f). To better understand the fixed point technique, consider Fig. 5.

2.4 Newton Raphson Method [5–9]

This numerical technique works faster than other methods mentioned above. However, in this method, the starting point of x_0 must be close enough to the desired root α, which is one of the disadvantages of this method. After explaining how this method works, you will clearly see that there is a special case of the simple repetition method mentioned. Now, with the help of Fig. 6, we will describe this method.

If x_0 is approximate to the desired root, α. As you can see in the figure, we draw the tangent line of the curve f(x) at point A to the coordinates (x_0, f(x_0)). We call the

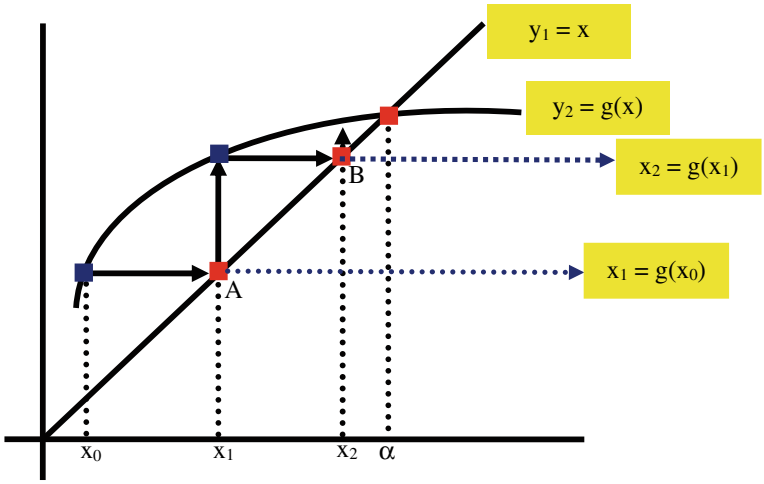


Fig. 5 Details of fixed point method [8]

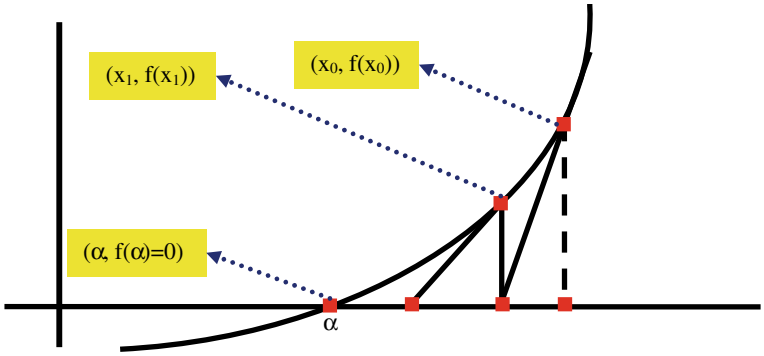


Fig. 6 Newton Raphson method [8]

location of this line the axis of length x_1. In fact, by placing a point $(x_1, 0)$ in Eq. 8 and continuing this process, we achieve a repetitive design of Newton method (9).

$$y - f(x_0) = f'(x_0)(x - x_0) \tag{8}$$

So we will have:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$\ldots$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{9}$$

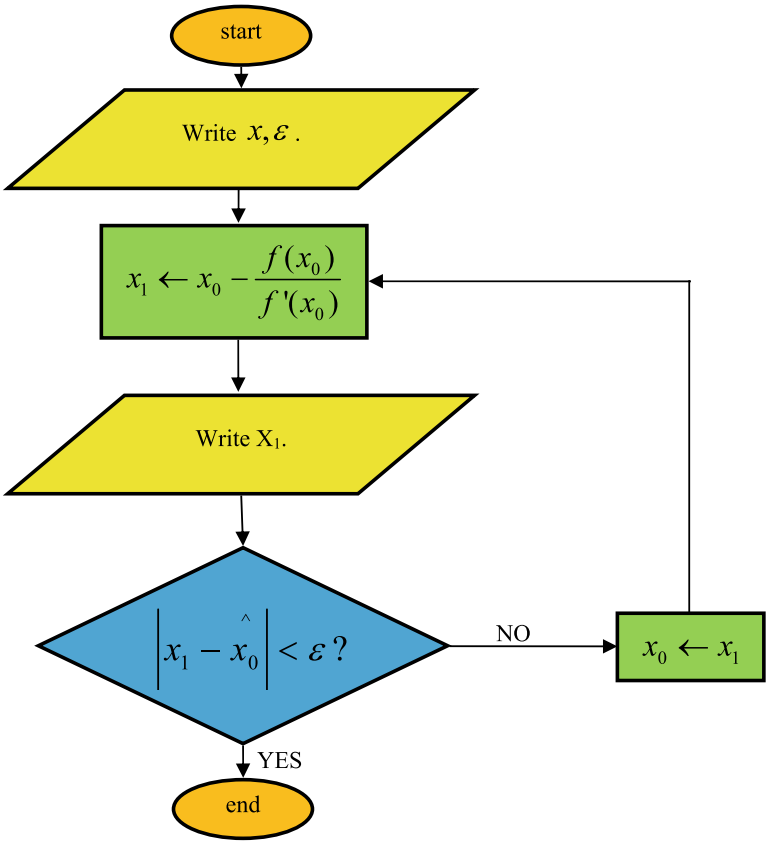


Fig. 7 Newton Raphson algorithm [8]

Given the Eq. 10, it is clear that Newton's method is a special case of the simple repetition method.

$$g(x) = x - \frac{f(x)}{f'(x)} \tag{10}$$

Newton Raphson method is shown in Fig. 7.

2.5 Secant Method [5–9]

First, consider the Eq. 11, which is the derivative definition.

$$\lim_{x \to x_n} \frac{f(x_n) - f(x)}{x_n - x} = f'(x_n) \tag{11}$$

If the value of x is close to the value of x_n, for example, x_{n-1}, then:

$$\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} = f'(x_n) \tag{12}$$

In the Newton formula, we use Eq. 12 instead of f '(x_n), and after simplification, we will have:

$$x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})} \tag{13}$$

Note that we need the values initial two of x_0 and x_1 in the Secant method for calculating x_n. Then the rest of the sequence of sentences is easily obtained through Eq. 13. The Secant method is not necessarily convergent and is slower than Newton in terms of speed, but much faster than the Bisection method and the Regula Falsi technique. Note Fig. 8 for further explanation.

One of the most important issues in smart grids is the discussion of MPPT from renewable sources such as photovoltaic systems and wind turbine systems. Because the mentioned sources are facing uncertainties such as the amount of sunlight and wind speed. In recent years, many studies and researches have been done in the field of MPPT in low voltage networks.

The researcher in the reference [11] used numerical techniques to investigate the tracking of the MPPT in photovoltaic systems. Figure 9 shows the MPPT controller for a photovoltaic system and Fig. 10 shows the performance of numerical techniques.

Numerical optimization methods can ensure the global optimal solution. Also, used methods in [12–15] are to achieve global optimal.

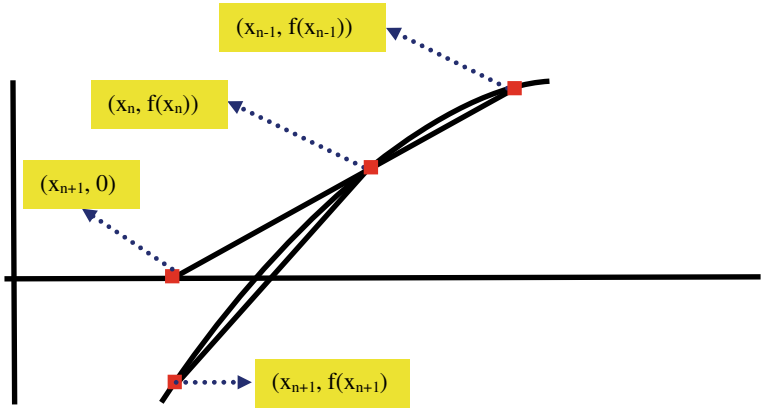


Fig. 8 Secant method [8]

