Georgios M. Kopanos · Pei Liu
Michael C. Georgiadis *Editors*

Advances in Energy Systems Engineering

Springer

Advances in Energy Systems Engineering

Georgios M. Kopanos · Pei Liu
Michael C. Georgiadis
Editors

Advances in Energy Systems Engineering

Springer

Editors
Georgios M. Kopanos
Cranfield University
Bedfordshire
UK

Michael C. Georgiadis
Aristotle University of Thessaloniki
Thessaloniki
Greece

Pei Liu
Tsinghua University
Beijing
China

ISBN 978-3-319-42802-4 ISBN 978-3-319-42803-1 (eBook)
DOI 10.1007/978-3-319-42803-1

Library of Congress Control Number: 2016948241

Printed on acid-free paper

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

Preface

Contemporary energy systems are changing fast and exhibit increasingly complex features. This is mainly due to the fact that conversion and supply of energy is not the only mission of energy systems nowadays. Besides providing energy, energy systems are also expected to secure their resources so that they can be operated in a long-term future, to minimize their impacts on our ecology and environmental system during construction and operation, to provide economically affordable energy to maximize their benefits to a larger population, and to contribute to the mitigation of carbon emissions for the globe's climate. New technologies have been continuously emerged toward addressing some of the above-mentioned targets. With the implementation of these technologies, renewable energy has been increasingly exploited and used, energy efficiency of conventional power generation has been continuously improved, and energy end-consumers have become more energy efficient and environmentally benign.

Albeit these achievements, large gaps are still observed between the current energy supply and demand sides. Dependence on conventional fossil fuels is still strong in most parts of the world. Alternative fossil fuel energy, shale oil and gas for instance, still needs further improvement during extraction and conversion processes. Renewable energy systems, especially wind and solar, still face the problem of intermittency, which is the major challenge against larger-scale applications of them. Novel energy consumption technologies, taking electric vehicle as an example, require huge investment in infrastructure and may exaggerate further the intermittency not only from the supply side but the demand side as well. All these existing problems of modern energy systems indicate clearly that the existing means of designing and operating the energy systems are far from perfect, and there is still a large potential for improvements.

Energy systems engineering provides a methodological scientific framework to arrive at realistic integrated solutions to complex energy systems problems, by adopting a holistic systems-based approach. This book demonstrates the potential of an energy systems engineering-based approach to systematically quantify different options at different levels of complexity (i.e., technology, plant, energy supply

chain network) through state-of-the-art modeling, simulation, control, and optimization-based frameworks. The successful implementation of these approaches in a number of real-life case studies highlights further the significance of this integrated system-wide approach. The aim of this book is to mirror the importance of fundamental and applied research in energy systems engineering applications, developing mechanisms for the transfer of the new methodology to industry.

This manuscript presents an in-depth account of recent novel methodologies, frameworks, and tools for the simulation, modeling, and optimization of integrated energy systems. This book contains 28 chapters of high-quality contributions from international leading researchers in the field of process and energy systems engineering. This book is mainly intended for academics, researchers, and industrial practitioners in energy systems engineering, who are involved in model-based energy systems activities, across engineering and applied science disciplines, as well as for educational purposes both in academia and industry.

According to the application domain and nature, the research works in this book are categorized into four parts, namely (I) shale gas, refineries, and polygeneration systems, (II) power and transport systems, (III) planning and operations of energy systems, and (IV) low-carbon energy systems. Each part comprises seven chapters. A brief description of the chapters of this book follows.

Part I focuses on the optimal design and operation of energy systems in shale gas, refining, and chemical/power industry. In Chap. 1, El-Halwagi and coworkers present a multi-period mathematical programming approach to address scheduling issues of shale gas production where needed infrastructure is not fully developed. In the subsequent chapter, You and coworkers provide a comprehensive review of the most recent approaches in four major challenging research areas within the shale gas industry, namely design and planning of shale gas supply chain, water management, sustainability, and shale gas processing. In Chap. 3, Reklaitis and coworkers illustrate how energy systems engineering methods can be applied to real-life problems via shale gas supply chains and interconnected power systems as two case studies. In Chap. 4, a computational fluid dynamics model of hydrogen production via steam methane reforming, for faster online operation optimization purposes, is presented by Christofides and coworkers. Dua and coworkers in Chap. 5 present data-based model reduction techniques for computationally complex process models and applications in oil refineries. In Chap. 6, Wang introduces coal polygeneration combining circulating fluidized bed and pyrolysis, from conceptual design, engineering application, to thermodynamic and economic analyses. In Chap. 7, Yi and coworkers present a carbon dioxide circulating design for coal polygeneration energy systems, along with a performance comparison with other types of polygeneration energy systems.

Part II concentrates on energy systems issues related to power and transport systems. In Chap. 8, Chen and coworkers present an approach for evaluating the contribution of energy storage as a means of supporting the renewables integration in power systems. In Chap. 9, Baldea discusses the possibility and opportunities to use chemical processes as energy storage utilities in the context of increasing demand for demand-side response. In the subsequent chapter, Liu and coworkers

analyze the future gap between supply and demand of transport fuel in China up to 2050, with the aid of a superstructure-based nationwide refining sector planning model. In Chap. 11, Pan presents a decomposition approach for China's regional energy consumption by sector and analyzes energy flows across the country. In Chap. 12, Ou and coworkers present comprehensive life-cycle assessment of alternative transport fuel pathways, together with a comparison with previous studies. In Chap. 13, Pan and coworkers use a Long-range Energy Alternatives Planning System (LEAP) to present energy consumption projections in the city of Shanghai, China. Voutetakis and coworkers in Chap. 14 provide a comprehensive review of the application of model predictive control on fuel cell systems, covering all stages from conceptual design to implementation.

PART III mainly focuses on the planning and operation of various scales and types of energy systems. In Chap. 15, Georgiadis and coworkers present a mixed-integer linear programming approach to power systems planning problems with constraints from an hourly level to a long-term capacity expansion level. In Chap. 16, Pistikopoulos and coworkers propose a mixed-integer dynamic programming approach for the integrated design and control of combined heat and power systems. In Chap. 17, Papageorgiou and coworkers propose a mathematical programming approach for the energy consumption-end management at smart homes. In Chap. 18, Grossmann and coworkers discuss energy demand-side management issues though a cryogenic air separation plant as a case study. In Chap. 19, Kopanos and coworkers provide optimization-based approaches for the integrated operational and maintenance planning of compressors networks, which are a part of industrial air separation plants. In Chap. 20, Harjunkoski and coworkers illustrate how advanced scheduling algorithms can be used in energy-intensive industries to cope with supply-side fluctuations. In Chap. 21, Feng and coworkers show how optimization-based approaches can be applied to address heat integration issues at intraplant and interplant levels.

Part IV focuses on low-carbon energy systems. In Chap. 22, Gani and coworkers illustrate how to use superstructure-based approaches for the process synthesis of biofuels. In Chap. 23, Manenti and coworkers propose generalized disjunctive programming and mixed-integer nonlinear programming methods for downstream production problems in biorefineries. In Chap. 24, Puigjaner and coworkers illustrate the application of optimization methods on biomass supply chain problems via three case studies of rather different scales. In Chap. 25, Maravelias and coworkers address biomass procurement problems via a composite-curve-based method. In Chap. 26, Gao and coworkers propose a conceptual design of a hydrogen storage-based energy system which can better accommodate intermittent power. In Chap. 27, Wang and coworkers provide a comprehensive overview of monitoring and diagnosis issues in wind turbines. In Chap. 28, Guillén-Gosálbez and coworkers provide a multi-objective optimization approach for the design of buildings considering both economic and environmental issues.

This book covers the most recently developed state-of-the-art computer-aided methodologies, algorithms, and tools for energy systems planning, design, operation, and control from unit, process, plant levels to supply chain and system-wide

levels, and illustrates their applications using real-life case studies. Notwithstanding, the research works presented here may not be able to cover all the underlying fields because the relatively newly born field of energy systems engineering is highly fast developing. Through this book, a much wider audience can be introduced to energy systems engineering and the fast development momentum of this field can be maintained.

Bedfordshire, UK — Georgios M. Kopanos
Beijing, China — Pei Liu
Thessaloniki, Greece — Michael C. Georgiadis

Contents

Part I
Shale Gas, Refineries and Polygeneration Systems

Chapter 1
Optimal Planning of Infrastructure for the Supply Chain of Shale Gas

Karla Arredondo-Ramírez, José María Ponce-Ortega and Mahmoud M. El-Halwagi

Abstract As a result of the discovery of substantial reserves of shale gas, it is important to develop effective strategies for the establishment of infrastructure supporting the growth of shale gas production and monetization. This chapter presents an optimization approach for the optimal planning of shale gas exploitation and infrastructure development in places that lack the infrastructure needed for production, treatment, and distribution. A multi-period optimization approach is presented to account for the variability in market. The different components of the infrastructure, the production schedules, and the time-value of money to maximize the net present value of the infrastructure are considered in the optimization model. The applicability of the proposed approach is shown through a case study from Mexico, where there are enormous reserves of shale gas that require the development of infrastructure. The results show attractive economic results for the exploitation and distribution of gas to satisfy a certain demand.

Nomenclature

Sets

i Well
m Market
p Multi-well pad
r Processing plant/Compression unit
t Time period

Parameters

β^{proc} Losses factor associated to the compression units and processing plants

K. Arredondo-Ramírez · J.M. Ponce-Ortega
Universidad Michoacana de San Nicolás de Hidalgo, Morelia, México

M.M. El-Halwagi (✉)
Texas A&M University, College Station, Texas, USA
e-mail: el-halwagi@tamu.edu

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_1

γPC	Exponent for the capital cost function for processing units
γpipe	Exponent for the capital cost function for pipelines
γwell	Exponent for the capital cost function for wells
$\mathrm{CapCostPC}_{r}^{\max}$	Maximum capital cost in processing plant/compression unit r
$\mathrm{CapCostPipeDisMk}_{\mathrm{r,m}}^{\max}$	Maximum capital cost for pipelines in processing plant/compression unit r to market m
$\mathrm{CapCostPipeWellPC}_{\mathrm{i,p,r}}^{\max}$	Maximum capital cost for pipelines in well i from well-pad p to processing plant/compression unit r
$\mathrm{CapCostWell}_{\mathrm{i,p}}^{\max}$	Maximum capital cost for well i in the well-pad p
FC_{r}^{PC}	Unit fixed cost for processing plant/compression unit r
$\mathrm{FC}_{r,m}^{PipeDisMk}$	Unit fixed cost for pipelines in processing plant/compression unit r to market m
$\mathrm{FC}_{i,p,r}^{PipeWellPC}$	Unit fixed cost for pipelines in well i from well-pad p to processing plant/compression unit r
$\mathrm{FC}_{\mathrm{i,p}}^{\mathrm{well}}$	Unit fixed cost for well i in the well-pad p
$\mathrm{f}_{\mathrm{m,t}}^{\mathrm{de-mk}}$	Demand in the market m over the time period t
$\mathrm{f}_{\mathrm{p}}^{\max}$	Maximum natural gas available in the well-pad p
$\mathrm{f}_{\mathrm{r,m}}^{\mathrm{max-dis-mk}}$	Maximum natural gas distributed from the processing plant/compression unit r to market m
$\mathrm{f}_{\mathrm{r}}^{\mathrm{max-pc}}$	Maximum natural gas in the well-pad p to processing plant/compression unit r
K_{Ft}	Factor used to annualize the inversion
$\mathrm{N}_{\mathrm{p}}^{\mathrm{max-well}}$	Maximum number of wells in the well-pad p
VP_{t}	Present value factor over time period t
$\mathrm{VC}_{\mathrm{i,p}}^{\mathrm{well}}$	Unit variable cost for well i in the well-pad p
VC_{r}^{PC}	Unit variable cost in processing plant/compression unit r
$\mathrm{VC}_{r,m}^{PipeDisMk}$	Unit variable cost for pipes in processing plant/compression unit r to market m
$\mathrm{VC}_{i,p,r}^{PipeWellPC}$	Unit variable cost for pipes in well i from well-pad p to processing plant/compression unit r