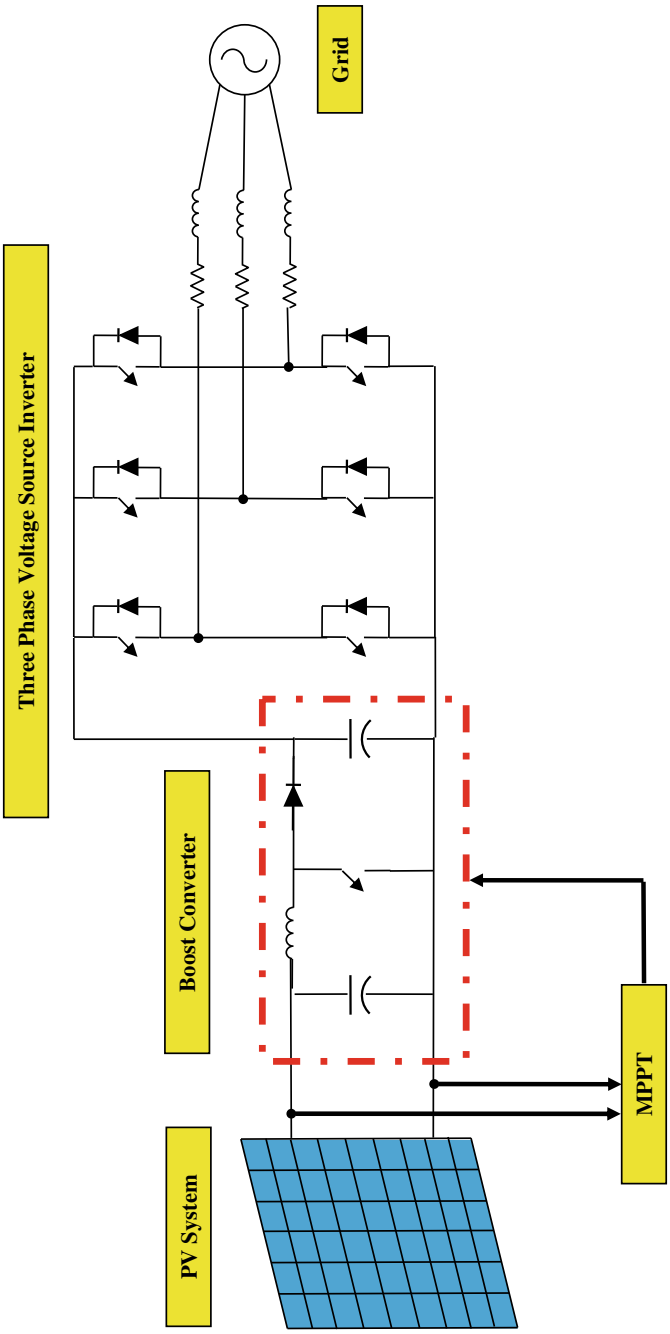


Fig. 9 MPPT controller for a photovoltaic system

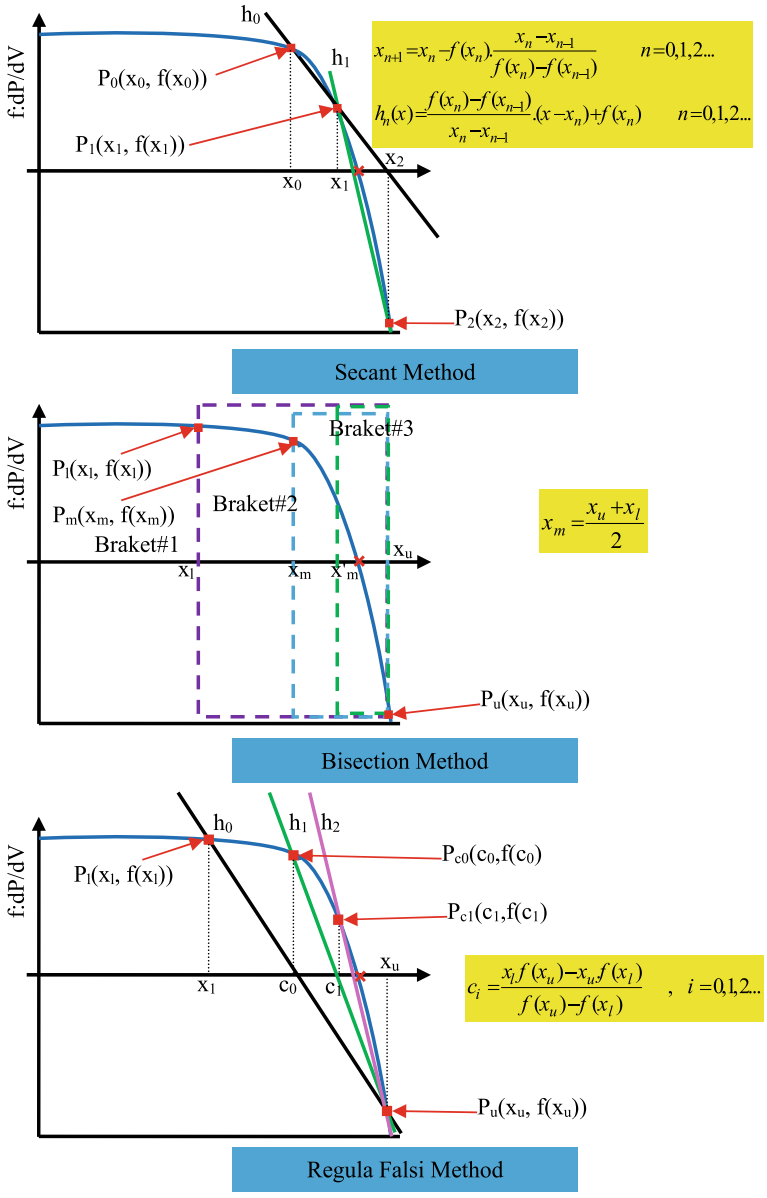


Fig. 10 Performance of numerical techniques [11]

3 FACTS Devices

The flow of active and reactive power (the flow of current leads to the flow of power.) In a transmission line between the two buses i and k shown in Fig. 11 of the Eqs. 14 and 15 are calculated [16]:

$$P_{ik} = -P_{ki} = \frac{V_i V_k}{x_{ik}} \sin(\delta_i - \delta_k) \tag{14}$$

$$Q_{ik} = \frac{1}{x_{ik}}[V_i^2 - V_i V_k \cos(\delta_i - \delta_k)] \tag{15}$$

V_i and V_k The magnitudes of the bus voltages i and k.
X_{ik} Reactor Transmission Line.
δ_{ik} power transfer angle.

FACTS controllers are divided into three classes based on the type of compensation, series (injecting voltage into the system in series), parallel (their connection is parallel and at the connection point, the current is injected into the system) and series–parallel are classified. The effect of some FACTS controllers on Eq. 14 is shown in Fig. 12 [16].

The classification of FACTS device controllers, the structure of the famous FACTS machines, and the use of the famous FACTS devices are shown in Figs. 13, 14 and 15, respectively. So, SSSC, Interline Power Flow Controller (IPFC), TCSC, Thyristor Switched Series Capacitor (TSSC), Thyristor Controlled Series Reactor

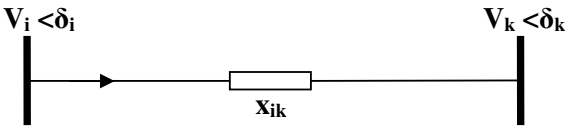


Fig. 11 Active and reactive power flow

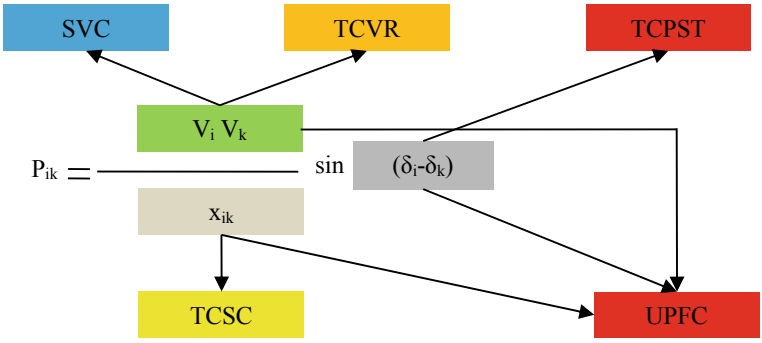


Fig. 12 Effect of FACTS controllers [16]

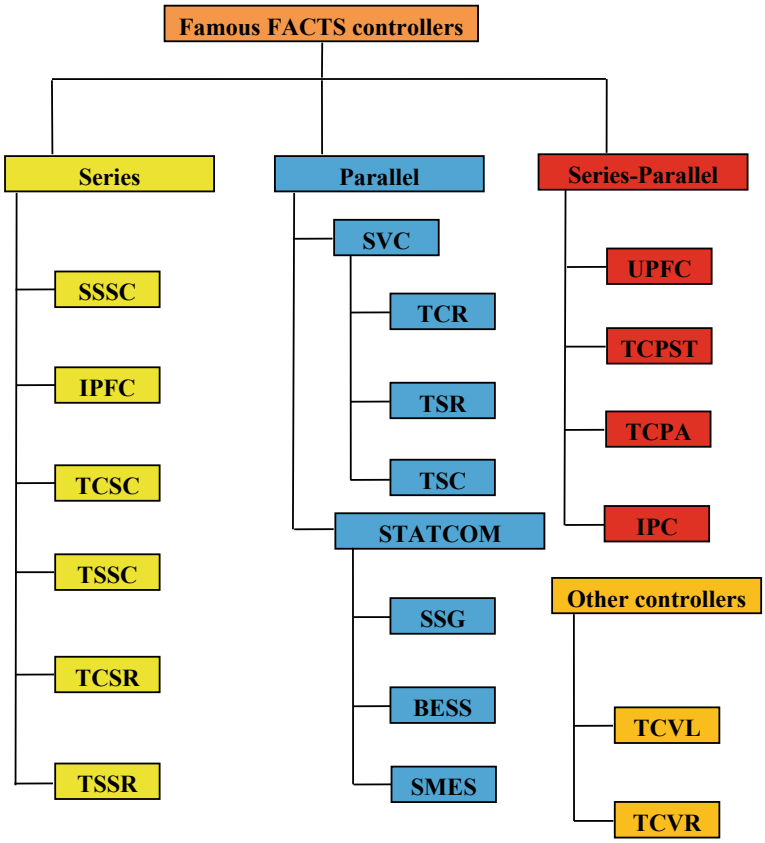


Fig. 13 FACTS controllers' classification [2]

(TCSR), Thyristor Switched Series Reactor (TSSR), SVC, Thyristor Controlled Reactor (TCR), Thyristor Switched Reactor (TSR), Thyristor Switched Capacitor (TSC), STATCOM, Static Synchronous Generator (SSG), Battery Energy Storage System (BESS), Superconducting Magnetic Energy Storage(SMES), UPFC, TCPST, Thyristor Controlled Phase Angle Regulator (TCPAR), Interphase Power Controller (IPC), Thyristor Controlled Voltage Limiter (TCVL) and TCVR are specified in the figures.

3.1 Power System Modeling

The first step in the study of power systems, such as the optimal placement of FACTS controllers, is to provide the correct model of the system under study. In other words,

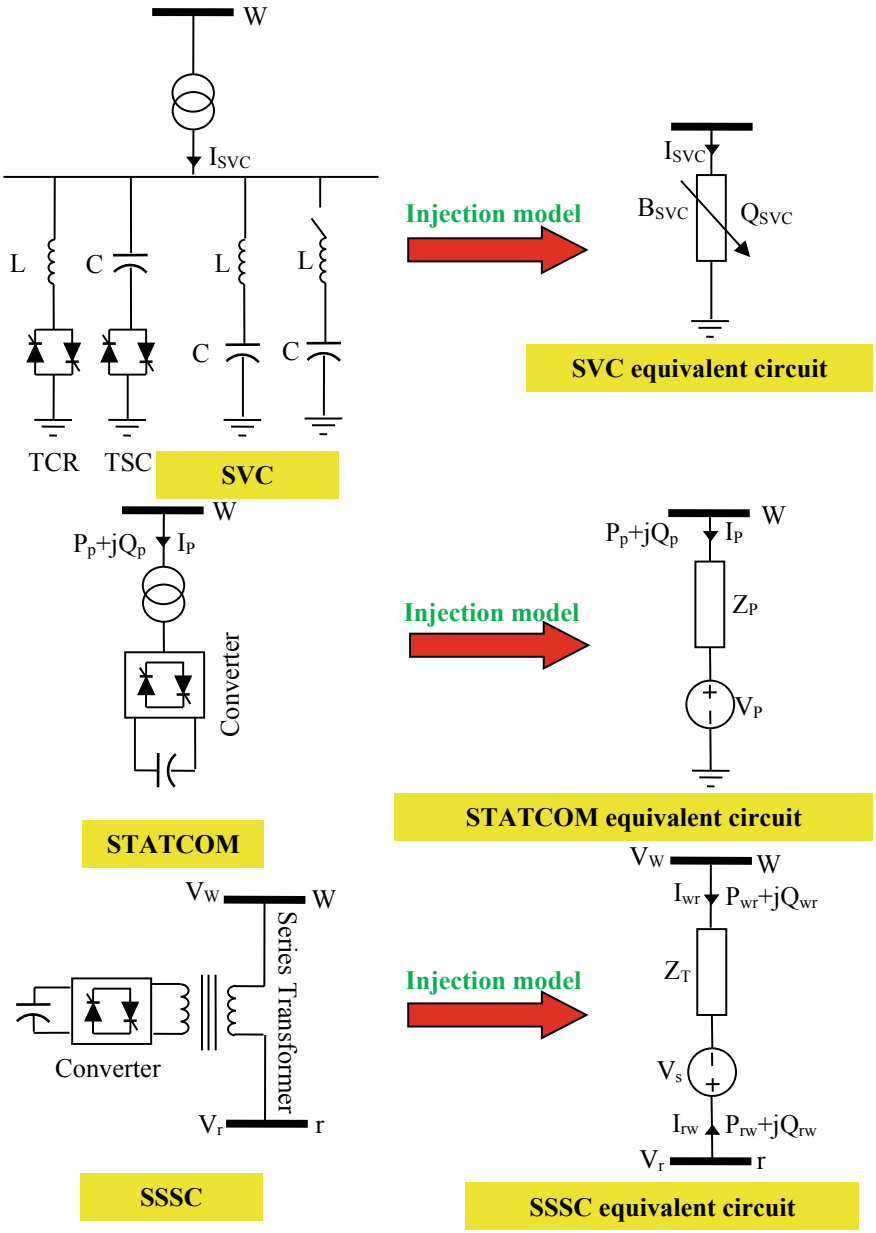
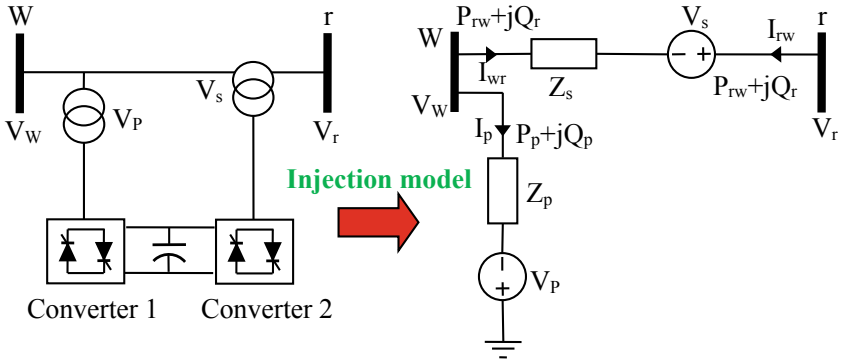
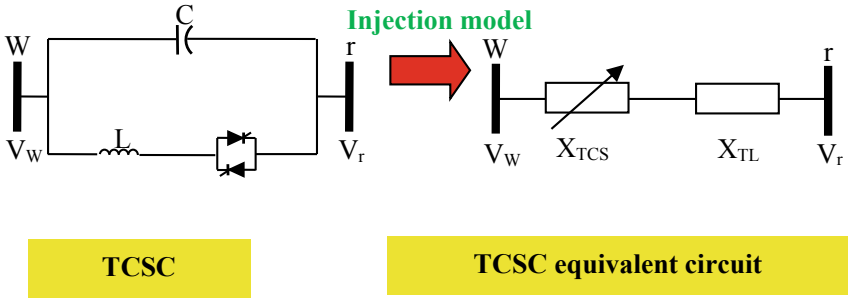


Fig. 14 The structure of the famous FACTS devices [17]





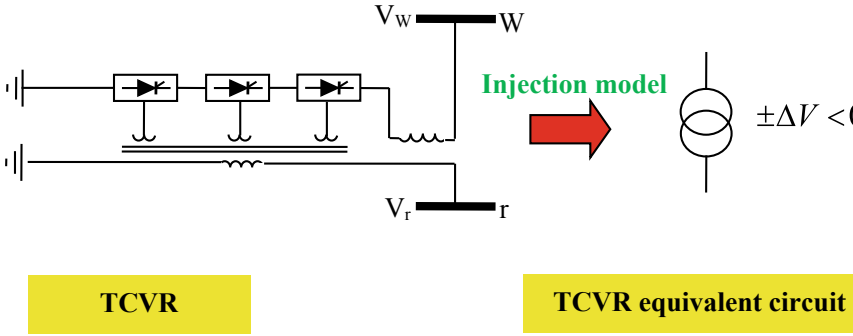


Fig. 14 (continued)

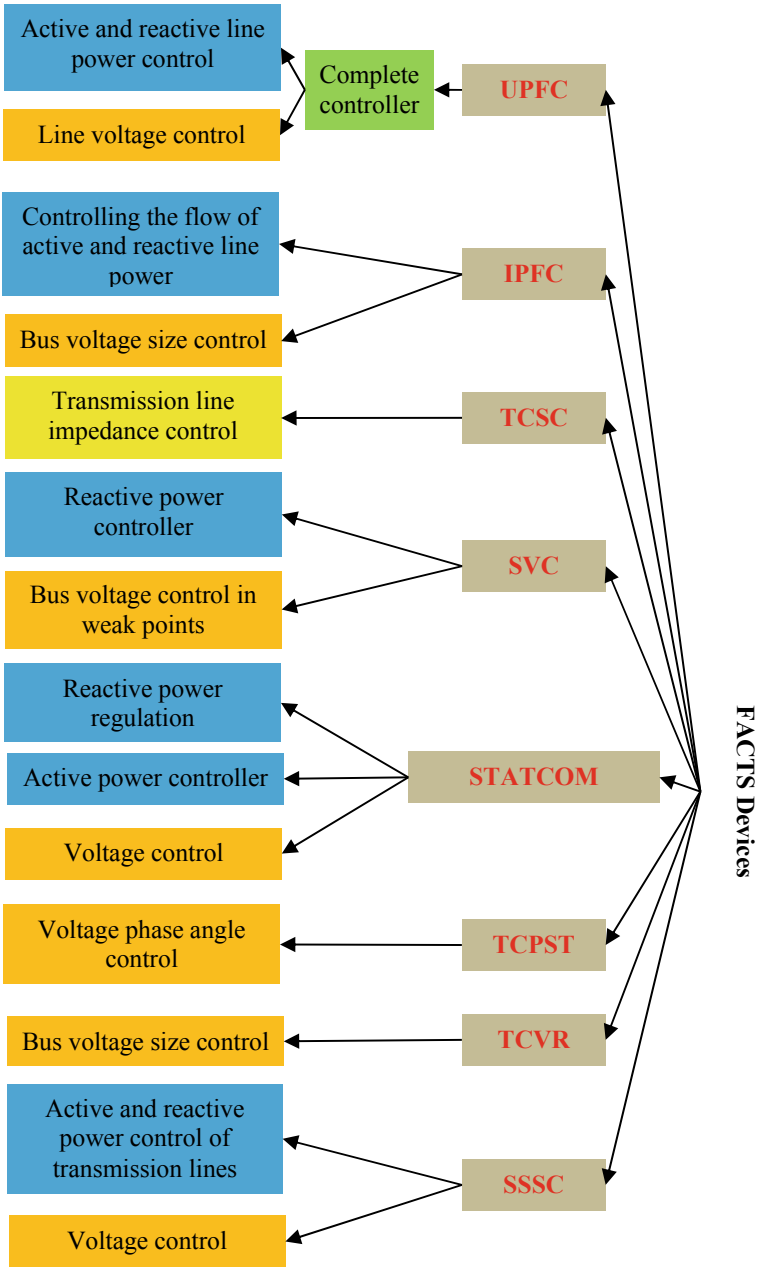


Fig. 15 Application of well-known FACTS devices [17]

if the model is not presented correctly, the objectives of the studies will not be achieved.

The mathematical model of the power system shown in Fig. 16 is defined in Eqs. 16–23 [18, 19].

$$\dot{\delta}_m = \omega_m$$
$$M\dot{\omega}_m = P_m - d_m\omega_m + E_m^2 Y_m \operatorname{Sin}\theta_m + E_m Y_m V \operatorname{Sin}(\delta - \delta_m - \theta_m) \tag{16}$$

$$k_{qw}\dot{\delta} = -k_{qV2}V^2 - k_{qV} + Q(\delta_m, \delta, V) - Q_0 - Q_1 \tag{17}$$

$$\begin{aligned} Tk_{q\omega}k_{pv}\dot{V} = {} & k_{p\omega}k_{qv2}V^2 + (k_{p\omega}k_{qv} - k_{q\omega}k_{pv})V \\ & + k_{qw}(P(\delta_m, \delta, V) - P_0 - P_1) \\ & - k_{p\omega}((Q(\delta_m, \delta, V) - Q_0 - Q_1) \end{aligned} \tag{18}$$

$$\begin{aligned} P(\delta_m, \delta, V) = {} & (Y_0' \operatorname{Sin}\theta_0' + Y_m \operatorname{Sin}\theta_m)V^2 \\ & - E_m Y_m V \operatorname{Sin}(\delta - \delta_m + \theta_m) \\ & - E_0' Y_0' V \operatorname{Sin}(\delta + \theta_0') \end{aligned} \tag{19}$$

$$\begin{aligned} Q(\delta_m, \delta, V) = {} & -(Y_0' \operatorname{Cos}\theta_0' + Y_m \operatorname{Cos}\theta_m)V^2 \\ & + E_m Y_m V \operatorname{Cos}(\delta - \delta_m + \theta_m) \\ & + E_0' Y_0' V \operatorname{Cos}(\delta + \theta_0') \end{aligned} \tag{20}$$

$$E_0' = [1 + C^2 Y_0^{-2} - 2CY_0^{-1} \operatorname{Cos}\theta_0]^{-\frac{1}{2}} \tag{21}$$

$$Y_0' = Y_0[1 + C^2 Y_0^{-2} - 2CY_0^{-1} \operatorname{Cos}\theta_0]^{\frac{1}{2}} \tag{22}$$

$$\theta_0' = \theta_0 + \tan^{-1}\left\{\frac{CY_0^{-1} \operatorname{Sin}\theta_0}{1 - CY_0^{-1} \operatorname{Cos}\theta_0}\right\} \tag{23}$$

The mathematical model of the load is shown in Eq. 24 [18]:

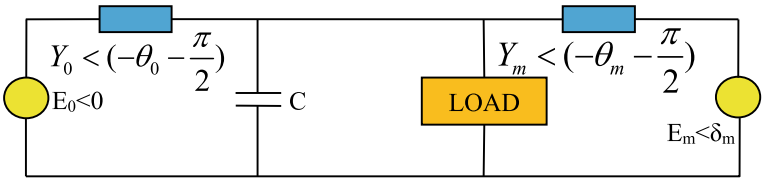


Fig. 16 Power system model [18]

$$\begin{cases} P = P_0 + P_1 + k_{p\omega}\dot{\delta} + k_{pv}V + k_{qv2}V^2 \\ Q = Q_0 + Q_1 + k_{q\omega}\dot{\delta} + k_{qv}V + k_{qv2}V^2 \end{cases} \quad (24)$$

C represents the constant capacitor, V represents the load voltage, δ represents the phase voltage of the load voltage, ω_m the rotor speed, δ_m the phase angle of the generator, M the inertia of the generator, d_m dumping and P_m express the mechanical power [18, 19].

3.2 SVC Modeling

In this device, changing of load and topology of the system will affect the bus voltage. Therefore, in case of improper control, the voltage will drop and may even lead to network collapse. Conventional devices such as capacitors, reactors (conventional devices controlled by mechanical switching), and rotating synchronous condenser can be used to adjust the voltage of the buses that are out of range. The injection or absorption of reactive power is automatically adjusted by the SVC and the rotating synchronous condenser to keep the voltage across the buses constant. SVC is a combination of capacitive and induction banks with thyristor control or mechanical switching control as shown in Fig. 7. To conduct numerical studies, SVC is classified into two models [20]:

1. shunt variable susceptance model
   susceptance is considered as a mode variable in a state-space equations and formulations. The shunt variable susceptance model is shown in Fig. 17 [21, 22].

2. Fire angle model
   Fire angle is considered as a mode variable in Eqs. and formulations. The fire angle model is shown in Fig. 18 [21, 22].

Figure 19 shows the dynamic characteristics and the steady-state voltage and current of the SVC controller. In the SVC controller, we only exchange reactive power. Therefore, the susceptance, current, and reactive power are changed to adjust the voltage according to a desirable characteristic with a suitable slope (slope value is usually considered between 1 and 5%) [21].