Binary Variables

$y_{r,t}^{PC}$	Binary variable associated to the existence of the processing plant/compressing unit r over the time period t
$y_{r,m,t}^{pipe-pc-mk}$	Binary variable associated to the existence of pipelines in processing plant/compression unit r to market m
$y_{i,p,r,t}^{pipe-well-pc}$	Binary variable associated to the existence of pipelines in well i from well-pad p to processing plant/compression unit r
$y_{i,p,t}^{well}$	Binary variable associated to the existence of well i in the well-pad p over time period t

Variables

$CapCostPC_r$	Capital cost for the processing plant/compression unit r
$CapCostPipeDisMk_{r,m}$	Capital cost associated to the pipelines used for natural gas from processing plant/compression unit r to market m
$CapCostPipeWellPC_{i,p,r}$	Capital cost associated to pipelines used for natural gas from well i in the well-pad p to the processing plant/compression unit r
$CapCostWell_{i,p}$	Capital cost for well i in the well-pad p
$f_{i,p,r,t}$	Shale gas from well i in the well-pad p to the processing plant/compression unit r over time period t
$f_{r,m,t}^{dis-mk}$	Natural gas distributed from the processing plant/compression unit p to the market m, over the time period t
$f_{r,m}^{dis-mk-pipe-cap}$	Natural gas capacity from the processing plant/compression unit r to the market m
$f_{r,t}^{in-pc}$	Shale gas inlet to processing plant/compression unit p over the time period t
$f_{r}^{in-pc-cap}$	Natural gas capacity inlet to the compression unit/processing plant r
$f_{m,t}^{mk}$	Natural gas distributed in the market m over the time period t
$f_{r,t}^{out-pc}$	Natural gas outlet from processing plant/compression unit r over the time period t
$f_{i,p,r}^{pc-pipe-cap}$	Natural gas capacity from well i in the well-pad p to processing plant/compression unit r
$f_{i,p,t}^{well}$	Shale gas extracted from well i in the well-pad p over time period t
$f_{i,p}^{well-cap}$	Natural gas capacity in the well i in the well-pad p

1.1 Global Perspective of Shale Gas Production

It is projected that by 2035 the global energy consumption increases by 37 % from today's levels with India and China accounting for half of the growth according to BP (2015). In order to give a broad view about hydrocarbon fuels, in 2014, 80.6 % of the total primary energy supply was based on fossil fuels such as oil, natural gas and coal. In 2014, the annual production of natural gas was 3,524 bcm (OECD/IEA 2015).

1.2 Strategic Planning and Operation of Shale Gas Fields

Process systems engineering approaches have been proposed in the area of shale gas production and monetization. With respect to the strategic planning shale gas fields, Rahman et al. (2001) presented an optimization approach for the hydraulic fracturing accounting for the return on investment. Knudsen et al. (2014) incorporated a Lagrangian approach for scheduling shut-in times in multi-well pads to stimulate the shale gas production in different wells. Cafaro and Grossmann (2014) proposed a mixed-integer nonlinear programming (MINLP) model for the optimal planning in shale gas production. He and You (2015) integrated shale gas processing with ethylene production incorporating exergy and the techno-economic analysis.

It should be noted that there are many shale plays that are rich in gas reserves but poor in infrastructure. Therefore, the goal of this chapter is to present an optimization approach for designing the entire infrastructure from shale gas resources to maximize the net present value (NPV) of the entire supply chain given a the demand of natural gas in the market under study while accounting for the following issues:

- Determining the number of wells drilling and operating in a certain multi-well pad in a given time period.
- Scheduling the drilling to generate an optimal planning for each source.
- Predict the drop in the profile production of the wells.
- Planning the shale gas production in the play over the time periods while accounting for the changes to the demand through the time.
- Determining the gas-processing plant capacity as well as the operation startup time for the shale gas play production.
- Designing and operating the transmission network over the time while accounting for the construction and startup times.
- Considering the long-term capital and the operating costs for the project.

1.3 Problem Statement

The addressed problem is schematically shown by Fig. 1.1. The following are key aspects of the problem: Consider a set of potential multi-well pads in the same geographical shale formation with a specific number of wells ($i = \{i \mid i = 1, 2, \ldots, Nn\}$), time periods ($t = \{t \mid t = 1, 2\ Nn\}$, markets ($m = \{m \mid m = 1,2,\ldots, \}$) and a set of potential processing plants ($p = \{p \mid p = 1, 2, \ldots, Nn\}$). The objective is to determine the optimal infrastructure, configuration, and production schemes for the production of natural gas using hydraulic fracturing, gas treatment plants and distribution network of natural gas for a specific time horizon. The objective function is the maximization of the net present value for the profit (NPVPROFIT), which

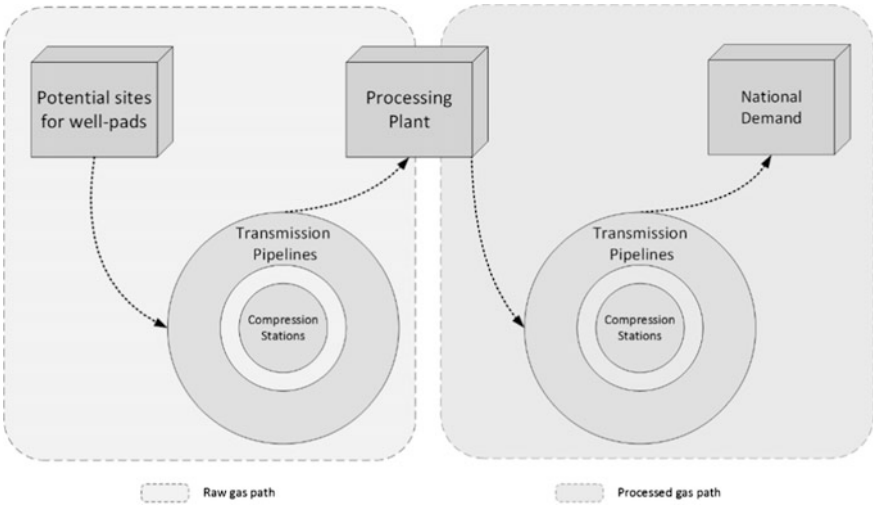

Fig. 1.1 Description of the strategic planning of the addressed problem (Arredondo-Ramirez et al. 2016)

includes the total sales in all markets over all the time periods of the considered time horizon, the total capital cost involved in the installation of the production network (TOTCAPPC) and the total operating cost associated with the network (TOTOPCOST).

The superstructure developed by Arredondo-Ramirez et al. (2016) has been used (see Fig. 1.2), and the optimization model is presented in the next section.

1.4 Optimization Model

The used indexes are the following: i represents the wells, p describes the multi-well-pads, m refers to the markets and r designates the processing plants/compression units. Furthermore, the index t is used to define the time periods over the given horizon. All the symbols used are described in the nomenclature section. Additional information on the optimization formulation are given by Arredondo-Ramirez et al. (2016).

1.4.1 Balance in Wells

$$f_{i,p,t}^{well} = \sum_{r} f_{i,p,r,t}, \forall\, i,p,t \tag{1}$$

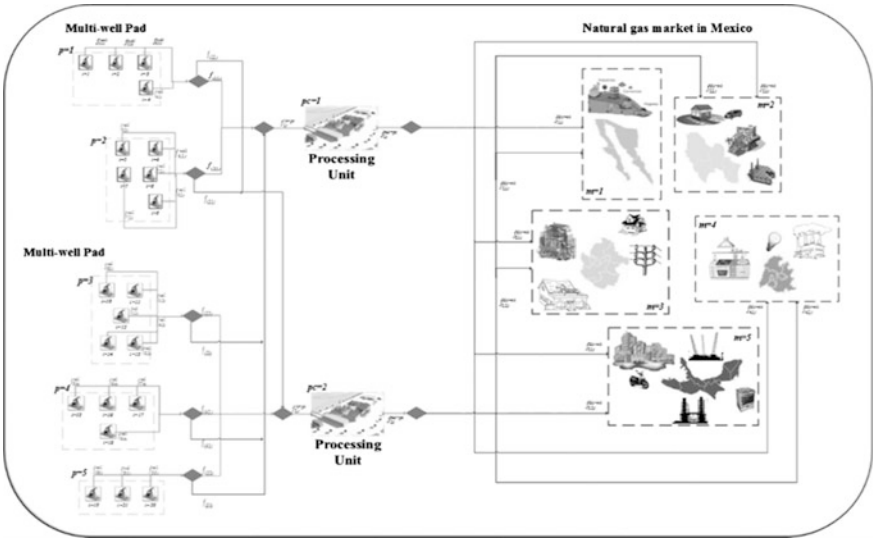

Fig. 1.2 Proposed superstructure for shale gas planning exploitation in Mexico (Arredondo-Ramirez et al. 2016)

1.4.2 Balance in the Inlet of Processing Plant/Compression Unit

$$f_{r,t}^{in-pc} = \sum_{i} \sum_{p} f_{i,p,r,t}, \forall\, r, t \tag{2}$$

1.4.3 Balance in the Outlet of Processing Plant

$$f_{r,t}^{out-pc} = \beta^{proc} \cdot f_{r,t}^{in-pc}, \forall\, r, t \tag{3}$$

1.4.4 Distribution of the Shale Gas to the Markets

$$f_{r,t}^{out-pc} = \sum_{m} f_{r,m,t}^{dis-mk}, \forall\, r, t \tag{4}$$

1.4.5 Total Gas in the Market

$$f_{m,t}^{mk} = \sum_{r} f_{r,m,t}^{dis-mk}, \forall\, m, t \tag{5}$$

1.4.6 Maximum Demand of Gas in the Market

$$f_{m,t}^{mk} \leq \mathrm{f}_{\mathrm{m,t}}^{\mathrm{de-mk}}, \forall\, m, t \tag{6}$$

1.4.7 Maximum Gas Produced in a Given Location

$$\sum_{t} \sum_{i} f_{i,p,t}^{well} \leq \mathrm{f}_{\mathrm{p}}^{\max}, \forall\, p \tag{7}$$

1.4.8 Maximum Number of Wells per Zone Over a Time Period

$$\sum_{i} y_{i,p,t}^{well} \leq \mathrm{N}_{\mathrm{p}}^{\mathrm{max-well}}, \forall p, t \tag{8}$$

1.4.9 Disjunctive Model to Determine the Existence of the Wells

$$f_{i,p,t}^{well} \leq \mathrm{f}_{\mathrm{p}}^{\max}, \forall\ i, p, t \tag{9}$$

$$f_{i,p}^{well-cap} \geq f_{i,p,t}^{well}, \forall\, i, p, t \tag{10}$$

To model the period of time here the well starts its operation, the following disjunction is developed:

$$\vee_{t} \begin{bmatrix} Y_{i,p,t}^{well} \\ f_{i,p,t'}^{well} \leq 0, \forall t' < t \\ f_{i,p,t'}^{well} \leq 0, \forall t' > t + t_{op} \\ CapCostWell_{i,p} = \mathrm{K_{Ft}} \mathrm{VP_t} [\mathrm{FC}_{\mathrm{i,p}}^{\mathrm{well}} + \mathrm{VC}_{\mathrm{i,p}}^{\mathrm{well}} \cdot (f_{i,p}^{well-cap})] \end{bmatrix}, \forall i, p$$

And previous disjunction is algebraically reformulated as follows:

$$\sum_{t} y_{i,p,t}^{well} = 1, \forall i, p \tag{11}$$

$$f_{i,p,t'}^{well} \leq \mathrm{f}_{\mathrm{p}}^{\max} \cdot (1 - y_{i,p,t}^{well}), i, p, t, \forall\, t' < t \tag{12}$$

$$f_{i,p,t'}^{well} \leq \mathrm{f}_{\mathrm{p}}^{\max} \cdot (1 - y_{i,p,t}^{well}), \forall\, i, p, t, \forall\, t' > t + t_{op} \tag{13}$$

$$\begin{aligned} CapCostWell_{i,p} \geq\ & \mathrm{K_{Ft}VP_t} \cdot [\mathrm{FC_{i,p}^{well}} + \mathrm{VC_{i,p}^{well}} \cdot (f_{i,p}^{well-cap})] \\ & - \mathrm{CapCostWell_{i,p}^{max}}(1 - y_{i,p,t}^{well}), \forall i, p, t \end{aligned} \tag{14}$$