The amount of slope depends on the desired voltage (the optimum voltage is defined in the controllable area), the production of reactive power between different

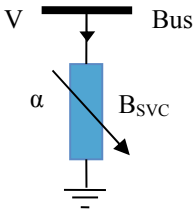


Fig. 17 Variable susceptance model [21, 22]

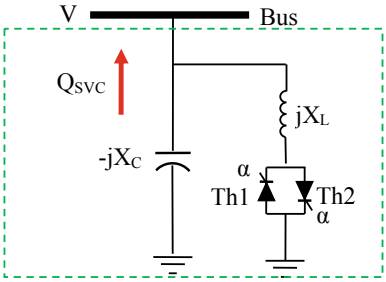


Fig. 18 Fire angle model [21, 22]

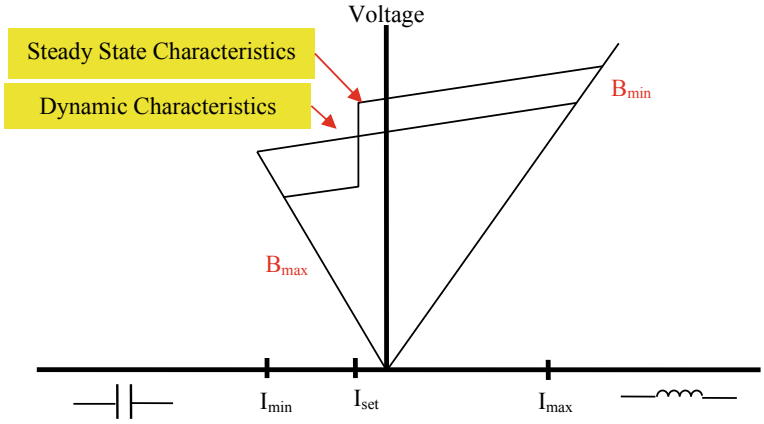


Fig. 19 Dynamic characteristics, steady state mode voltage, and SVC current controller [21]

sources (considering optimal conditions for the system), and attention to other needs of the system. Figure 19 shows that the SVC compensates for the reactive power in the system in a capacitive region such as a shunt capacitor and in an inductive section such as a shunt reactor. In the following, we provide a numerical analysis of the fire angle model and the variable susceptance model of the SVC controller (considering sine voltage). The I_{SVC} is defined according to Eq. 25 [21].

$$I_{SVC} = -jB_{SVC}V_k \tag{25}$$

X_{TCR} is defined according to Eq. 26 [21, 22].

$$X_{TCR} = \frac{\pi X_L}{\sigma - \sin\sigma} \tag{26}$$

We know that $X_L = \omega L$, now we have a Eq. by placing $\sigma = 2(\pi - \alpha)$ [21, 22].

$$X_{TCR} = \frac{\pi X_L}{2(\pi - \alpha) + \sin(2\alpha)} \tag{27}$$

In the Equations mentioned, σ indicates conductivity, and α indicates fire angles. If α = 90, the TCR will be fully conductive. In other words, X_{TCR} will be equal to X_L. If α = 180, the TCR conductivity will be zero, and the reactance will be infinite. SVC reactance is a combination of a capacitor and TCR reactance listed in Eq. 28 [21, 22].

$$X_{SVC}(\alpha) = \frac{\pi X_C X_L}{X_C[2(\pi - \alpha) + \sin 2\alpha] - \pi X_L} \tag{28}$$

We know that $X_C = \frac{1}{\omega C}$, then we define Q_K according to Eq. 29 [21, 22].

$$Q_k = -V_k^2 \left\{ \frac{X_C[2(\pi - \alpha) + \sin 2\alpha}{\pi X_C X_L} \right\} \tag{29}$$

On the other hand, the SVC equivalent susceptance is defined according to Eq. 30 [21, 22]:

$$B_{SVC} = -\frac{1}{X_{SVC}} \tag{30}$$

3.2.1 Power Flow Model

The equations for the linearized power flow of the SVC controller are shown in Table 2 based on the shunts variable susceptance models and fire angle [20, 23].

Table 2 The shunt variable susceptance models and fire angle [20, 23]

Model	Description
*S*hunt variable susceptance $\begin{bmatrix} \Delta P_k \\ \Delta Q_k \end{bmatrix}^i = \begin{bmatrix} 0 & 0 \\ 0 & Q_k \end{bmatrix}^i \begin{bmatrix} \Delta\theta_k \\ \Delta B_{SVC}/B_{SVC} \end{bmatrix}^i$	$B_{SVC}^{i+1} = B_{SVC}^i + \left(\frac{\Delta B_{SVC}}{B_{SVC}}\right)^i B_{SVC}^i$
Fire angle $\begin{bmatrix} \Delta P_k \\ \Delta Q_k \end{bmatrix}^i = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\partial Q_k}{\partial \alpha} \end{bmatrix}^i \begin{bmatrix} \Delta\theta_k \\ \alpha \end{bmatrix}^i$	$\frac{\partial Q_k}{\partial \alpha} = \frac{2V_k^2}{X_L}(\cos(2\alpha) - 1)$ $\alpha^{i+1} = \alpha^i + \Delta\alpha^i$ $B_{eq} = -\frac{X_L - \frac{X_C}{\pi}(2(\pi - \alpha) + \sin(2\alpha))}{X_C X_L}$

3.2.2 SVC Controller Model

The SVC diagram block with a classic PI controller is shown in Fig. 20. X_{SVC} is modified using fire angle control (this change is done with the help of a classical controller) to adjust the voltage within the allowable range [21, 22].

According to the figure above, the state equations $\dot{X}_{1SVC}$, $\dot{X}_{2SVC}$, and $\dot{X}_{3SVC}$ according to Eq. 31 are defined [21, 22].

$$\begin{aligned}
\dot{X}_{1SVC} &= \frac{1}{T_m}[V_{SVC}(1 + KX_{3SVC}) - X_{1SVC}] \\
\dot{X}_{2SVC} &= K_I\left[V_{ref,SVC} - X_{1SVC}\right] \\
\dot{X}_{3SVC} &= \frac{1}{T_c}\left[X_{2SVC} + K_P(V_{ref,SVC} - X_{1SVC}) - X_{3SVC}\right]
\end{aligned} \tag{31}$$

Q_{SVC} reactive power is expressed according to Eq. 32 [21]:

$$Q_{SVC} = V_{SVC}^2 X_{3SVC} \tag{32}$$

In the following, we will perform the Eqs. 31 and 32 [21, 22]:

$$\begin{aligned}
\Delta Q_{SVC} &= 2V_{SVC0}\Delta V_{SVC}X_{3SVC0} + V_{SVC0}^2\Delta X_{3SVC} \\
\Delta\dot{X}_{1SVC} &= \frac{1}{T_m}[\Delta V_{SVC}(1 + KX_{3SVC0}) + V_{SVC0}K\Delta X_{3SVC} - \Delta X_{1SVC}] \\
\Delta\dot{X}_{2SVC} &= K_I(\Delta V_{ref,SVC} - \Delta X_{1SVC}) \\
\Delta\dot{X}_{3SVC} &= \frac{1}{T_C}\left[\Delta X_{2SVC} + K_P(\Delta V_{ref,SVC} - \Delta X_{1SVC}) - \Delta X_{3SVC}\right]
\end{aligned} \tag{33}$$

Finally, we will have the Eq. 34 [21, 22]:

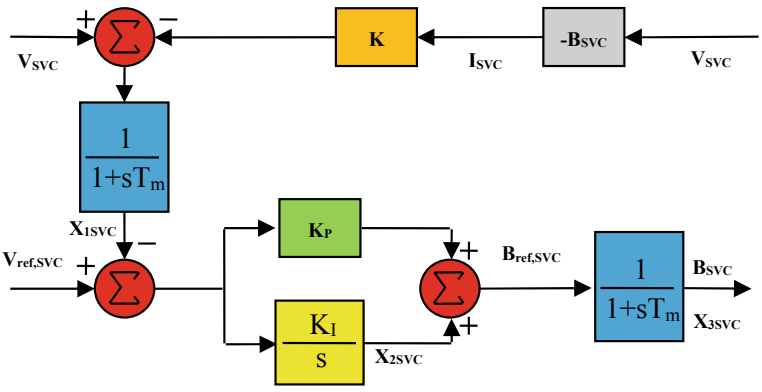


Fig. 20 Dynamic characteristics and constant voltage mode—SVC current controller [21, 22]

$$\begin{bmatrix} \Delta\dot{X}_{1SVC} \\ \Delta\dot{X}_{2SVC} \\ \Delta\dot{X}_{3SVC} \end{bmatrix} = \begin{bmatrix} \frac{-1}{T_m} & 0 & \frac{KV_{SVC0}}{T_m} \\ -K_I & 0 & 0 \\ \frac{-K_P}{T_C} & \frac{1}{T_C} & \frac{-1}{T_C} \end{bmatrix} \begin{bmatrix} \Delta X_{1SVC} \\ \Delta X_{2SVC} \\ \Delta X_{3SVC} \end{bmatrix} + \begin{bmatrix} \frac{1}{T_m}(1 + KX_{3SVC0}) \\ 0 \\ 0 \end{bmatrix} [\Delta V_{SVC}] \tag{34}$$

Equation 34 can be summarized in the form of Eq. 35 [21, 22]:

$$\Delta\dot{X}_{SVC} = A_{SVC}\Delta X_{SVC} + B_{SVC}\Delta V_{SVC} \tag{35}$$

3.2.3 SVC Connection in Multi-machine Power Systems

Equilibrium power equations for the optimal placement of the SVC controller in multi-machine power systems are defined according to Eq. 36 [21, 22].

$$\begin{aligned} &P_{SVCi} + P_{Li}(V_i) - \sum_{k=1}^{n} V_i V_K Y_{ik} \cos(\theta_i - \theta_k - \alpha_{ik}) = 0 \\ &i = m+1, ..., n \\ &Q_{SVCi} + Q_{Li}(V_i) - \sum_{k=1}^{n} V_i V_k Y_{ik} \sin(\theta_i - \theta_k - \alpha_{ik}) = 0 \\ &i = m+1, ..., n \end{aligned} \tag{36}$$

After the linearization of the above equations, the Eq. 37 is obtained [21, 22].

$$\begin{aligned} &D_{SVC}\Delta X_{SVC} + D_1\Delta V_g + D_{2SVC}\Delta V_l = 0 \\ &D_{2SVC} = C_{SVC} + D_2 \end{aligned} \tag{37}$$

Considering the SVC controller, the DAE model is defined according to Eq. 38 [21, 22]:

$$\begin{bmatrix} \Delta\dot{X} \\ \Delta\dot{X}_{SVC} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} A_{1\,\mathrm{mod}} & P_{1SVC} & A_{2new} & A_{3new} \\ P_{2SVC} & A_{SVC} & P_{3SVC} & B_{SVCnew} \\ K_2 & P_{4SVC} & K_{1new} & C_{4new} \\ G_1 & D_{SVC} & D_{1new\_SVC} & D_{2new\_SVC} \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta X_{SVC} \\ \Delta z \\ \Delta v \end{bmatrix} + \begin{bmatrix} E \\ 0 \\ 0 \\ 0 \end{bmatrix} \Delta U$$

$$\Delta\dot{X}_{sys\_svc} = A_{sys\_svc}\Delta X_{sys\_svc} + E_{SVC}\Delta U$$

$$A_{SYS\_SVC} = A_{SV1} - (A_{SV2} * (inv(A_{SV4}) * A_{SV3})$$

$$A_{SV1} = \begin{bmatrix} A_{1\,\mathrm{mod}} & P_{1svc} \\ P_{2svc} & A_{SVC} \end{bmatrix}, A_{SV2} = \begin{bmatrix} A_{2new} & A_{3new} \\ P_{3svc} & B_{svcnew} \end{bmatrix}$$

$$A_{SV3} = \begin{bmatrix} K_2 & P_{4svc} \\ G_1 & D_{SVC} \end{bmatrix}, A_{SV4} = \begin{bmatrix} K_{1new} & C_{4new} \\ D_{1new\_svc} & D_{2new\_svc} \end{bmatrix} \quad (38)$$

3.3 TCSC Modeling

As mentioned in the FACTS devices, the TCSC controller is a series of compensators and increases the transmission capacity by changing the appearance impedance of the transmission line. The structure of the TCSC series controller is shown in Fig. 21. The X_{TCSC} steady-state impedance is defined by Eq. 39 [21].

$$X_{TCSC}(\alpha) = \frac{X_C X_L(\alpha)}{X_L(\alpha) - X_C} \quad (39)$$

X_L (α) is defined according to Eq. 40 [21]:

$$X_L(\alpha) = X_L \frac{\pi}{\pi - 2\alpha - \sin 2\alpha}, \quad X_L \le X_L(\alpha) \le \infty \quad (40)$$

3.3.1 TCSC Controller Model

The TCSC controller is very similar in structure to the FC-TCR SVC controller. Equivalent impedance is presented for this series compensator in Eq. 41 [21, 22].

$$X_{TCSC} = X_C \left[\begin{array}{l} 1 - \frac{k}{k^2-1} \cdot \frac{\sigma + \sin\sigma}{\pi} + \\ \frac{4.k^2.\cos^2(\sigma/2)}{\pi(k^2-1)^2} \cdot (k \tan\frac{k\sigma}{2} - \tan\frac{\sigma}{2} \end{array} \right] \quad (41)$$

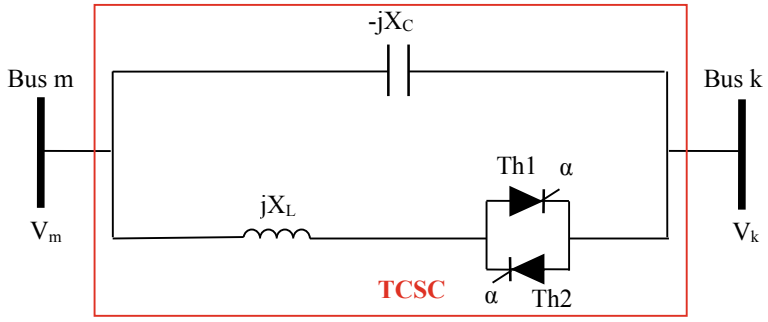


Fig. 21 TCSC series controller structure [21]

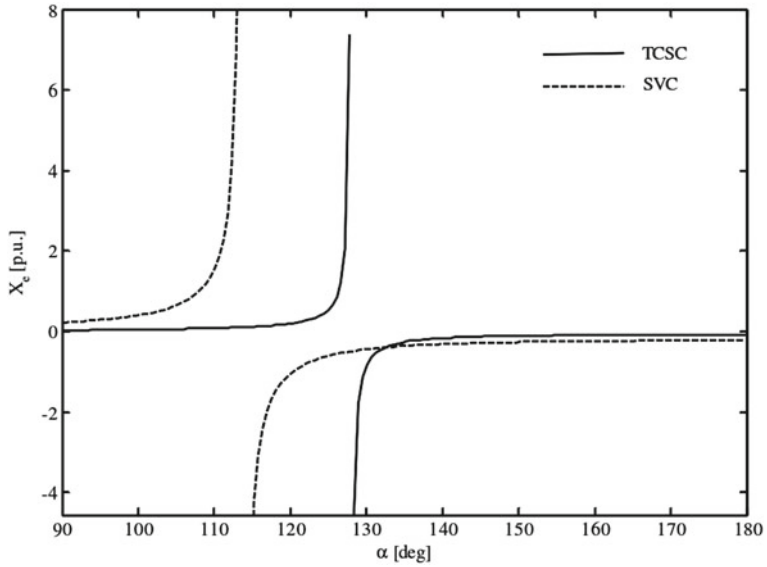


Fig. 22 Equivalent reactance to the SVC controller [22]

α: Fire angle.
σ: conductivity angle.
k: TCSC ratio.

The TCSC controller usually operates in the capacitor area. A very important point in the series controllers is the resonance. Figures 22 and 23 show the equivalent reactance and susceptance of SVC and TCSC controllers, respectively. The resonance in the SVC controller does not appear to be a problem, but in the TCSC series controller, a range must be defined for the fire angle. Therefore, the fire angle range for the TCSC controller is defined by considering the resonance point according to Eq. 42 [22]:

$$\alpha_{res} \leq \alpha \leq 180^{\circ} \tag{42}$$

3.3.2 TCSC Controller Model

As Fig. 24, the TCSC controller is located between bus k and m. If the TCSC controller losses are neglected, the power equilibrium and B_{TCSC} equations are defined according to Eq. 43 [21].

$$P_k = V_k V_m B_{TCSC} \sin(\theta_k - \theta_m)$$
$$Q_k = V_k^2 B_{TCSC} - V_k V_m B_{TCSC} \cos(\theta_k - \theta_m)$$

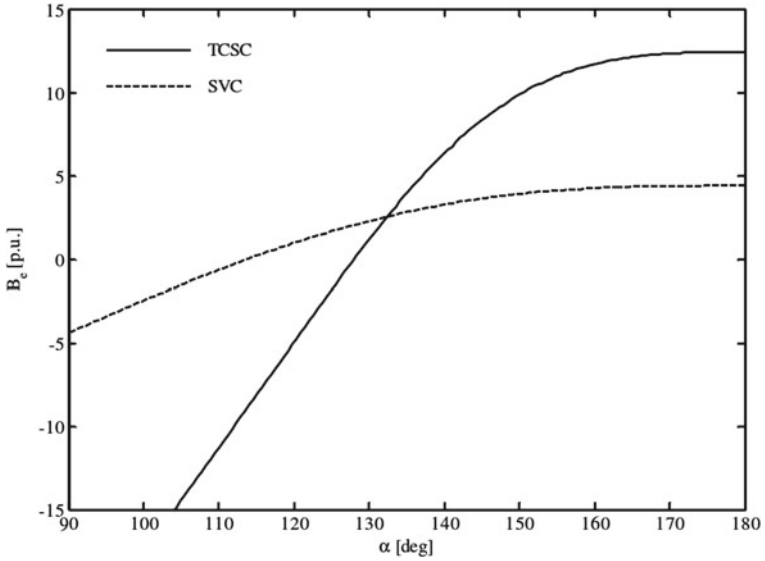


Fig. 23 Equivalent reactance to the TCSC controller [22]

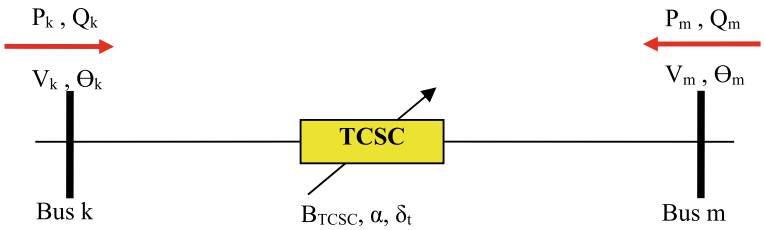


Fig. 24 TCSC connection in multi-machine power systems [21]

$$
\begin{aligned}
P_m &= V_k V_m B_{TCSC} \sin(\theta_m - \theta_k) \\
Q_m &= V_m^2 B_{TCSC} - V_k V_m B_{TCSC} \cos(\theta_m - \theta_k) \\
B_{TCSC} &= -\pi(k^4 - 2k^2 + 1)\cos k(\pi - \alpha)/ \\
&\left[\begin{array}{l}
X_C(\pi k^4 \cos k(\pi - \alpha) \\
-\pi \cos k(\pi - \alpha) - \\
2k^4 \alpha \cos k(\pi - \alpha) \\
+2\alpha k^2 \cos k(\pi - \alpha) \\
-k^4 \sin 2\alpha \cos k(\pi - \alpha) \\
+k^2 \sin 2\alpha \cos k(\pi - \alpha) \\
-4k^3 \cos^2 \alpha \sin k(\pi - \alpha) \\
-4k^2 \cos \alpha \sin \alpha \cos k(\pi - \alpha)
\end{array}\right]
\end{aligned}
\tag{43}
$$

Table 3 Control strategies for TCSC [21, 22]

Number	The name of the control strategy	Description
1	Reactance control	$B_{set} - B_{TCSC} = 0$
2	Power control	$P_{set} - P = 0$
3	Current control	$I_{set} - I = 0$
4	Transmission angle control	$\delta_{set} - \delta = 0$

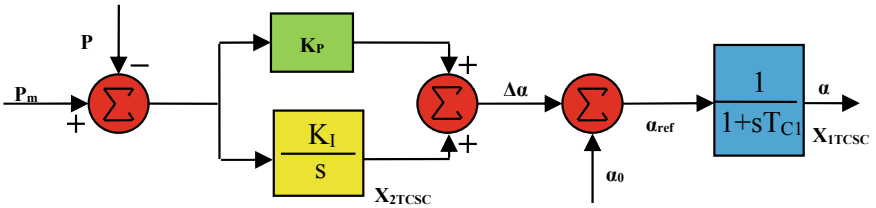


Fig. 25 Block diagram of the TCSC controller with the PI controller [21, 22]

Table 3 are shown 4 control strategies for TCSC in [21, 22].

Each of the strategies listed in the table above can be used to achieve TCSC controller goals. Here are some numerical studies of the power control strategy. Figure 25 shows the block diagram of the TCSC controller with the PI controller [21, 22].

X_{2TCSC}, $\dot{X}_{2TCSC}$ and $\dot{X}_{1TCSC}$ are defined according to Eq. 44 [21, 22].

$$\begin{aligned} X_{2TCSC} &= \frac{K_I}{S}(P_{set} - P) \\ \dot{X}_{2TCSC} &= K_I P_{set} - K_I P \\ \dot{X}_{1TCSC} &= \frac{-X_{1TCSC}}{T_{C1}} + \frac{X_{2TCSC}}{T_{C1}} + \frac{K_P P_{set}}{T_{C1}} - \frac{K_P P}{T_{C1}} + \frac{\alpha_0}{T_{C1}} \end{aligned} \tag{44}$$

After linearization we will have [21, 22]:

$$\begin{aligned} \Delta\dot{X}_{TCSC} &= A_{TCSC}\Delta X_{TCSC} + B_{TCSC}\begin{bmatrix} \Delta\theta_k \\ \Delta V_k \\ \Delta\theta_m \\ \Delta V_m \end{bmatrix} \\ \begin{bmatrix} \Delta P_k \\ \Delta Q_k \\ \Delta P_m \\ \Delta Q_m \end{bmatrix} &= C_{TCSC}\Delta X_{1TCSC} + D_{TCSC}\begin{bmatrix} \Delta\theta_k \\ \Delta V_k \\ \Delta\theta_m \\ \Delta V_m \end{bmatrix} \end{aligned} \tag{45}$$

The DAE model of the TCSC controller is defined by Eq. 46 [21, 22].

$$\begin{bmatrix} \Delta\dot{X} \\ \Delta\dot{X}_{TCSC} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} A_{1\,\mathrm{mod}} & P_{1TCSC} & A_{2new} & A_{new} \\ P_{2TCSC} & A_{TCSC} & B_{TCSC1new} & B_{TCSCnew} \\ K_2 & P_{4TCSC} & K_{1new} & C_{4new} \\ G_1 & C_{TCSC} & D_{1new\_TCSC} & D_{2new\_TCSC} \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta X_{TCSC} \\ \Delta z \\ \Delta v \end{bmatrix} + \begin{bmatrix} E \\ 0 \\ 0 \end{bmatrix} + \Delta U \tag{46}$$

As a result, we will have [21, 22].

$$\begin{aligned} \Delta\dot{X}_{SYS\_TCSC} &= A_{SYS\_TCSC}\Delta X_{SYS\_TCSC} + E_{TCSC}\Delta U \\ A_{SYS\_TCSC} &= A_{TC1} - (A_{TC2} * inv(A_{TC4})) * A_{TC3}) \\ A_{TC1} &= \begin{bmatrix} A_{1\,\mathrm{mod}} & P_{1TCSC} \\ P_{2TCSC} & A_{TCSC} \end{bmatrix} \\ A_{TC2} &= \begin{bmatrix} A_{2new} & A_{new} \\ B_{TCSC1new} & B_{TCSCnew} \end{bmatrix} \\ A_{TC3} &= \begin{bmatrix} K_2 & P_{4TCSC} \\ G_1 & C_{TCSC} \end{bmatrix} \\ A_{TC4} &= \begin{bmatrix} K_{1new} & C_{4new} \\ D_{1new\_TCSC} & D_{2new\_TCSC} \end{bmatrix} \end{aligned} \tag{47}$$

3.4 UPFC Modeling

As shown in Fig. 26, the UPFC Series–Parallel Controller consists of two series and parallel converters connected by a DC link [24].