1.4.10 Disjunctive Model to Determine the Existence of the Processing Unit

$$f_{r,t}^{in-pc} \leq \mathrm{f}_{r}^{\max-\mathrm{pc}}, \forall\, r, t \tag{15}$$

$$f_{r}^{in-pc-cap} \geq f_{r,t}^{in-pc}, \forall\, r, t \tag{16}$$

Similarly, the following disjunction is also needed for determining the time period when a processing unit needs to be installed:

$$\vee_{t} \begin{bmatrix} Y_{r,t}^{pc} \\ f_{r,t'}^{in-pc} \leq 0, \forall\, t' \leq t \\ CapCostPC_r = \mathrm{K_{Ft}VP_t}[\mathrm{FC_r^{PC}} + \mathrm{VC_r^{PC}} \cdot (f_r^{in-pc-cap})] \end{bmatrix}, \forall\, r$$

Which is reformulated as follows:

$$\sum_{t} y_{r,t}^{PC} = 1, \forall r \tag{17}$$

$$f_{r,t'}^{in-pc} \leq \mathrm{f}_{\mathrm{r,p}}^{\max-\mathrm{pc}} \cdot (1 - y_{r,t}^{PC}), \forall r, t, \forall t' < t \tag{18}$$

$$CapCostPC_r \geq \mathrm{K_{Ft}VP_t} \times [\mathrm{FC_r^{PC}} + \mathrm{VC_r^{PC}} \cdot (f_r^{in-pc-cap})] - \mathrm{CapCostPC_r^{max}}(1 - y_{r,t}^{PC}), \forall\ r, t \tag{19}$$

1.4.11 Disjunctive Model to Determine the Existence of the Distribution Network

1.4.11.1 Shale Gas Distribution from Well-Pad to Processing Units

$$f_{i,p,r,t} \leq \mathrm{f}_{\mathrm{i,p,r}}^{\mathrm{max-pc}}, \forall\, i,p,r,t \tag{20}$$

$$f_{i,p,r}^{pc-pipe-cap} \geq f_{i,p,r,t}, \forall\, i,p,r,t \tag{21}$$

Also, the following disjunction is needed:

$$\vee_t \begin{bmatrix} Y_{i,p,t}^{pipe-well-pc} \\ f_{i,p,r,t'} \leq 0, \forall\, t' \leq t \\ CapCostPipeWellPC_{i,p,r} = \mathrm{K_{Ft}VP_t[FCPipeWellPC_{i,p,r}} + \mathrm{VCPipeWellPC_{i,p,r}} \cdot (f_{i,p}^{pc-pipe-cap})] \end{bmatrix}, \forall\, i,p,r$$

Which is reformulated as follows:

$$\sum_t y_{i,p,r,t}^{pipe-well-pc} = 1, \forall i,p,r \tag{22}$$

$$f_{i,p,r,t'} \leq \mathrm{f}_{\mathrm{i,p,r}}^{\mathrm{max-pc}} \cdot (1 - y_{i,p,r,t}^{pipe-well-pc}), \forall i,p,r,t, \forall t' < t \tag{23}$$

$$\begin{aligned} CapCostPipeWellPC_{i,p,r} \geq\ & \mathrm{K_{Ft}VP_t[FCPipeWellPC_{i,p,r}} + \mathrm{VCPipeWellPC_{i,p,r}} \cdot (f_{i,p}^{pc-pipe-cap})] \\ & - \mathrm{CapCostPipeWellPC}_{\mathrm{i,p,r}}^{\max}(1 - y_{i,p,r,t}^{pipe-well-pc}), \forall i,p,r,t \end{aligned} \tag{24}$$

1.4.11.2 Natural Gas Distribution from Processing Units to the Market

$$f_{r,m,t}^{dis-mk} \leq \mathrm{f}_{\mathrm{r,m}}^{\mathrm{max-dis-mk}}, \forall\, r,m,t \tag{25}$$

$$f_{r,m}^{dis-mk-pipe-cap} \geq f_{r,m,t}^{dis-mk}, \forall\, r,m,t \tag{26}$$

And the following disjunction is used:

$$\vee_t \begin{bmatrix} Y_{r,m,t}^{pipe-pc-mk} \\ f_{r,m,t'}^{dis-mk} \leq 0, \forall\, t' \leq t \\ CapCostPipeDisMk_{r,m} = \mathrm{K_{Ft}VP_t[FCPipeDisMk_{r,m}} + \mathrm{VCPipeDisMk_{r,m}} \cdot (f_{r,m}^{dis-mk-pipe-cap})] \end{bmatrix}, \forall\, r,m$$

Which is reformulated as follows:

$$\sum_{t} y_{r,m,t}^{pipe-pc-mk} = 1, \forall\, r, m \tag{27}$$

$$f_{r,m,t'}^{dis-mk} \leq \mathrm{f}_{\mathrm{r,m}}^{\max-\mathrm{dis}-\mathrm{mk}} \cdot (1 - y_{r,m,t}^{pipe-pc-mk}), \forall\, r, m, t, \forall\, t' < t \tag{28}$$

$$\begin{aligned} CapCostPipeDisMk_{r,m} \leq\ & \mathrm{K_{Ft}VP_t}[\mathrm{FCPipeDisMk_{r,m}} + \mathrm{VCPipeDisMk_{r,m}} \cdot (f_{r,m}^{dis-mk-pipe-cap})] \\ & - \mathrm{CapCostPipeDisMk_{r,m}^{max}}(1 - y_{r,m,t}^{pipe-pc-mk}), \forall r, m, t \end{aligned} \tag{29}$$

1.4.12 Total Cost

First the total capital cost is:

$$\begin{aligned} TotCAPC = & \sum_{i}\sum_{p} \mathrm{CapCostWell_{i,p}} + \sum_{r} \mathrm{CapCostCP_r} \\ & + \sum_{i}\sum_{p}\sum_{r} \mathrm{CapCostPipeWellCP_{i,p,r}} \\ & + \sum_{r}\sum_{m} \mathrm{CapCostPipeDisMk_{r,m}} \end{aligned} \tag{30}$$

And the operating cost:

$$\begin{aligned} TotOpCost = & \sum_{i}\sum_{p}\sum_{t} \mathrm{K_{0_t}} \cdot \mathrm{VC_{i,p}^{well}} \cdot f_{i,p,t}^{well} + \sum_{r}\sum_{t} \mathrm{K_{0_t}} \cdot \mathrm{VC_r^{pc}} \cdot \mathrm{f_{r,t}^{in-pc}} \\ & + \sum_{i}\sum_{p}\sum_{r}\sum_{t} \mathrm{K_{0_t}} \cdot \mathrm{VC_{i,r,p}^{pump-well-pc}} \cdot f_{i,p,r,t}^{pc} \\ & + \sum_{r}\sum_{m}\sum_{t} \mathrm{K_{0_t}} \cdot \mathrm{VC_{r,m}^{pump-pc-mk}} \cdot f_{r,m,t}^{dis-mk} \end{aligned} \tag{31}$$

1.4.13 Total Sales at Present Value

$$Sales = \sum_{m}\sum_{t} K_{0_t} \cdot VC^{gas} \cdot f_{m,t}^{mk} \tag{32}$$

Fig. 1.3 Coastal shale formation in Mexico

1.4.14 Objective Function

The net present profit is given as follows:

$$NPVprofit = Sales - TotCAPC - TotOpCost \tag{33}$$

1.5 Case Study

A case study from Mexico was considered (see Fig. 1.3), where the main data were taken from Lozano-Maya (2013) (see Figs. 1.4 and 1.5).

Two time horizons were considered for the case study: five and ten years of production. The production was assumed to supply the total demand of natural gas in Mexico from 2015−2019 for the first scenario and from 2015−2024 for the second one. The data are summarized in Table 1.1. The total demand in the country was obtained from the Mexican government (Sener 2013), which offers forecast in the different markets in Mexico that are divided in five zones.

1.6 Results Analysis

A five-year horizon was considered, for which the optimization formulation consists of 760 continuous variables and 300 binary variables and includes 1,891 constraints. The proposed model was coded in the software GAMS (Brooke et al. 2016), and the CPU time consumed was 7,600 s.

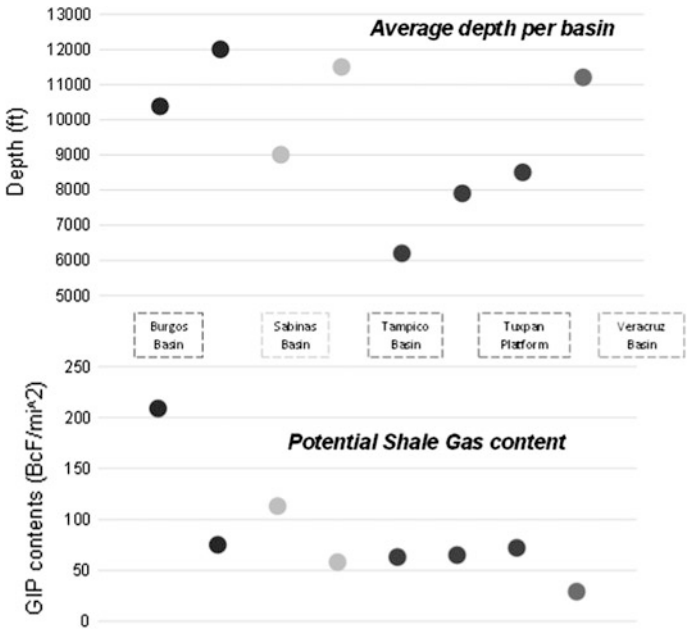

Fig. 1.4 Preliminary analysis of shale formations in Mexico

The optimal solution is shown in Fig. 1.6. The optimal planning includes drilling six wells (i1, i2, i4, i9, i10 and i16). The existence of the well is shown and it presents the year (t1, t2, t3, t4, t5) when each well stars its operation. The non-existing wells are faded in the figure. The operation of two processing plants (r1, r2) is considered in the first year (t1) of production. It is important to discuss the total rate processed in each plant. Table 1.2 shows greater activity in Processing Plant 1, which produces 75 % of total natural gas. Figure 1.6 shows the existence of the pipelines over the different time periods with different colors used for each year. Figure 1.6 also presents the piping network for the raw gas from the wells to the processing plants where two pipelines are used in the first year (t1) from two wells (i1, i16) to the corresponding processing plants (r1, r2). For the second year (t2), the same well (i10) requires two different pipelines to the processing plants (r1, r2). Towards the third and fourth years (t3, t4), well number two (i2) and well number nine (i9) are activated and send the raw gas to the processing plant one (r1) and plant two (r2). Finally, in the fifth year (t5), well number four (i4) is required to transport raw gas to processing plant one (r1). Another important element shown in Fig. 1.6, is the pipeline network from the processing plants to markets. Table 1.3 shows information about the pipeline network. It should be noted that while the construction of well ten (i10) is in the second year (t2), it starts to send raw gas to the processing plant 1 (r1) during the third year (t3). Figure 1.7 shows that processing plant one (r1) distributes to the five markets (m1, m2, m3, m4, m5) in the first year (t1) of operation. On the other hand, processing plant two (r2) distributes

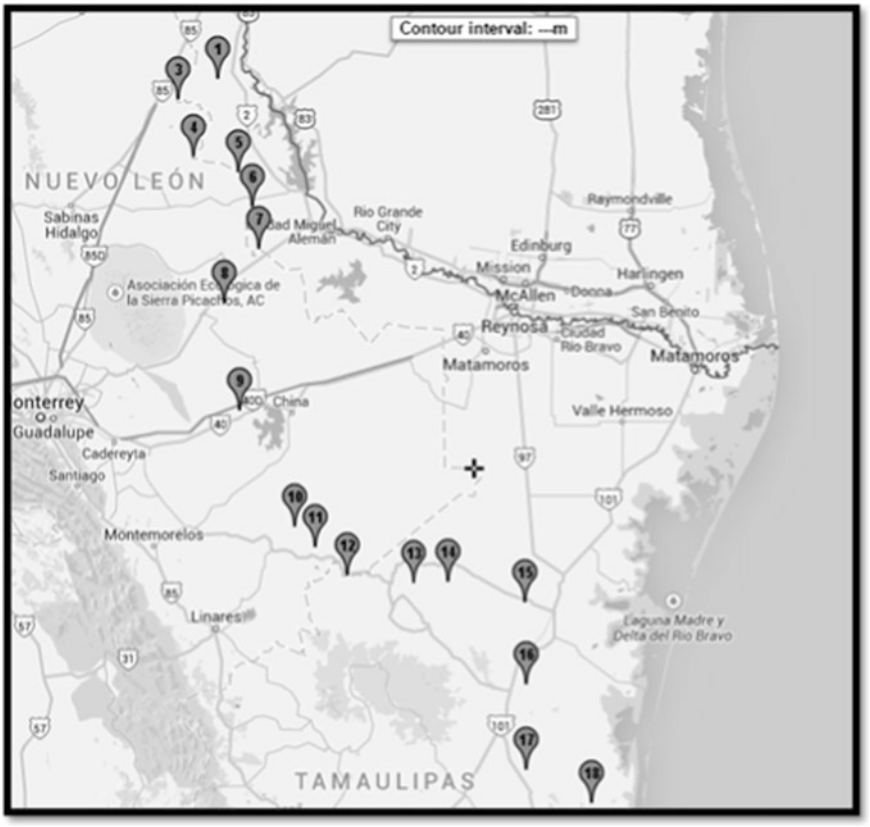

Fig. 1.5 Well localization in the Burgos basin

Table 1.1 Natural gas demanded per market region in Mexico (Sener 2013)