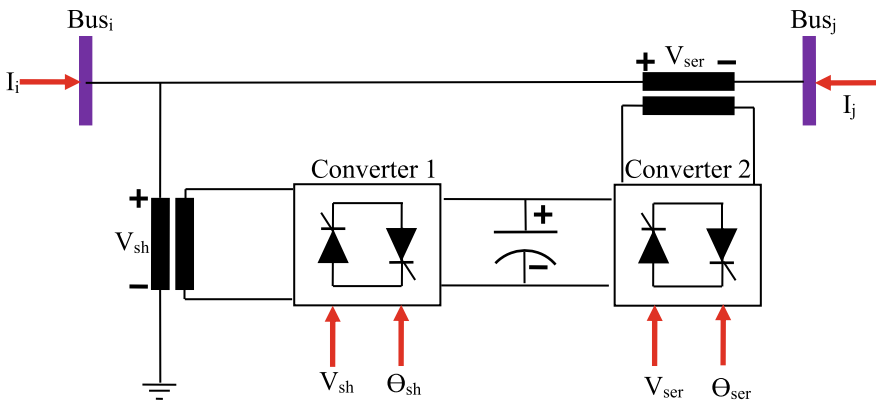


Fig. 26 UPFC structure [24]

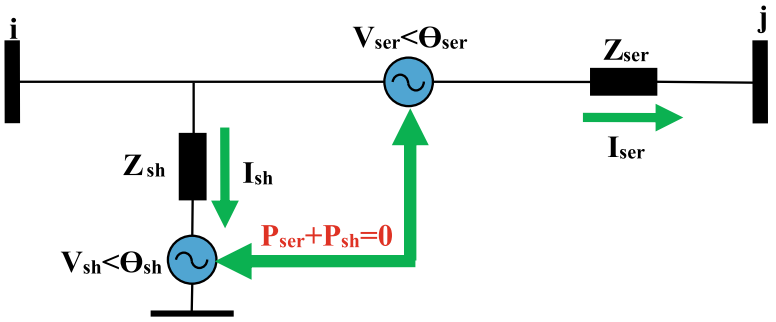


Fig. 27 UPFC injection model [24]

In the following, we will study UPFC numerical studies using Newton Raphson method. According to Fig. 27, the UPFC equivalent circuit consists of two ideal voltage sources. Z_{ser} and Z_{sh} represent the series and parallel impedance, respectively, and consist of a positive sequence of leakage resistance and inductance. The relationship between ideal voltage sources is expressed in Eq. 48 [24, 25].

$$\begin{aligned} V_{ser} &= V_{ser}(\cos\theta_{ser} + j\sin\theta_{ser}) \\ V_{sh} &= V_{sh}(\cos\theta_{sh} + j\sin\theta_{sh}) \end{aligned} \tag{48}$$

According to the injection model (equivalent circuit) UFPC, Eqs. 49 and 50 introduce the limitations of V_{ser} and Θ_{ser}, respectively [24, 25]:

$$V_{ser\,\min} \le V_{ser} \le V_{ser\,\max} \tag{49}$$

$$0 \le \theta_{ser} \le 360 \tag{50}$$

According to the UFPC equivalent circuit, Eqs. 51 and 52 show the constraints of V_{sh} and Θ_{sh}, respectively [24, 25].

$$V_{sh\,\min} \le V_{sh} \le V_{sh\,\max} \tag{51}$$

$$0 \le \theta_{sh} \le 360 \tag{52}$$

According to the above equations, the equations for active and reactive power of UPFC in node i are given in Eqs. 53 and 54, respectively [24, 25]:

$$\begin{aligned} P_i = {} & V_i^2 G_{ii} + V_i V_j \begin{pmatrix} G_{ij}\cos(\delta_i - \delta_j) \\ +B_{ij}\sin(\delta_i - \delta_j) \end{pmatrix} + V_i V_{sh} \begin{pmatrix} G_{ij}\cos(\delta_i - \theta_{ser}) \\ +B_{ij}\sin(\delta_i - \theta_{ser}) \end{pmatrix} \\ & + V_i V_{sh} \begin{pmatrix} G_{sh}\cos(\delta_i - \theta_{sh}) \\ +B_{sh}\sin(\delta_i - \theta_{sh}) \end{pmatrix} \end{aligned} \tag{53}$$

$$Q_i = -V_i^2 B_{ii} + V_i V_j \begin{pmatrix} G_{ij}\sin(\delta_i - \delta_j) \\ -B_{ij}\cos(\delta_i - \delta_j) \end{pmatrix} + V_i V_{ser} \begin{pmatrix} G_{ij}\sin(\delta_i - \theta_{ser}) \\ -B_{ij}\cos(\delta_i - \theta_{ser}) \end{pmatrix} + V_i V_{sh} \begin{pmatrix} G_{sh}\sin(\delta_i - \theta_{sh}) \\ -B_{sh}\cos(\delta_i - \theta_{sh}) \end{pmatrix} \quad (54)$$

The Eqs. of active and reactive power of UPFC in node j are also defined according to Eqs. 55 and 56 [24, 25].

$$P_j = V_j^2 G_{jj} + V_j V_i \begin{pmatrix} G_{ij}\cos(\delta_j - \delta_i) \\ +B_{ij}\sin(\delta_j - \delta_i) \end{pmatrix} + V_j V_{ser} \begin{pmatrix} G_{jj}\cos(\delta_j - \theta_{ser}) \\ +B_{jj}\sin(\delta_j - \theta_{ser}) \end{pmatrix} \quad (55)$$

$$Q_j = -V_j^2 B_{jj} + V_j V_i \begin{pmatrix} G_{ij}\sin(\delta_j - \delta_i) \\ -B_{ij}\cos(\delta_j - \delta_i) \end{pmatrix} + V_j V_{ser} \begin{pmatrix} G_{jj}\sin(\delta_j - \theta_{ser}) \\ -B_{jj}\cos(\delta_j - \theta_{ser}) \end{pmatrix} \quad (56)$$

The shunt converter reactive and active power relations are described in the following equation [24, 25].

$$P_{sh} = -V_{sh}^2 G_{sh} + V_{sh} V_i \begin{pmatrix} G_{sh}\cos(\theta_{sh} - \delta_i) \\ +B_{sh}\sin(\theta_{sh} - \delta_i) \end{pmatrix} \quad (57)$$

$$Q_{sh} = V_{sh}^2 B_{sh} + V_{sh} V_i \begin{pmatrix} G_{sh}\sin(\theta_{sh} - \delta_i) \\ -B_{sh}\cos(\theta_{sh} - \delta_i) \end{pmatrix} \quad (58)$$

The active power relation of the series converter is introduced in the following equation [24, 25]:

$$P_{ser} = -V_{ser}^2 G_{jj} + V_{ser} V_i \begin{pmatrix} G_{ij}\cos(\theta_{ser} - \delta_i) \\ +B_{ij}\sin(\theta_{ser} - \delta_i) \end{pmatrix} + V_{ser} V_j \begin{pmatrix} G_{jj}\cos(\theta_{ser} - \delta_j) \\ +B_{jj}\sin(\theta_{ser} - \delta_j) \end{pmatrix} \quad (59)$$

The reactive power relative of the series converter is introduced in the following equation [24, 25]:

$$Q_{ser} = -V_{ser}^2 B_{jj} + V_{ser} V_i \begin{pmatrix} G_{ij}\sin(\theta_{ser} - \delta_i) \\ -B_{ij}\cos(\theta_{ser} - \delta_i) \end{pmatrix} + V_{ser} V_j \begin{pmatrix} G_{jj}\sin(\theta_{ser} - \delta_j) \\ -B_{jj}\cos(\theta_{ser} - \delta_j) \end{pmatrix} \quad (60)$$

Y_{ii}, Y_{jj}, Y_{ij}, and Y_{ser} admittances are defined according to the following equations: [24]:

$$
\begin{aligned}
Y_{ii} &= G_{ii} + jB_{ii} = Z_{ser}^{-1} + Z_{sh}^{-1} \\
Y_{jj} &= G_{jj} + jB_{jj} = Z_{ser}^{-1} \\
Y_{ij} &= Y_{ji} = G_{ij} + jB_{ij} = -Z_{ser}^{-1} \\
Y_{ser} &= G_{ser} + jB_{ser} = -Z_{sh}^{-1}
\end{aligned} \tag{61}
$$

Assuming low switching losses of series and parallel converters, the following equation will be established [24, 25]:

$$P_{ser} + P_{sh} = 0 \tag{62}$$

In the next step of the study, we must combine the linearized power equations of the FACTS controller with the linear equations of the system (the rest of the network). Therefore, the following equation is defined [24]:

$$f(x) = [J][\Delta x] \tag{63}$$

$$[f(x)] = \left[\Delta P_i \ \Delta P_j \ \Delta Q_i \ \Delta Q_j \ \Delta P_{ij} \ \Delta Q_{ij} \ \Delta P_{bb} \right]^T \tag{64}$$

In the mentioned relationship, [ΔX] the solution vector, [J] the Jacobin matrix and ΔP_{bb} matrixes of the incompatibility of the power obtained from the Eq. 62. The solution vector is defined according to Eq. 65 [24].

$$[\Delta x] = \left[\Delta\delta_i \ \Delta\delta_j \ \frac{\Delta V_{sh}}{V_{sh}} \ \frac{\Delta V_j}{V_j} \ \Delta\theta_{ser} \ \frac{\Delta V_{ser}}{V_{ser}} \ \Delta\theta_{sh} \right]^T \tag{65}$$

Jacobin matrix is given according to [24, 25] in Eq. 66:

$$
J = \begin{bmatrix}
\frac{\partial P_i}{\partial \delta_i} & \frac{\partial P_i}{\partial \delta_j} & \frac{\partial P_i}{\partial V_{sh}} V_{sh} & \frac{\partial P_i}{\partial V_j} V_j & \frac{\partial P_i}{\partial \theta_{ser}} & \frac{\partial P_i}{\partial V_{ser}} V_{ser} & \frac{\partial P_i}{\partial \theta_{sh}} \\
\frac{\partial P_j}{\partial \delta_i} & \frac{\partial P_j}{\partial \delta_j} & 0 & \frac{\partial P_j}{\partial V_j} V_j & \frac{\partial P_j}{\partial \theta_{ser}} & \frac{\partial P_j}{\partial V_{ser}} V_{ser} & 0 \\
\frac{\partial Q_i}{\partial \delta_i} & \frac{\partial Q_i}{\partial \delta_j} & \frac{\partial Q_i}{\partial V_{sh}} V_{sh} & \frac{\partial Q_i}{\partial V_j} V_j & \frac{\partial Q_i}{\partial \theta_{ser}} & \frac{\partial Q_i}{\partial V_{ser}} V_{ser} & \frac{\partial Q_i}{\partial \theta_{sh}} \\
\frac{\partial Q_j}{\partial \delta_i} & \frac{\partial Q_j}{\partial \delta_j} & 0 & \frac{\partial Q_j}{\partial V_j} V_j & \frac{\partial Q_j}{\partial \theta_{ser}} & \frac{\partial Q_j}{\partial V_{ser}} V_{ser} & 0 \\
\frac{\partial P_{ij}}{\partial \delta_i} & \frac{\partial P_{ij}}{\partial \delta_j} & 0 & \frac{\partial P_{ij}}{\partial V_j} V_j & \frac{\partial P_{ij}}{\partial \theta_{ser}} & \frac{\partial P_{ij}}{\partial V_{ser}} V_{ser} & 0 \\
\frac{\partial Q_{ij}}{\partial \delta_i} & \frac{\partial Q_{ij}}{\partial \delta_j} & 0 & \frac{\partial Q_{ij}}{\partial V_j} V_j & \frac{\partial Q_{ij}}{\partial \theta_{ser}} & \frac{\partial Q_{ij}}{\partial V_{ser}} V_{ser} & 0 \\
\frac{\partial P_{bb}}{\partial \delta_i} & \frac{\partial P_{bb}}{\partial \delta_j} & \frac{\partial P_{bb}}{\partial V_{sh}} V_{sh} & \frac{\partial P_{bb}}{\partial V_j} V_j & \frac{\partial P_{bb}}{\partial \theta_{ser}} & \frac{\partial P_{bb}}{\partial V_{ser}} V_{ser} & \frac{\partial P_{bb}}{\partial \theta_{sh}}
\end{bmatrix} \tag{66}
$$

The initial conditions of the series and parallel sources are shown in Eqs. 67 and 68, respectively [24].

$$\theta_{ser}^0 = \tan^{-1}\left(\frac{P_{jref}}{CI}\right)$$

$$V_{ser}^{0} = \left(\frac{X_{ser}}{V_j^0}\right)\sqrt{P_{jref}^2 + CI^2}$$
$$CI = Q_{jref} - \frac{V_j}{X_{ser}}(V_j^0 - V_i^0) \quad if\ V_j^0 \neq V_i^0$$
$$CI = Q_{jref} \quad if\ V_j^0 = V_i^0 \tag{67}$$

$$\theta_{sh}^{0} = -\sin^{-1}\left(\frac{(V_i^0 - V_j^0)V_{ser}^0 X_{sh}\sin(\theta_{ser})}{V_{sh}^0 V_i^0 X_{ser}}\right) \tag{68}$$

4 Formulation the Problem of FACTS Devices Placement

To solve the problem of economic placement of FACTS controllers in power systems, the objective functions and constraints of the problem must first be determined. Therefore, we will describe the mentioned cases [26].