Year	Central zone (MMSCF)	Midwestern (MMSCF)	Northwest (MMSCF)	Northeastern (MMSCF)	South-Southeast (MMSCF)
2015	2.7	3.0	1.7	7.1	12.4
2016	2.8	3.0	2.1	7.4	12.5
2017	3.1	3.2	2.5	7.6	12.8
2018	3.4	3.5	2.7	7.5	13.0
2019	3.5	3.9	2.7	7.9	12.4
2020	3.7	4.2	2.8	8.1	11.7
2021	4.0	4.6	3.0	8.2	11.5
2022	4.1	4.8	3.0	8.2	12.0
2023	4.1	4.9	3.1	8.4	12.7
2024	4.3	5.0	3.3	8.5	13.1

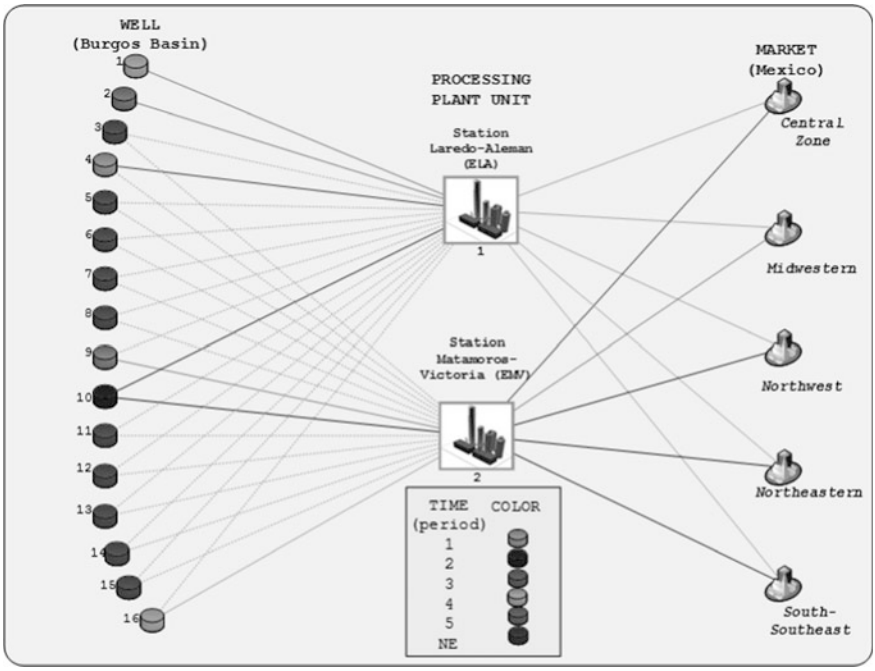

Fig. 1.6 Optimal solution for the case study

Table 1.2 Total gas processed in plants for the case study

Period	Processing plant 1 (MMSCF)	Processing plant 2 (MMSCF)
1	36.2	17.6
2	36.6	19
3	37.6	20.8
4	53.2	7
5	53	7.8

Table 1.3 Construction period for the pipes from well-pads to processing plants for the case study

Well	Processing plant 1 (period)	Processing plant 2 (period)
Well 1-Period 1	1	–
Well 2-Period 3	3	–
Well 4-Period 5	5	–
Well 9-Period 4	–	–
Well 10-Period 2	3	1
Well 16-Period 1	–	1

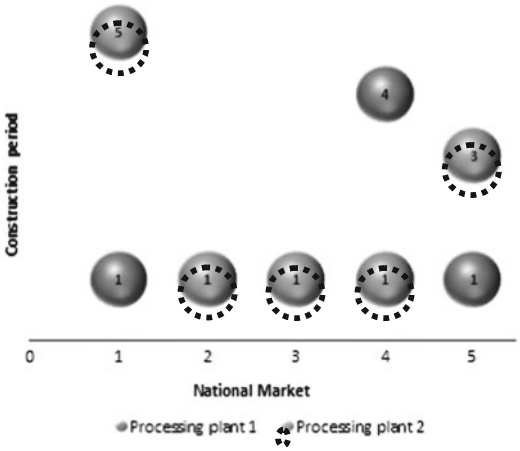

Fig. 1.7 Piping construction for distributing the gas to the markets

Table 1.4 Construction period for the pipes from the processing plants to the markets

	Central zone	Midwestern	Northwest	Northeastern	South-Southeast
Processing plant 1-Period 1	1	1	NE	4	1
Processing plant 2-Period 1	5	1	1	1	3

to the Central (m1), Northeastern (m4) and South-Southeast (m5) markets in the second year, and it distributes to the Midwestern (m2) market during the third year (t3), as well as to the Northwest (m3) market during the fifth year (t5). Table 1.4 shows that the construction time for the distribution network varies for every plant. For instance, processing plant 2 (r2) supplies every market (m2, m3, m4) in the first year, with the exception of the Central Zone (m1) which is supplied in the fifth year (t5) and the South-Southeast (m5) in the third year (t3).

Figure 1.8 shows that the annual production ranges between 54 and 61 Mscf. It should also be noted that there is a requirement to drill one well each year with the exception of the fifth year when it there is needed to operate wells one and sixteen (i1, i16). Figure 1.9 shows how the gas demands are satisfied through the different time periods, which sum a total amount of 144.4 Mscf. The South-Southeast market represents the highest consumer because it is the most industrialized zone in Mexico. Finally, the net present profit is 512,890 Billions of USD with a capital cost of 335 Billions of USD and an operating cost of 1,060 Billions of USD. The production in the six wells were 288.8 Mscf as shown in Table 1.5.

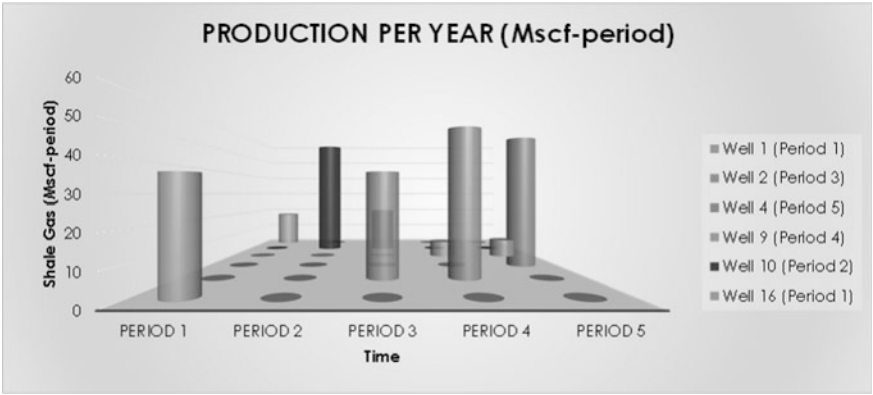

Fig. 1.8 Production per year in the Burgos basin

Fig. 1.9 Sales to markets per year in the Burgos basin

Table 1.5 Summary for the economic results at present value

	Scenario 1
Total capital cost (Billions USD)	335.5
Total operating cost (Billions USD)	1059.9
Sales (Billions USD)	514290.0
Net present profit (Billions USD)	512890.0
Total produced gas (MMSCF)	288.8

1.7 Conclusions

This chapter has presented a multi-period optimization approach for planning and developing an infrastructure for shale gas production, treatment, and distribution. The optimization model is based on a superstructure that accounts for the installation of infrastructure for exploiting new wells, new processing facilities, the piping network for the raw gas to treatment, and the distribution network for the treated gas to the markets. The objective is to maximize the net present profit of the infrastructure. The model determines the optimal size and operation for the new infrastructure as well as the construction time for satisfying the demands while accounting for the time value of the money.

As case study from Mexico has been considered for the optimal planning of shale gas production, treatment, and distribution; where there is currently little infrastructure to support shale gas production and distribution. The results show that it is possible to satisfy the total national demands in while making attractive profits. The generated results highlight the importance of production and infrastructure planning and scheduling while accounting for the market demands and the time value of money. Finally, the proposed optimization formulation is general and can be applied for the optimal planning in different regions that have shale gas with no production and distribution infrastructure.

References

Arredondo-Ramírez, K., Ponce-Ortega, J. M., & El-Halwagi, M. M. (2016). Optimal planning and infrastructure development or shale gas production. *Energy Conversion and Management, 119*, 91–100.

BP. (2015). BP Energy Outlook 2035, BP.

Brooke, A., Kendrick, D., Meeruas, A., & Raman, R. (2016). *GAMS: A users guide*. Washington, DC: GAMS Development Corporation.

Cafaro, D. C., & Grossmann, I. E. (2014). Strategic planning, design, and development of the shale gas supply chain network. *AIChE Journal, 60*, 2122–2142.

He, C., & You, F. (2015). Process design and optimization of an integrated shale gas process for green chemicals production. *Computed Aided Chemical Engineering., 37*, 1397–1402.

Knudsen, B. R., Grossmann, I. E., Foss, B., & Conn, A. R. (2014). Lagrangian relaxation based decomposition for well scheduling in shale-gas systems. *Computers & Chemical Engineering, 64*, 234–249.

Lozano-Maya, J. R. (2013). The United States experience as a reference of success for shale gas development: The case of Mexico. *Energy Policy., 62*, 70–78.

OECD/IEA (2015). Key World Statistics 2015, IEA.

Rahman, M. M., Rahman, M. K., & Rahman, S. S. (2001). An integrated model for multiobjective design optimization of hydraulic fracturing. *Journal of Petroleum Science and Engineering, 31*, 41–62.

Sener. (2013). *Prospectiva del Mercado de Gas Natural 2013-2027*. Mexico: Mexico City.

Chapter 2
Shale Gas Process and Supply Chain Optimization

Jiyao Gao, Chang He and Fengqi You

Abstract In recent decades, large-scale production of shale gas has been considered as a major issue in the U.S. energy industry. In accordance with its great economic potential and environmental concerns, shale gas process and supply chain optimization has become one of the most popular research areas. In this chapter, we provide a comprehensive overview of the supply chain management and process design problems in shale gas industry. We summarize four major research challenge areas, namely the design and planning of shale gas supply chain, water management in hydraulic fracturing, sustainability concerns in shale gas industry, and design and optimization in shale gas processing system. We further provide review and discussions of the major publications corresponding to each of the aforementioned topics. Potential opportunities in the shale gas system are presented as well to illuminate the future research.

2.1 Introduction

Shale gas is known as unconventional natural gas extracted from shale rock layer and has emerged as one of the most promising energy sources within the last few decades. With the discovery of huge shale gas reserves all over the world, a "shale revolution" starts in the U.S. and keeps spreading out in other countries. In 2005, the U.S. barely produces any natural gas from shale formations. Nowadays, nearly 44 % of the total natural gas withdrawal is from shale gas wells (EIA 2011).

J. Gao · F. You (✉)
Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, 318 Olin Hall, Ithaca, NY 14853, USA
e-mail: fengqi.you@cornell.edu

C. He
School of Chemical Engineering and Technology, Sun Yat-Sen University, Zhuhai 519082, People's Republic of China

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_2

According to the Annual Energy Outlook 2015 by the U.S. Energy Information Administration (EIA 2015), natural gas production is expected to grow by an average rate of 1.6 % per year from 2012 to 2040. As a result, the percentage of the U.S. total natural gas production from shale gas is expected to increase to 53 % by 2040.

The remarkable development of large-scale shale gas production would not be possible without the hydraulic fracturing and horizontal drilling technologies (Gregory et al. 2011; Vengosh et al. 2013). Different from the conventional natural gas, shale gas is embedded in the shale rocks that can be a few thousand feet deep. Therefore, special techniques are required to withdraw this unconventional resource. By using hydraulic fracturing, millions of gallons of fracturing fluid (about 90 % water, 9.5 % sand, and 0.5 % chemical additives) is pumped underground under high pressure (up to 70 Mpa), fracturing the rock layer and holding fractures open, thus forcing the shale gas and/or oil to flow back to the surface. The horizontal drilling, on the other hand, is a drilling process in which the well is turned horizontally at depth. Compared with the vertical drilling, horizontal drilling allows us to drill multiple wells at a single shale site/pad. As a result, a horizontal well site is able to produce more energy with fewer wellbores, which significantly reduces the capital investment and improves the efficiency for shale gas production (Hughes 2013).

Shale gas production of a single well normally features a high initial production rate followed by an astounding decline from 60 to 90 % after the first three years (Hughes 2014). This characteristic is mainly caused by the pressure depletion and inherently low permeability of the reservoir. Consequently, operators need to regularly drill new wells to maintain a stable production profile, which results in a scheduling problem. Additionally, based on the composition, shale gas can be classified as dry gas and wet gas, and the key difference is the content of natural gas liquids (NGLs). The NGLs are defined as light hydrocarbons including ethane, propane, butane and heavier components, which typically have substantially higher market values than methane gas. Dry gas is almost pure methane with trace NGLs. Although methane is still the dominant component in wet gas, the amount of NGLs is significant enough to require further processing. Depending on the location, both the estimated ultimate recovery (EUR) and shale gas composition of a shale well may have significant variance. All of these issues render the optimal design and operation in the shale gas industry a challenging topic.

Apart from the shale gas itself, the wide application of hydraulic fracturing has resulted in another serious issue, known as the shale water management problem (Nicot and Scanlon 2012; Mauter et al. 2013, 2014; Small et al. 2014; Yang et al. 2014; Mauter and Palmer 2014). The hydraulic fracturing operation during the shale gas production process requires a massive amount of freshwater (Jiang et al. 2014). Till 2012, a total of 63,000 shale wells have been reported in the U.S., and each well requires approximately 4–6 million gallons of water for fracturing and production (EPA 2011; API 2010; Paper 2008; Nicot et al. 2014). Such a huge

consumption of fresh water resource can cause severe consequences on local water supply especially in water-scarce territories. Moreover, during hydraulic fracturing process, a large portion of injected fluid flows back to the surface as highly contaminated wastewater, which contains high concentration of total dissolved solids (TDS) as well as other toxic and radioactive dissolved constituents (Soeder and Kappel 2009; Rahm and Riha 2012). Due to the increasing public concern on the water related environmental issues, it is imperative to develop an effective approach to address challenges in the shale water management problem.

The shale gas produced at shale site is generally considered as raw shale gas that needs further processing. The processing service is typically provided by midstream processors. Through the shale gas processing, impurities such as compounds and gases, oil, and water mixed with the natural gas are removed. Two major products, "pipeline-quality" sales gas and NGLs, are extracted and sold separately. The sales gas is normally delivered to major intrastate and interstate pipeline transmission systems and further sent to final customers. The NGLs, on the other hand, can be used as feedstock for the production of value-added chemicals, such as olefins and gasoline blending stocks. With the rapid development of shale gas industry, excessive supply of NGLs requires more cost-effective shale gas processing designs and conversion alternatives for a better use of this valuable byproduct (He and You 2014; Ehlinger et al. 2014).

Despite the great economic potential stimulated by the shale gas, one major concern impeding the expansion of shale gas industry is its negative impact on the climate change. Methane is about 25 times more potent greenhouse gas (GHG) than carbon dioxide based on the 100-year global warming potential (GWP). A small amount of methane leakage could lead to enormous greenhouse gas footprint (Howarth 2014). Additionally, supply chain activities such as shale gas production, processing, transportation, and gas-based power generation could incur large amount of GHG emissions as well (Allen 2014a, b; Zavala-Araiza et al. 2015). There have been extensive studies published evaluating the life cycle carbon footprint of shale gas (Jiang et al. 2014; Harto 2013; Dale et al. 2013; Burnham et al. 2011; Stephenson et al. 2011; DOE/NETL 2011; Laurenzi and Jersey 2013; Alvarez et al. 2012; Brandt et al. 2014). However, the shortage of decision-support tools and methodologies still exists, which requires the development of corresponding optimization models for more sustainable design alternatives in the shale gas industry.

This chapter reviews the most recent advances in application of mathematical programming techniques for optimization of shale gas process and supply chain designs. The remaining content of this chapter is organized as follows. We first present a comprehensive review of recent literature on strategic design and planning of shale gas supply chain. Next, we introduce the water management problem in the shale gas system and explore the application of life cycle optimization approach in the shale gas supply chain. After that, we focus on the shale gas processing and conversion alternatives from a process design perspective. Finally, we present challenges and opportunities for future shale gas development and conclude the chapter.

2.2 Design and Planning of Shale Gas Supply Chain and Water Management

2.2.1 Shale Gas Supply Chain Network

The shale gas supply chain is a complex system involving shale gas suppliers in the upstream, shale gas processors in the midstream, and distributors in the downstream (Chima 2011; Seydor et al. 2012). In this section, we provide a comprehensive overview of a typical shale gas supply chain structure and corresponding design decisions. The major stages of a shale gas supply chain are summarized in the following Fig. 2.1.

2.2.1.1 Shale Site Construction

First, potential shale wells are explored through geologic evaluation. Once a potential shale site is identified, the well operator needs to reach a lease agreement with the corresponding landowner and then obtain the drilling permits. It is the operator's responsibility to guarantee that all the following drilling and production activities will be carried out in accordance with relevant regulations. After the approval of the operator's permit by local environmental regulation agencies, the site construction and well drilling can start.

The shale site is typically constructed following these steps: the first step involves the clearance of proposed area and the accommodation of equipment. Meanwhile, a road way is constructed to provide access to the shale site. Subsequently, pits/impoundments are constructed to properly handle the fluids during drilling and hydraulic fracturing. Next, pipelines associated with shale sites are installed, including gathering lines, injection lines, and water supply lines. Other infrastructures such as storage tanks are built as well.

2.2.1.2 Drilling, Fracturing, and Production

After the shale site construction, the drilling rig is moved on site and assembled. A conductor hole is predrilled, and then conductor pipes are inserted to prevent soft rocks from caving and conduct drilling mud from bottom to the surface during drilling process. Depending on the number, depth, and length of horizontal wells to

Fig. 2.1 Overview of shale gas supply chain

be drilled, the drilling stage can last for a few months, which requires a constant supply of drilling fluid and proper handling of sediments and wastewater. Once the drilling is completed, protective casing and cementing are used.

The following stage is the well completion, which mainly involves the hydraulic fracturing operation. A mixture of water, sand, and chemical additives is injected underground at a high pressure to break up shale-rock formations, such that fractures are created and held open by proppant, and then shale gas and oil can be extracted. Typically, the horizontal wells are stimulated by stages, depending the specific fracturing schedule and technology applied, the hydraulic fracturing stage could last for several months. Once fracturing is completed, a wellhead is constructed, and the local gathering pipelines are prepared for the controlled extraction of natural gas.

2.2.1.3 Transport and Processing

The transportation of shale gas is mainly carried out by pipelines. The pipeline system includes all the equipment such as pumps, valves, meters, and monitoring devices. After shale gas is produced, water and condensate are typically removed at or near wellhead. Then, gathering lines (6–20-in. diameter) are used to take the raw shale gas to processing facilities, where sales gas that meets the pipeline specifications and valuable NGL products are obtained. Gathering lines are typically considered as the demarcation between upstream production and midstream processing and transmission to market.

The shale gas processing plant is a dedicated separation train that consists of a series of processes, including gas sweetening, dehydration, NGL recovery, and N_2 rejection. If economically feasible, NGLs may be further separated into ethane, propane, butanes, and pentanes+streams in a NGLs fractionation train. Notably, for dry gas that is almost pure methane, processing might be unnecessary (Cafaro and Grossmann 2014). As a result of the rapid development of shale gas industry, expansion of infrastructure including pipelines and processing plants is getting necessary, and the high capital investment of these infrastructures requires an optimal strategy for the corresponding design decisions.

2.2.1.4 Storage and Distribution

Like other commodities, the produced shale gas can either be directly transported to interstate/intrastate pipeline system for distribution or transported to underground reservoirs and stored for an indefinite period of time. There are three principle types of underground storage sites in the U.S. today, including the depleted natural gas or oil fields, aquifers, and salt caverns (EIA Underground Natural Gas Storage 2014). The underground reservoirs can behave as a “buffer” in the shale gas supply chain to accommodate fluctuations in natural gas demand and price.

The distribution system is in charge of delivering the shale gas to end-use customers. While large customers such as power plants may receive gas directly from the interstate or intrastate pipelines, most individual customers buy gas from distribution companies. The transmission cost could account for up to half of the total distribution cost, in addition to the commodity cost. Depending on the specific function, the transmission pipeline can be 6–48 in. in diameter. As a result, the corresponding capital investment of these pipelines can be significantly different. Thus, it is important to design proper capacity for the distribution pipeline system.

2.2.1.5 End Use

Natural gas is priced and traded as a commodity at different locations throughout the country. These locations are known as market hubs that are normally located at the intersection of major pipeline systems. The biggest market hub is Henry Hub located in Louisiana, and the spot and future natural gas prices set at Henry Hub are generally considered as the primary price set for the North American natural gas market. Depending on the specific usage, the natural gas consumption can be classified into three major partitions, including lease and plant fuel, pipeline and distribution use, and volumes delivered to consumers. There are four types of end-customers, namely power plants, industrial customers, commercial customers, and residential customers (EIA Natural Gas Consumption by End Use 2015).

2.2.2 Optimization Models for Shale Gas Supply Chain

Conventional natural gas supply chain has been fully studied in the literature. A variety of models have been proposed addressing the design and planning problems in oil/gas supply chains (Duran and Grossmann 1986; Iyer et al. 1998; van den Heever and Grossmann 2000; Gupta and Grossmann 2012). In recent decades, with the rapid development of shale gas industry, many research studies arise specifically focusing on the optimization of shale gas system. Knudsen and Foss (2013) present a novel operational scheme for enhanced utilization of late-life shale gas systems. In this work, a large number of geographically distributed wells and pads are considered, which are producing at low erratic rates due to reservoir pressure depletion and well liquid loading. By using a shale-gas well and reservoir proxy model, a generalized disjunctive program (GDP) is formulated. The proposed cyclic shut-in and production strategy is expected to avoid well liquid loading and improve the overall production. Following this work, Knudsen et al. (2014) propose a Lagrangian relaxation-based decomposition scheme for solving large field-wide well scheduling problems in shale gas systems. Furthermore, they explore the integration of shale gas supply with local natural-gas power generation. A large-scale mixed-integer linear program (MILP) is proposed, and the results

indicate a potential increase of profit for shale gas operators by improving well scheduling (Knudsen et al. 2014).

The first mathematical model addressing the long-term strategic planning and design of the shale gas supply chain is proposed by Cafaro and Grossmann (2014). The authors propose a large-scale nonconvex mixed-integer nonlinear program (MINLP) model to determine the optimal design of shale gas supply chain. Critical decisions, including the drilling and fracturing plan, the location and sizing of processing plants, the length and location of gas and liquid pipelines, and the power of gas compressors, are simultaneously addressed. The overall objective is defined as maximizing the net present value (NPV) of the project over a 10-year planning horizon. Features such as decreasing production rate of shale wells, variant shale gas composition, and concave cost functions for capital investment are captured. A real-world case study based on Marcellus shale play is considered and solved by a tailored branch-and-refine-optimization approach. From the results of this work, the importance of NGLs to the economics of the project is identified especially considering the low natural gas price. In addition, the optimal drilling and fracturing scheduling could maximize the utilization of processing and transportation infrastructure, thus significantly improving the overall economics and energy/resource efficiency.