4.1 Objective Functions

Important objective functions in FACTS controller placement studies include voltage deviation, system overload, and active power losses. In other words, we are dealing with a multi-objective optimization problem when discussing the placement of FACTS devices. Formulation of a multi-objective optimization problem in Eq. 69 is given in which x is the decision variable, Ω is the solution range, C_j (x) is the equal constraint, H_k (x) is the unequal constraint, F_v (x) is the objective function of the voltage deviation, F_s (x) is the objective function of the system overload and F_{PL} (x) is the objective function of active power losses [26].

$$Min\ F(x) = [F_V(x), F_S(x), F_{PL}(x)]$$
$$Subject\ to\ x \in \Omega$$
$$C_j(x) = 0 \quad j = 1, ..., n$$
$$H_k(x) \leq 0 \quad k = 1, ..., p$$
$$F_V = \sum_i \left| V_i - V_i^{ref} \right|^2$$
$$F_S = \sum_j \left(\frac{S_j}{S_j^{\max}}\right)^2$$
$$F_{PL} = \sum_i P_{Li} \tag{69}$$

4.2 Constraints

Constraints on the optimal placement of FACTS controllers to two categories of equality constraints (e.g. active power equilibrium) and unequal constraints (e.g. allowable voltage range, reactive power output limit, FACTS devices limitations such as compensation range) Are divided [26].

4.2.1 Equality Constraints

In Eq. 70, the equilibrium of active and reactive power is mentioned [26]:

$$\begin{aligned} P_{Gi} - P_{Di} - \sum_{j=1}^{N} V_j [G_{ij} \cos(\theta_{ij}) + B_{ij} \sin(\theta_{ij}) = 0 \\ Q_{Gi} - Q_{Di} - \sum_{j=1}^{N} V_j [G_{ij} \sin(\theta_{ij}) - B_{ij} \cos(\theta_{ij}) = 0 \end{aligned} \tag{70}$$

4.2.2 Inequality Constraints

In Eq. 71, the limitation of reactive power generation and in Eq. 72 the limitation of some FACTS controllers such as SVC, TCSC and UPFC are given [26].

$$Q_{Gi}^{\min} < Q_{Gi} < Q_{Gi}^{\max} \quad for \quad i = 1, \ominus, N \tag{71}$$

$$\begin{aligned} &For \quad SVC \quad Q_C^{\min} < Q_{SVC} < Q_C^{\max} \\ &For \quad TCSC \quad r^{\min} < r_{TCSC} < r^{\max} \\ &For \quad UPFC \quad \gamma^{\min} < \gamma < \gamma^{\max} \\ &\qquad and \quad Q_C^{\min} < Q_{conv1} < Q_C^{\max} \end{aligned} \tag{72}$$

The objective functions used to optimally placement of FACTS devices in researchers' studies are shown in Fig. 28 [17].

4.3 Optimal Placement of TCSC, SVC, and UPFC

The correct choice of bus and line in the system under study to optimal placement of the FACTS devices controllers such as SVC, TCSC, and UPFC depends on the network structure. For example, SVCs (TCR, TSR, and TSC) are used to correct

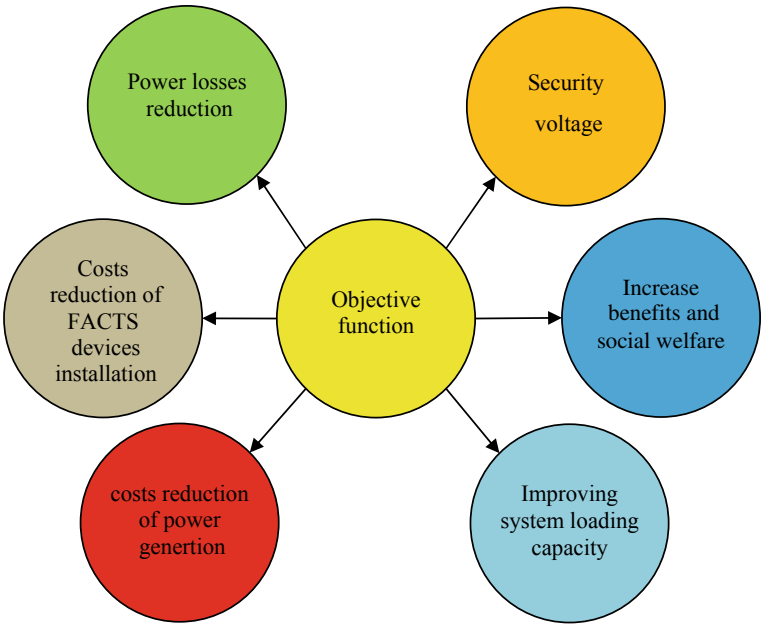


Fig. 28 Objective functions for optimal placement of FACTS devices [17]

the total reactive power flow between buses. TCSCs are commonly used to change line impedance to control the power flow in transmission lines. UPFC, as shown in Fig. 15, is a complete controller and is used to improve voltage and increase line capacity [27].

TCSCs are used as FACTS devices series controllers using the Lmn index specified in weak lines. Lmn is one of the appropriate indicators that can be analyzed in lines that are unstable. This indicator says that if Lmn is equal to one, the state of the critical line is. If Lmn is less than one, it indicates the stable state of the line. The Lmn index is given in Eq. 73 [27, 28].

$$Lmn = \frac{4XQ_r}{[V_s \mathrm{Sin}(\theta - \delta)]^2} \tag{73}$$

x represents the line, Q_r indicates the demand for reactive power, V_s indicates the bus voltage of the sending side, Θ indicates the difference between the bus angles, and δ indicates the impedance angle. The optimal placement of the SVC controller, which is connected in parallel, is done by analyzing the PV curves. It should be noted that PV curves are made using the CPF technique. The predictive and correction scheme used in CPF is shown in Fig. 29. Optimal placement by UPFC is done on lines where the high active power flow has, it should also be noted that UPFCs are placed at the beginning of the buses to play their role as a complete FACTS devices series–parallel controller [27, 29].

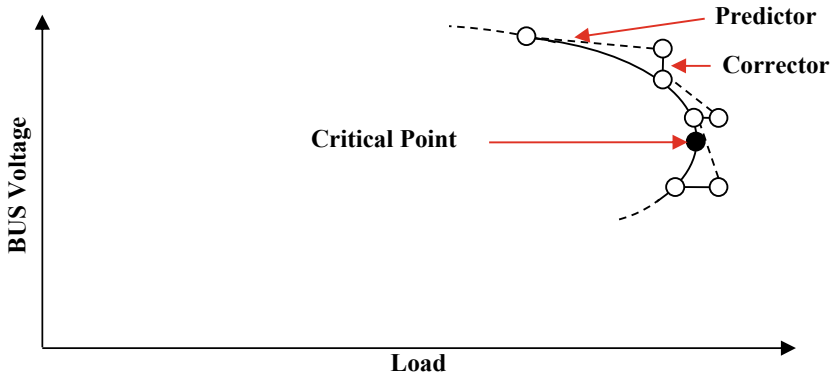


Fig. 29 Predictive and correction scheme used in CPF [29]

The objective functions and the optimal location and size limitations of FACTS devices [27] are shown in Table 4:

Table 4 The objective functions and the optimal location and size limitations of FACTS devices [27]

Number	Objective function	Description
1	$\min\{C_{PL} + C_{FACTS}\}$ $C_{PL} = (Active \quad Power \quad loss) \times (0.09\$/Kwh) \times 365 \times 24$ $C_{FACTS} = C_{TCSC} + C_{SVC} + C_{UPFC}$	C_{PL}: the cost of active loss function C_{FACTS}: the cost function of the FACTS devices
2	$C_{TCSC} = 0.0015t^2 - 0.7130t + 153.75(\frac{\$}{KVAr})$	C_{TCSC}: cost function of TCSC
3	$C_{SVC} = 0.0003s^2 - 0.3051s + 127.38(\frac{\$}{KVAr})$	C_{SVC}: cost function of SVC
4	$C_{UPFC} = 0.0003u^2 - 0.2961u + 188.22(\frac{\$}{KVAr})$	C_{UPFC}: cost function of UPFC
Number	Constraints	Description
1	$0.95 \leq V_j \leq 1.05$	Bus voltage limit
2	$S_{\min} \leq S_L \leq S_{\max}$	Thermal limit
3	$Q_{g,\min} \leq Q_g \leq Q_{g,\max}$	Reactive power limit of generators
4	$T_{i,\min} \leq T_i \leq T_{i,\max}$	Transformer tap setting limit
5	$-0.9 \leq Z_{SVC} \leq 0.9(pu)$	SVC size limit
6	$-0.8X_L \leq X_{TCSC} \leq 0.2X_L(pu)$	TCSC size limit

Example 1: In the 14-bus distribution system is shown in Fig. 30, the 5 SVCs placement perform with a capacity of 5 MW. Then, the SVCs place in the selected locations. Compare the voltage profile with SVC mode. Perform the placement with the Newton Raphson method to improve the voltage profile. Select buses to be placed in buses without any power plant. Then place the SVC in the selected locations and use Newton Raphson method to check the system's performance in terms of voltage and power losses before and after the SVC.