In spite of the optimistic forecast of shale gas development, a recent report unveils the fact that the actual profitability of a shale well can be significantly affected by various uncertainties, especially uncertainty of EUR (Hughes 2014). There are some studies evaluating the influence of uncertainty in shale gas supply chain based on general analysis and numerical models (Gracceva and Zeniewski 2013; Jayakumar and Rai 2012; Chaudhri 2012; Harding 2008). However, few of them provide a systematic approach to tackle the shale gas supply chain optimization problem under uncertainty.

Identifying such a knowledge gap in the shale gas supply chain literature, Gao and You (2015) present a comprehensive shale gas supply chain model that not only considers design and planning decisions, but also properly addresses EUR uncertainty of shale wells. In this work, a shale gas supply chain covering the upstream producers to the downstream end-customers is considered, as shown in Fig. 2.2. The EUR distribution is derived from real data reported in literature (EIA 2011; Swindell 2014). By reviewing the distribution of EUR, a distinct "long tail" is observed. Thus, a scenario-based stochastic programming approach is adopted as the suitable method accounting for EUR uncertainty. The resulting problem is formulated as a two-stage stochastic mixed-integer linear fraction programming (SMILFP) model. The objective is to minimize the expected levelized cost of energy (LCOE) under all the realization of scenarios, which is normally regarded as the cost at which energy must be generated to break-even over the life time of the project. The stochastic program is known for its computationally challenging property, because its problem size scales up exponentially as the number of scenarios increases. Therefore, a sample average approximation approach is applied to generate scenarios based on real EUR distribution data. Additionally, a novel algorithm integrating the parametric algorithm and the L-shaped method is

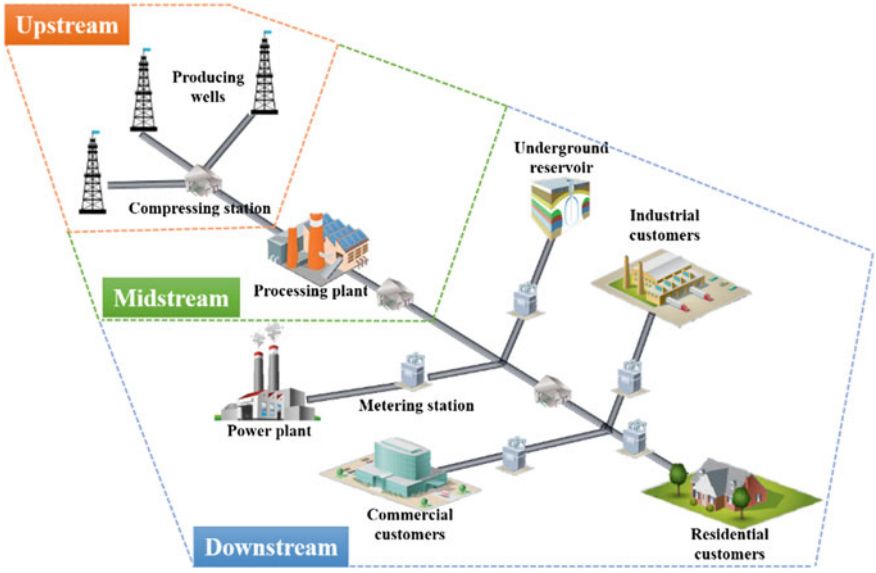

Fig. 2.2 Shale gas supply chain network from shale well to end-user (Gao and You 2015)

developed to solve the resulting large-scale SMILFP problem. The proposed modeling framework and solution method are demonstrated by a case study based on the Marcellus shale play. By solving the corresponding optimization problem, an LCOE of \$0.0038/MJ is obtained. Moreover, by comparing the result with that of the perfect information model (all the EUR data are known beforehand), the LCOE is only 3 % less in the stochastic model. Meanwhile compared with the deterministic model (using nominal EUR data to determine design decisions), the LCOE obtained from the stochastic model is improved by 30 %. This result verifies the great influence of EUR uncertainty on the shale gas supply chain optimization, and the stochastic programming model is proven a superior choice for determining optimal shale gas supply chain design under EUR uncertainty.

Recently, Drouven and Grossmann (2016) propose multi-period planning, design and strategic models addressing the long-term quality-sensitive shale gas development problem. As important extensions of the previous work by Cafaro and Grossmann (2014), some new developments are presented. First, a novel superstructure for shale gas development problem is proposed, which captures the distinctive "tree"-structure of typical shale gas gathering system. Different delivery options including processing sales routes and direct delivery sales arcs are explicitly distinguished. In addition, discrete sizes of pipeline diameters and compressors are considered. Thus, the corresponding capacity constraints are captured by mixed-integer linear constraints, and the economies of scale is taken into account

without involving concave cost functions. Moreover, this work extends the scope of shale gas development problem to include strategic decisions, such as the selection of delivery nodes, arrangement of delivery agreements, and procurement of delivery capacity. Most importantly, the spatial composition variations of shale gas are explicitly addressed in this work. The quality of shale gas is required to satisfy the delivery specifications at delivery nodes. Besides, depending on the price forecast, the upstream operator will target different shale gas qualities. The resulting problem is a large-scale, nonconvex, MINLP problem with an objective to maximize the NPV over the planning horizon. Based on the results, the shale gas development is proven to be quality sensitive. Additionally, the profitability of shale gas development projects can be improved by a few million U.S. dollars through the optimization of return-to-pad operations, equipment utilization, and strategic delivery agreements.

2.3 Sustainable Design and Operations of Shale Gas Supply Chain

Although the shale gas industry is well known for its great economic potential, another aspect that needs at least equal attention is its overall sustainability performance. There are two major topics focusing on the environmental impacts of shale gas development, namely the water management issue and the life cycle environmental impacts of shale gas. In this section, we provide a comprehensive description of the related problems and give a selected review of the recent optimization literature.

2.3.1 Water Management in Shale Gas System

Water use is associated with each step of drilling and shale gas production process. It is known that the hydraulic fracturing operation requires millions of gallons of freshwater (Jiang et al. 2014). In the Eagle Ford play of south Texas, for instance, large consumption of water resources in hydraulic fracturing would make the droughts even worse (Mauter et al. 2014). In other regions, such as Marcellus where water scarcity is not a severe problem, spatial and seasonal variability in stream flow rates still raises the risk that water withdrawals may negatively impact water resources (Mauter and Palmer 2014). Furthermore, as fracturing fluid is injected underground, a portion of water will flow back to the surface as highly contaminated water. The management of the flowback water and produced water is recognized as a greater challenge.

2.3.1.1 Overview of Water Management Options

The wastewater generated in shale gas development typically contains the following compositions: dissolved salts, minerals, residual fracturing fluid additives, heavy metals, bacteria, suspended solids, naturally occurring radioactive material, volatile organics, hydrocarbons, and ammonia (Karapataki 2012). In general, this water can be classified by the amount of total dissolved solids (TDS) per liter. Based on the operational definition, water produced during the well completion stage is defined as flowback water. On the other hand, water is referred to as produced water when the well is under production. Notably, the volumetric flow rate of flowback water is significantly larger than that of produced water, and the produced water tends to have higher concentration of TDS, likely because of its longer residence time downhole as well as smaller flow rate. As a whole, we can observe the flow rate of wastewater decreases along with time while the salinity of wastewater increases with time (Slutz et al. 2012; Gaudlip and Paugh 2008). The wastewater can be handled in multiple ways: direct injection into the Class II disposal wells, centralized wastewater treatment, and onsite treatment for reuse in hydraulic fracturing operations (Slutz et al. 2012; Veil 2010). These options are briefly introduced in the following sections.

Class II Disposal Wells

Disposal wells for injection of brine associated with oil and gas operations are classified as Class II disposal wells. We use disposal wells for abbreviation in the rest of this chapter. The wastewater from shale gas production can be directly sent to disposal wells and pumped into deep impermeable rock layers. This option is chosen when nearby disposal wells are available and permitted for underground injection. For states where disposal wells are abundant, the underground injection is a cost-effective option. If it is not the case, such as Marcellus in Pennsylvania where only several disposal wells are reported with limited capacity, wastewater will have to be transported to out-of-state locations, such as Ohio, for disposal, making this option much less attractive due to high transportation cost. Additionally, the underground injection option is criticized for the risks of causing water contamination and inducing seismicity (EPA 2011; Vidic et al. 2013). Therefore, the application of underground injection must be strictly subject to corresponding well capacity and regulations in case of unexpected environmental issues.

Centralized Wastewater Treatment Facility

The wastewater from shale gas production can also be transported to centralized wastewater treatment (CWT) facilities for treatment. The treated water is then discharged to surface water or recycled to shale sites for reuse. The resulting concentrated brine is sent to underground injection or taken down to zero liquid discharge condition and disposed as solid waste (API 2010; Puder and Veil 2006). CWT facilities involve a sequence of treatment processes. In general, the first step is fine particle filtration where the wastewater is screened to remove large objects, and then water is pumped into settling tanks to allow settling of heavy solids and removal of free oil. The second step is softening, where feedstock goes through agitation, aeration, pH adjustment, etc. to soften the water. The third step is

ultrafiltration, where particulates and macromolecules are removed. This step is followed by reverse osmosis/nanofiltration/thermal distillation, where most salts and other effluent materials are removed. Finally, certain toxic elements such as boron are removed to meet the specifications of surface discharge or reuse (Veil 2010). The advantages of CWT treatment option are its lower treatment cost compared with onsite treatment option and the potential to reduce freshwater consumption by recycling. However, the economic viability of this option depends on the proximity of CWT facilities to shale sites as well.

Onsite Treatment for Reuse

The last option is onsite treatment for reuse, where some mobile water treatment units are installed at shale site to treat the wastewater. There are three levels of onsite treatment, each of which corresponds to different technologies and treatment results (Veil 2010; Acharya et al. 2011). The primary treatment is all about clarification, where suspended matter, free oil and grease (FOG), iron, and microbiological contaminants are removed. Technologies for primary treatment include coagulation, flocculation and disinfection, electro-coagulation microfiltration/ultrafiltration, adsorption, ozonation, etc. The secondary treatment involves softening, where harness ions are removed. The corresponding technologies include lime softening, ion exchange, activated carbon, and so on. The tertiary treatment focuses on desalination to reduce the TDS. Major technologies for tertiary treatment include membrane separation, electrically driven membrane separation, thermal distillation, and zero liquid discharge. Since the primary and secondary treatments only partially treat the wastewater, a certain amount of make-up water is required for blending to satisfy the reuse specification for TDS concentration. For tertiary treatment, blending with freshwater is considered as a way of pretreatment to reduce the TDS concentration, so that the feed water can be treated more effectively.

2.3.1.2 Optimization Models for Water Management

Most existing publications addressing water management problem focus on evaluating the environmental impacts of hydraulic fracturing and providing techno-economic analysis of specific water management options (Slutz et al. 2012; Vidic et al. 2013; Goldstein 2014; McHugh et al. 2014; Horner et al. 2011). However, considering the shale water supply chain as a complex system that involves numerous technology alternatives, it is important to develop an integrated approach to address all the challenges and opportunities of the water-energy system simultaneously.

Yang et al. (2014) first propose an optimization model for shale gas water management. Key aspects for water use in hydraulic fracturing, including source water acquisition, wastewater production, reuse and recycle, subsequent transportation, storage, and disposal are considered. A discrete-time two-stage stochastic MILP model is proposed to address the uncertainty of water availability. In this work, two specific problems are considered. The first problem mainly focuses on the water acquisition stage, and the goal is to find the optimal water acquisition

strategy regarding different water sources. The uninterruptible sources are available throughout the year but require expensive truck transportation. The interruptible sources can be transported by pipeline with lower cost but are affected by seasonal availability. Storage option is also taken into account to coordinate water acquisition and demand. The objective is defined as minimizing the expected trucking and pumping cost of the water required to complete all the well sites. According to the optimization results of an example case, the total expected cost is reduced by $2.4 million compared against a heuristic schedule. The second problem addresses a more comprehensive model, where the handling of wastewater and revenue from gas production are taken into account. The goal is to determine the optimal fracturing schedule, as well as logistics for water acquisition, flowback reuse, and treatment. The objective is maximizing the expected profit from shale gas production after considering the operating cost for water management. Through this optimization, the expected profit is increased by 37 % from $156.41 million to $214.15 million, and the total cost is reduced from $25.02 million to $23.41 million compared to the heuristic schedule. Notably, this work is later extended by Yang et al. (2015) to optimize the long-term investment decisions using a deterministic MILP model. Multiple design decisions including capacity of water impoundments, pipeline options, treatment technologies and facility locations, as well as the fracturing schedule are addressed.

In order to explore more sustainable solutions for the shale water management. Gao and You (2015) develop a novel modeling framework for integrated design and operations of a water supply chain network for shale gas production. In this work, both the strategic design decisions (selection of freshwater sources, selection of transportation modes and capacities, installation of onsite treatment facilities, etc.), and operational decisions (water acquisition, water treatment, storage and transportation, etc.), are considered. Tradeoff between cost effectiveness and freshwater conservancy is captured with explicit consideration of multiple water management options and corresponding technologies. A novel fractional objective function is considered, which reflects the profit associated with unit net consumption of freshwater. The resulting problem is formulated as an MILFP and further solved by some tailored solution algorithms. To demonstrate the advantage of the proposed MILFP model, an MILP model targeting for maximizing the total profit is considered for comparison. By solving a large-scale case study based on Marcellus shale play, the results show that: (1) By consuming the same amount of freshwater resource, the MILFP model generates 14 % more profit and saves 12 % net water consumption than the MILP model. (2) Although the CWT is more cost effective in terms of water treatment, the onsite treatment option generates more balanced solution with better water efficiency, and the resulting reduction in water consumption can be significant. (3) The optimal water management strategy can be variant depending on the specific shale play and corresponding regulations (e.g. see Fig. 2.3 for the results obtained from the MILFP model).

Recently, Lira-Barragán et al. (2016) present a mathematical programming formulation for synthesizing water networks associated with shale gas hydraulic fracturing operations while accounting for uncertainties. The uncertainties

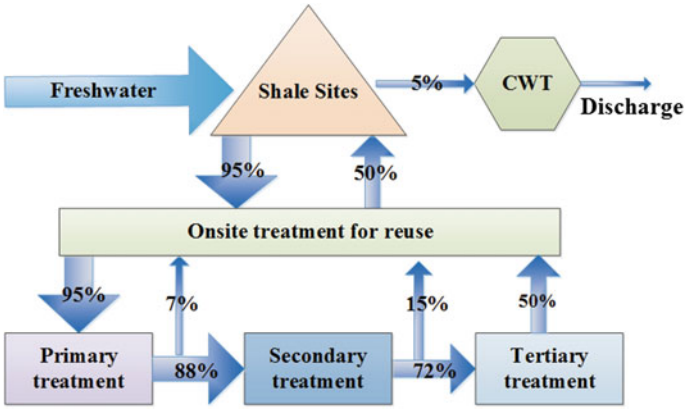

Fig. 2.3 Optimal water management strategy for MILFP model (Gao and You 2015)

correspond to the water requirements to fracture each well and the portion of this water that returns as flowback water. A two-stage stochastic programming model is developed and solved to minimize the total expected annual cost. The goal of the introduced approach is to provide guidance for decision makers in managing water resources and in acquiring properly sizes for water management systems. Meanwhile, Lira-Barragán et al. (2016) propose a similar model formulation to include the water consumption in the optimization. The economic objective function consists of the minimization of the total annualized cost, and the environmental goal aims at minimizing the total fresh water requirements.

2.3.2 *Life Cycle Optimization of Shale Gas Supply Chains*

The life cycle carbon footprint of shale gas is another popular topic that has been extensively addressed by using life cycle assessment approaches (Weber and Clavin 2012; Heath et al. 2014). Nevertheless, a research need of decision-support tools and methodologies dedicated to the sustainable design and operations of a shale gas supply chain system can be identified. To fill the knowledge gap in this research area, Gao and You (2015) develop a functional-unit-based life cycle optimization model for the optimal design and operations of shale gas supply chains. Such a shale gas supply chain covers the "well-to-wire" life cycle of shale gas, in which decisions regarding network design, drilling scheduling, water management, technology selection, facility location and sizing, natural gas storage, and transportation are fully captured. The levelized cost of electricity generated from shale gas is chosen as the economic indicator, and the environmental indicator is defined as GHG emissions per unit amount of electricity generation. A general superstructure of this shale gas supply chain is shown in Fig. 2.4.

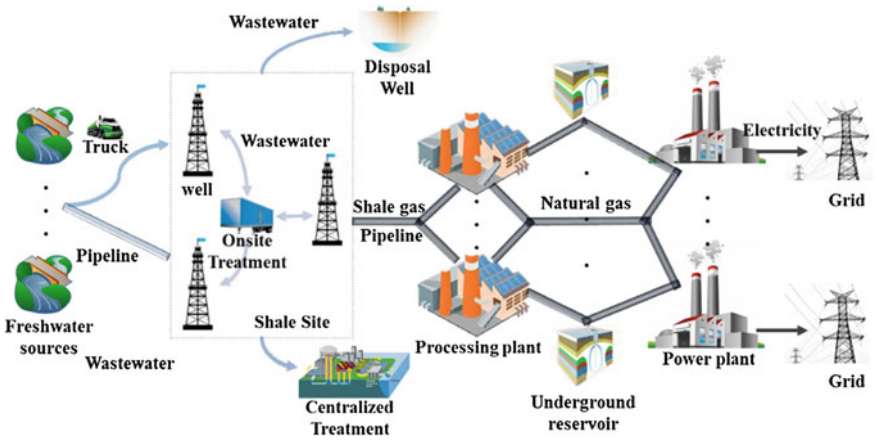

Fig. 2.4 Superstructure of the "well-to-wire" shale gas supply chain (Gao and You 2015)

The resulting problem is a multi-objective, multi-period MINLP problem, which is further solved by applying a tailored global optimization method integrating both the parametric algorithm and branch-and-refine approach. The proposed modeling framework and solution algorithm are demonstrated by a specific case study based on the Marcellus shale play in southwest PA. By solving this life cycle optimization problem, a Pareto-optimal curve consisting of 10 Pareto optimal solutions is obtained. The corresponding figure is presented in Fig. 2.5, which shows the trade-offs between the levelized cost of electricity and unit GHG emissions.

As can be seen in Fig. 2.5, point A has the lowest GHG emissions per unit electricity generation of 443 kg CO_2e/MWh, and it has the highest levelized cost of

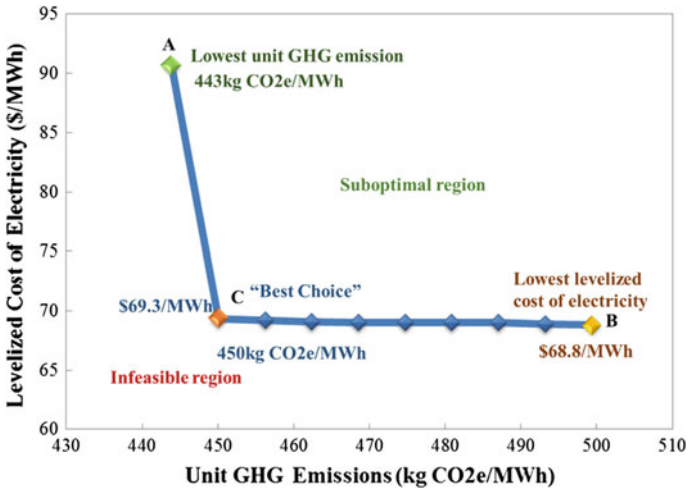

Fig. 2.5 Pareto-optimal curve of the shale gas supply chain (Gao and You 2015)

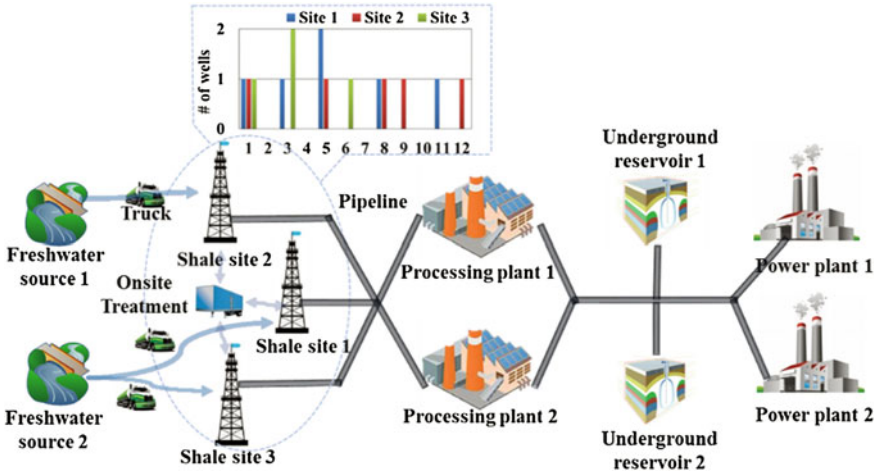

Fig. 2.6 Optimal shale gas supply chain strategies for "Best Choice" solution (Gao and You 2015)

electricity of $69/MWh. Point B has the lowest levelized cost of electricity of $69/MWh and the highest GHG emissions per unit electricity generated of 499 kg CO_2e/MWh. Point C is considered as the "best choice" solution, whose levelized cost of electricity is $69/MWh and unit GHG emissions us 450 kg CO_2e/MWh. By reviewing the cost and emission breakdowns, we identify the power generation as the primary source of GHG emissions, and the shale gas production and processing contribute to the largest portion of total cost. The optimal shale gas supply chain strategies for the "Best Choice" (point C) are given in Fig. 2.6.

The optimal strategies can be summarized by the following points: (1) Trucks are more flexible and cost-effective option for freshwater transportation considering the life cycle performance. (2) More pipeline links can potentially improve the overall economic performance of shale gas supply chain in the long term. (3) A relatively "evenly-distributed" shale well drilling schedule is the key to maintain a stable shale gas production profile, and thus benefiting the overall supply chain performance. (4) Managing wastewater with reverse osmosis (RO) technology onsite is identified as the most sustainable wastewater treatment strategy. (5) Underground reservoirs are important "buffers" coordinating the drilling activities and market demand for a better economic performance.

2.4 Shale Gas Processing and Conversion

The previous sections mainly discuss the supply chain optimization problems in shale gas industry. In this section, we focus on the shale gas processing optimization from a process design perspective. Potential opportunities regarding co-production of various chemicals from shale gas feedstock are discussed.

2.4.1 Shale Gas Processing Plant

A general design for the conventional shale gas processing plant is presented in Fig. 2.7. As shown below, first the raw shale gas pipelined from wellheads goes through a gas sweetening section, where the acid impurities, such as H_2S and CO_2, are removed. Considering the mutable characteristic of shale gas composition, three alternative schemes are employed to efficiently neutralize the raw gas as follows: (1) when the gas is only slightly sour, fixed-bed type scavenger process would be a cost-effective approach to H_2S removal; (2) chemical absorption-based acid gas removal (AGR) process followed by a scavenger process, works well for raw gas with moderate to high content of CO_2 and small content of H_2S; and (3) for moderate amounts of H_2S, the sulfur must be captured by a sulfur recovery unit when its amount exceeds a limit specified by the environmental regulations (Parks et al. 2010). The shale gas input after the gas sweetening is considered as sweet gas, which will pass through a dehydration section to remove the water vapor typically using the regenerable adsorption in liquid triethylene glycol. Next, a NGL recovery section uses a turbo-expansion configuration combined with an external refrigerant designed to recover about 80 % of the ethane from the dry gas. As seen in this figure, the remaining gas (mainly methane) is finally compressed as pipeline gas or sent to an N_2 rejection section depending on the N_2 concentration; otherwise, high N_2 content (>4 mol%) would make the heating value of the pipeline gas lower than specified (Natural Gas Processing 2006). The nitrogen rejection process employs a cryogenic distillation unit integrated with a heat pump system. Besides, marketable NGL products including ethane, propane, butanes, and pentanes, etc. are sequentially extracted by passing through a fractionation train consisting of a deethanizer, a depropanizer, and a debutanizer. Likewise, some strict specifications on NGL products should be considered like Y-grade (Y-Grade Product Specifications 2012).