Solution: In this example, the SVC was replaced to improve the voltage profile. The proposed method is that to placement the first bus voltages are obtained by Newton–Raphson method. Then, according to the obtained voltages, 5 buses with the lowest voltage are selected for SVC installation. The buses selected for placement are: 4-5-10-13-14. Then, to check the performance of SVC in the system, we put the information about SVC in the base system. By placing SVC in these buses and comparing the voltage before and after SVC as shown in Fig. 31, we see that the

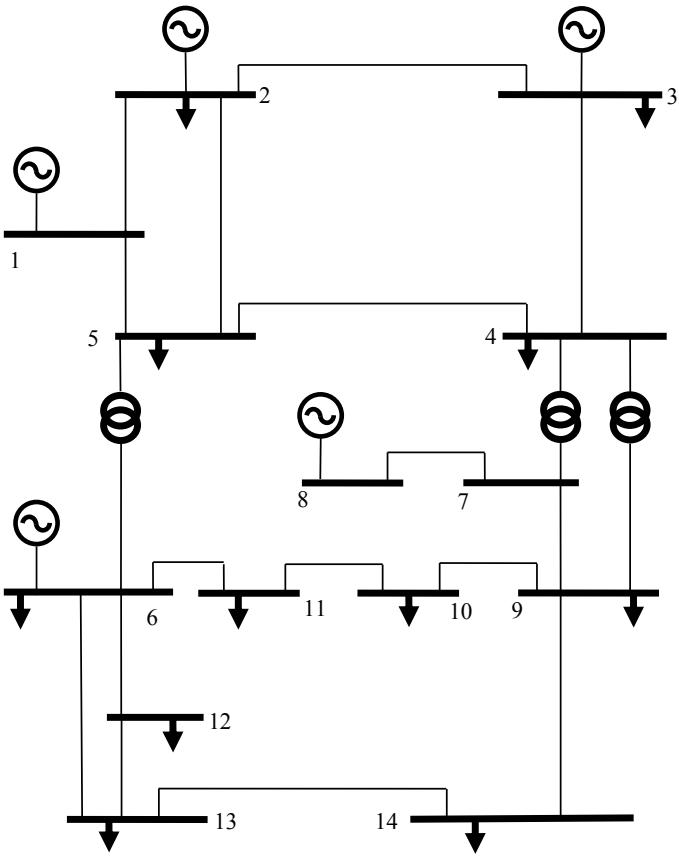


Fig. 30 IEEE 14-bus test system

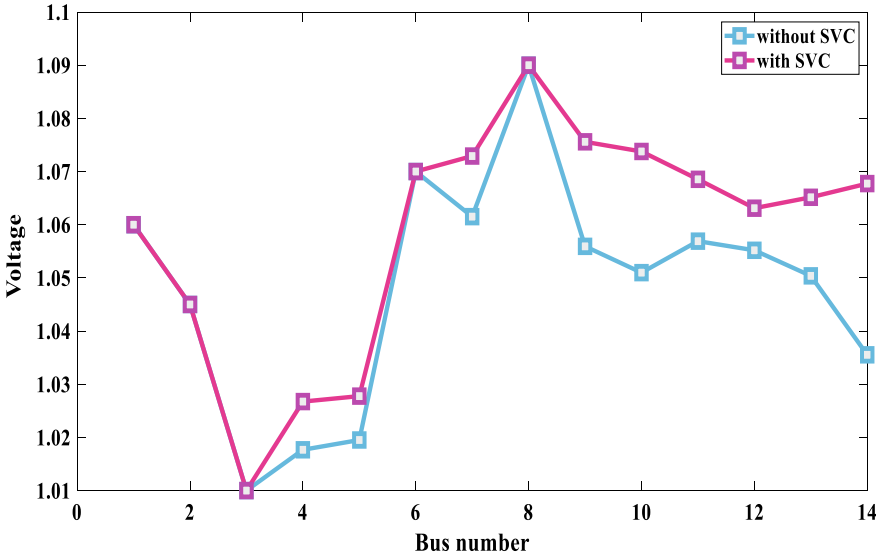


Fig. 31 Voltage profile before and after SVC placement

voltage has improved significantly that indicate suitable performance of in improving system voltage levels. The active and reactive power losses without SVC mode are 13,393 MW and 54.54 MVAR respectively. While the power losses after SVC installation are 13,345 MW and 54.22 MVAR. By adding SVC, 48 MW is reduced for the active power and 320 kVAR for reactive power, indicating suitable performance of SVC in reducing system power losses. To better check the voltage profile, we can check the voltage stability indices in offline in both cases using the following relationships:

$$VSF_{i+1} = (2V_{i+1} - V_i) \tag{74}$$

$$VSF = \frac{\sum_{i=2}^{N_b} VSF_{i+1}}{(N_b - 1)} \tag{75}$$

That VSF is: Voltage stability Factor. Considering the above two equations, we check the voltage stability in each case with and without SVC. without SVC mode, VSF is 1.0514 p.u. While in SVC mode, the VSF increases by 1.0625 p.u.

Example 2: In Fig. 30, install one UPFC with the goal of improving the voltage profile. Then draw the voltage profile before and after the placement. Also determine the active and reactive power losses before and after the placement UPFC using the Newton Raphson method.

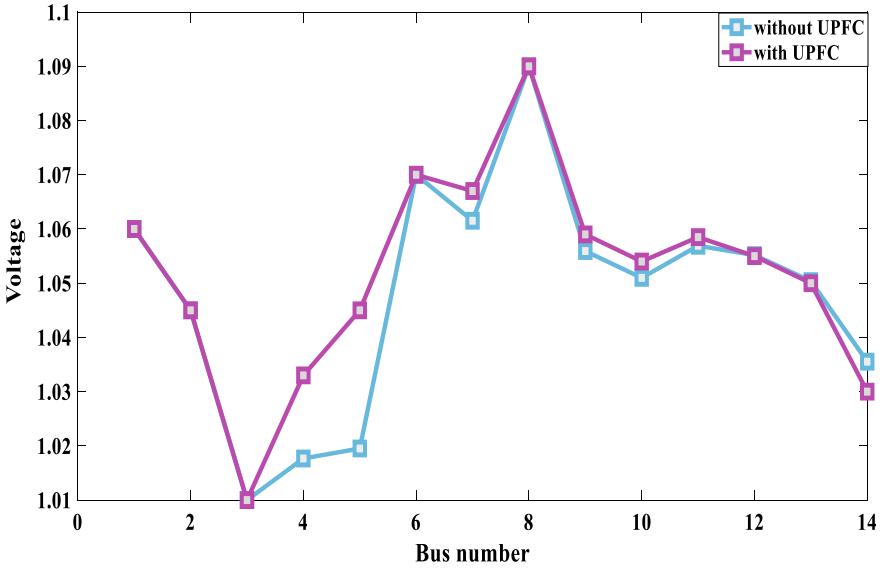


Fig. 32 Voltage profile before and after UPFC placement

Solution: The optimal place for UPFC is selected in the line between buses 2 and 5. The voltage profile before and after placing the UPFC is shown in Fig. 32. The voltage is improved significantly after installing UPFC. The active and reactive power losses without UPFC mode are 13,393 MW and 54.54 MVAR respectively, While the power losses after installation of UPFC are 13,368 MW and 54.38 MVAR.

5 Conclusions

In this chapter, first the benefits of FACTS devices have been studied and proposed as a suitable solution for improvement of power system characteristics. Then mathematical modeling of some widely used FACTS devices such as SVC, TCSC and UPFC are described. In the following, numerical optimization methods for FACTS devices analysis are presented. Numerical optimization methods in many cases can ensure global optimization. Among the introduced methods, one of the most widely used methods in studies of power system is the Newton Raphson method. The important functions and constraints in the discussion of placement and modeling of FACTS devices are presented in this chapter for controlling power flow, regulating the bus voltage and reducing of power losses. The Newton Raphson method is considered for the equation solving of the objective functions. Finally, to demonstrate the advantage of FACTS devices, simulations for SVC and UPFC are performed on the IEEE 14

bus system. According to the results, optimal placement of FACTS devices improves most of the power system specifications.

References

1. Bayliss C, Hardy B (2012) Transmission and distribution electrical engineering. Elsevier
2. Hingoranl NG (2000) Understanding facts. Wiley
3. Singh AK, Pal B (2018) Dynamic estimation and control of power systems. Academic Press
4. Xu K (2014) Time dependent performance analysis of wireless networks. Dissertation, University of Pittsburgh
5. Burden RL, Faires JD, Burden AM (2010) Numerical analysis: Cengage Learning
6. Cheney EW, Kincaid DR (2012) Numerical mathematics and computing. Cengage Learning
7. Hamming R (2012) Numerical methods for scientists and engineers. Courier Corporation
8. Babolian E (2020) Numerical analysis 1. Payame Noor University. Pnueb. https://www.pnueb.com/courses/ls/228/86. Accessed 23 May 2020
9. Ralston A, Rabinowitz P (2001) A first course in numerical analysis. Courier Corporation
10. Xing H et al (2014) Distributed bisection method for economic power dispatch in smart grid. IEEE Trans Power Syst 30(6):3024–3035
11. Seunghyun C (2015) Integration of mathematics for sustainable energy applications. In: 2015 ASEE annual conference & exposition, Seattle, Washington, June 2015, ASEE Conferences
12. Choopani K, Hedayati M, Effatnejad R (2020) Self-healing optimization in active distribution network to improve reliability, and reduction losses, switching cost and load shedding. Int Trans Electr Energy Syst 30(5):e12348
13. Choopani K, Effatnejad R, Hedayati M (2020) Coordination of energy storage and wind power plant considering energy and reserve market for a resilience smart grid. J Energy Storage 30:101542
14. Vazinram F et al (2020) Self-healing model for gas-electricity distribution network with consideration of various types of generation units and demand response capability. Energy Convers Manag 206:112487
15. Vazinram F et al (2019) Decentralised self-healing model for gas and electricity distribution network. IET Gener Transm Distrib 13(19):4451–4463
16. Ghahremani E, Kamwa I (2012) Optimal placement of multiple-type FACTS devices to maximize power system loadability using a generic graphical user interface. IEEE Trans Power Syst 28(2):764–778
17. Ahmad AL, Sirjani R (2019) Optimal placement and sizing of multi-type FACTS devices in power systems using metaheuristic optimisation techniques: an updated review. Ain Shams Eng J
18. Dobson I et al (1988) A model of voltage collapse in electric power systems. In: Proceedings of the 27th IEEE conference on decision and control, vol 3. IEEE
19. Liaw D-C, Chen J-W, Huang Y-H (2018) A Numerical study of the SVC-controlled electric power system dynamics. Int J New Technol Res 4(6):10–15
20. Ambriz-Perez H, Acha E, Fuerte-Esquivel CR (2000) Advanced SVC models for Newton-Raphson load flow and Newton optimal power flow studies. IEEE Trans Power Syst 15(1):129–136
21. Singh B et al (2018) Utilities of differential algebraic equations (DAE) model of SVC and TCSC for operation, control, planning & protection of power system environments. Int J Futur Revolut Comput Sci Commun Eng 4(1):61–67
22. Welhazi Y et al (2014) Power system stability enhancement using FACTS controllers in multimachine power systems. J Electr Syst 10(3):276–291
23. Pali BS, Bhowmick S, Kumar N (2012) Newton-Raphson power flow models of static VAR compensator. In: 2012 IEEE 5th India international conference on power electronics (IICPE). IEEE

24. El-Sadek MZ et al (2007) Comprehensive Newton-Raphson model for incorporating unified power flow controller in load flow studies. J Eng Sci 35(1):189–205
25. Acha E et al (2004) FACTS: modelling and simulation in power networks. Wiley
26. El-Arini MM, Ahmed RS (2012) Optimal location of FACTS devices to improve power systems performance. J Electr Eng 12(3):73–80
27. Nadeem M et al (2020) Optimal placement, sizing and coordination of FACTS devices in transmission network using whale optimization algorithm. Energies 13(3):753
28. Moghavvemi M, Omar FM (1998) Technique for contingency monitoring and voltage collapse prediction. IEE Proc Gener Transm Distrib 145(6):634–640
29. Ajjarapu V, Christy C (1992) The continuation power flow: a tool for steady state voltage stability analysis. IEEE Trans Power Syst 7(1):416–423

Correction to: Numerical Methods for Energy Applications

Naser Mahdavi Tabatabaei and Nicu Bizon

Correction to:
N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems, https://doi.org/10.1007/978-3-030-62191-9

The original version of these chapters "Numerical Methods in Selecting Location of Distributed Generation in Energy Network" and "Numerical Methods for Power System Analysis with FACTS Devices Applications" are inadvertently published with branch name after the institution name. It has now been corrected.

The updated version of these chapters can be found at
https://doi.org/10.1007/978-3-030-62191-9_34
https://doi.org/10.1007/978-3-030-62191-9_35

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9_36

Index

N. Mahdavi Tabatabaei and N. Bizon (eds.), *Numerical Methods for Energy Applications*, Power Systems,
https://doi.org/10.1007/978-3-030-62191-9

C

D

F

G

S