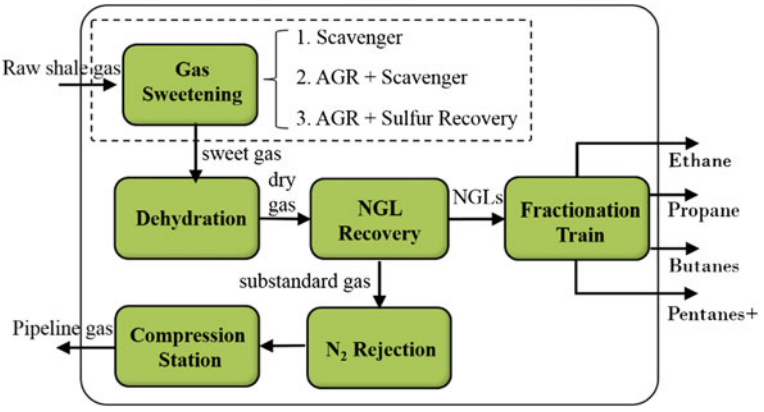

Fig. 2.7 Conventional shale gas processing plant

2.4.2 Shale Gas Conversion Scheme

Shale gas boom significantly increased interest in C_1, C_2, and C_3 chemistry to convert methane, ethane, and propane to value-added products, respectively, as shown in Fig. 2.8. In chemical industries, one-carbon-molecule compounds (like methanol, formaldehyde, formic acid, methylene chloride, etc.) tend to be derived from C_1 methane or carbon monoxide (itself derived from methane), two-carbon compounds (ethylene, acetaldehyde, ethanol, acetic acid, ketene, ethylene glycol, etc.) from C_2 ethane, three-carbon compounds (propylene, acetone, isopropyl alcohol, acrylic acid, etc.) from C_3 propane, and so on (Siirola 2014; Mitchell and Shantz 2015). In addition, chemicals with larger molecules could be formed through reactive functional groups including ether, esters, amides, etc. In particular, it is well-known that methane is always the dominant but less valuable ingredient in shale gas. The current industrial practice of steam reforming (SMR) produces syngas, which can be further used as an intermediate to produce other chemical commodities. Besides, methane and ethane are recognized as two major chemicals that potentially lead to integration opportunities with other chemical systems (He and You 2014; Ehlinger et al. 2014; Martín and Grossmann 2013). These opportunities include: (1) producing liquid fuels from methane, (2) producing methanol from methane, and (3) producing ethylene from ethane. This section focuses on relevant novel process designs and integrations and review of corresponding literature.

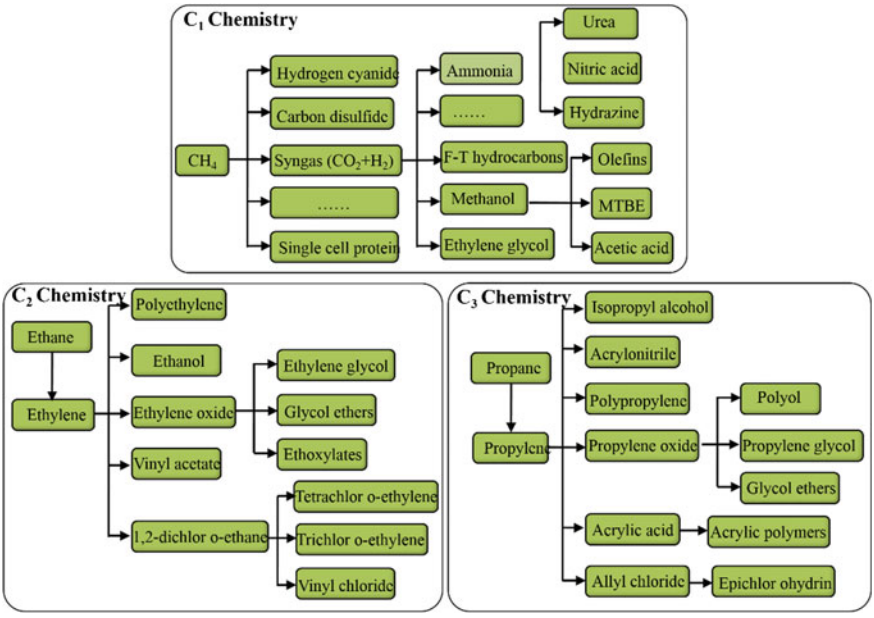

Fig. 2.8 Typical C_1, C_2, and C_3 production schemes (Siirola 2014)

To address the utilization of shale-derived natural gas, Martín and Grossmann (2013) present a superstructure optimization approach for the simultaneous production of liquid fuels and hydrogen from switchgrass and shale gas. This process is based on Fischer-Tropsch (FT) technology, in which shale gas is reformed with steam, while the switchgrass is gasified and reformed. After the raw gas is cleaned up and its composition is adjusted, the sour gases are removed. Finally, the liquid fuels, namely fuel gas, gasoline, and diesel, are produced using an FT reactor. By implementing a superstructure optimization approach, this integrated production system is able to achieve a production cost lower than $1/gal if biomass price is below $100/t and shale gas price is below $11.5/MMBTU.

In addition to producing liquid fuels from natural gas. Ehlinger et al. (2014) present a shale gas processing design that aims to produce methanol with shale gas feedstock. The raw shale gas has to undergo a set of preprocessing steps, including acid gas removal, dehydration, and nitrogen removal. Then the "pipeline quality" gas can be sent to a methanol plant as feedstock. In order to produce methanol, the first step is to produce syngas, which can be achieved via partial oxidation, stream reforming, or autothermal reforming. In another work by Noureldin et al. (2014), an optimization-based model is proposed targeting on modeling and selection of reforming approaches for syngas generation from natural/shale gas. In this work, the partial oxidation is selected as the reforming technology. After going through the partial oxidation reactor, vapor-liquid equilibrium separator (VLE-SEP), water-gas shift reactor (WGS), another VLE-SEP, and a CO_2 separator, the syngas is obtained and ready for the methanol production. The methanol synthesis occurs as a combination of two reactions in the syngas mixture in the MeOH reactor under low temperatures (503–533 K) and high pressures (50–100 atm).

$$CO_{(g)} + H_2O_{(g)} \leftrightarrow H_{2(g)} + CO_{2(g)} \quad \Delta H^{298} = -41\,\text{kJ/mol} \tag{1}$$

$$CO_{2(g)} + 3H_{2(g)} \leftrightarrow CH_3OH_{(l)} + H_2O_{(g)} \quad \Delta H^{298} = -50\,\text{kJ/mol} \tag{2}$$

By conducting a techno-economic analysis based on the gas-to-methanol process, production of methanol from shale gas is proven to be profitable for a broad range of methanol selling price and shale gas costs. A desirable 31 % ROI can be achieved in the case study. Following this work, an extensive economic and environmental analysis for the production of methanol from shale gas is presented (Julián-Durán et al. 2014), in which four types of reforming technologies, namely partial oxidation, steam methane reforming, autothermal reforming, and combined reforming are considered. The partial oxidation and autothermal reforming are identified with better economic performances, while the combined reforming outperforms other alternatives with the best environmental performance.

Apart from methane, ethane is an important feedstock for ethylene production. Therefore, integration between shale gas processing and ethylene production is another popular research direction. He and You (2014) propose three novel process designs for integrating shale gas processing with ethylene production. These novel

process designs include shale gas sequential processing (SSP), ethane cracking gas (a processable mixture including C_2H_4, H_2, C_2H_6, CH_4, etc.) recycling to NGL recovery (CRN), and cracking gas recycling to dehydration (CRD) designs. A unique feature of the proposed process designs is the co-processing strategy of shale gas and cracking gas. The SSP design has an identical processing structure as that of a conventional shale gas processing plant, but the ethane extracted from NGLs is pipelined to a local ethylene plant as feedstock. In the ethylene plant, ethane is converted into cracking gas using thermal steam cracking technology. The cracking gas then passes through the dehydration area, followed by a cracking gas separation area where ethylene and other products are produced and unconverted ethane is recycled. The CRN design adopts a modified NGL recovery section and fractionation train section. As a result, the cracking gas from ethylene plant can be co-processed by the modified area with dry gas after the dehydration area. Moreover, the cracking gas separation area can be eliminated from this design. The CRD design considers another recycling strategy, in which the produced cracking gas together with sweet gas are introduced to a centralized dehydration area for further purification and separation. Compared with the CRN design, the dehydration area in the ethylene plant is removed. Based on the proposed shale gas processing designs, the authors develop detailed thermo-economic models and energy analysis for the process designs. The economic analysis reveals that the estimated NPVs in the proposed SSP, CRN, and CRD designs are 1.7–2.4 times greater than that of conventional processing design. Furthermore, the NGL-richer shale gas generally produces 3.17–5.12 times estimated NPV than NGL-leaner shale gas.

Following this work, the authors extend the scope and further develop a novel process design for making chemicals from shale gas and bioethanol (He and You 2015). Such a process design consists of four major process areas: gas treatment, gas to chemicals, methane-to-ethylene, and bioethanol-to-ethylene. As always, the gas treatment aims to purify the raw shale gas, where acid gas and water vapor are either removed or recovered, and dry gas is ready to enter the next process area, namely the gas to chemicals area. In this area, heavy NGLs (butanes and natural gasoline) and low-boiling gases (methane, H_2, etc.) are successively taken from the dry gas using an NGL cutting unit. The C_2 and C_3 hydrocarbons in the remaining steam are then sent to an olefin separation unit followed by an olefin production unit. Meanwhile, the H_2-rich gas goes through a pressure swing adsorption (PSA) unit for hydrogen recovery. Next, methane from the NGL cutting unit is sent to the methane-to-ethylene area and then partially oxidized via the oxidative coupling of methane (OCM) reaction in an O_2-rich environment. The OCM product includes ethylene, ethane, hydrogen, carbon dioxide, carbon monoxide, and unconverted methane. Finally, in the bioethanol-to-ethylene area, bioethanol is used as a renewable feedstock for producing ethylene with tail gas from PSA as the dehydration fuel. Such a shale gas process design is further optimized by using a simulation-optimization method based on the NSGA-II algorithm. In the "good choice" optimal design, the minimum ethylene selling price can be reduced to \$877.2/ton, and the unit GWP of ethylene is 0.360 kg CO_2-eq/kg in the high carbon shale gas scenario. In the low carbon scenario, the results are \$655.1/ton and

0.030 kg CO_2-eq/kg, respectively. The results reveal that shale gas can be converted to more cost-effective and greener chemicals with proper process design, integration, and optimization.

Most recently, in order to decipher the true production costs and environmental impacts of shale gas-to-olefin (STO) projects, He and You (2016) develop a mega-scale shale gas supply chain olefin production network model with explicit consideration of process designs, integration methods, and alternative technologies. A techno-economic-environmental life cycle analysis is conducted for systematically evaluating the energy-water-carbon nexus. Four major shale regions of the U. S., including Appalachian, Gulf Coast, Mid-Continent, and Rocky Mountain regions are considered. A total of 594,922 shale wells are involved. In the proposed STO process, the petrochemical plant is co-located with shale plays and gas processing facilities. The raw shale gas is processed in multisite distributed processing facilities, and the recovered NGLs are moved to a centralized steam cracking plant, where the NGLs are fractionated and pyrolyzed. The results obtained from this work indicates that the STO located at the Mid-Continent region has relatively low environmental impacts. Besides, shale gas is still a low-carbon feedstock though its GHG emissions are 15 % higher than NG-ethane on average. Based on the sensitivity analysis, the well lifetime is identified as the critical factor in evaluating the overall environmental footprints.

In addition to the aforementioned processes and integrations, there are multiple alternatives that remain to be further expanded. For instance, Wang et al. (2013) propose a highly efficient cold energy integrating scheme by integrating NGL recovery from shale gas and LNG regasification at receiving terminals. The goal of this study is to recycle the cold energy from LNG regasification process to assist the NGL recovery process for economic improvement and energy saving. A general methodology framework is proposed and further decomposed into four steps: the first step is to develop the process superstructure and prepare corresponding data; then, a simulation-assisted MILP model is developed and solved for the optimal process synthesis; next, heat exchange network design and analysis are performed based on the pinch technology to accomplish the maximum energy saving target; finally, rigorous plant-wide simulations are conducted to validate the feasibility and capability of the entire process design coupling of separation and heat integration. By comparing the optimal integrated design proposed in this work with independent LNG regasification and NGL recovery processes, the authors observe a 61.8 % reduction in hot utility and 100 % reduction of cold utility.

2.5 Future Directions

Although the shale gas industry has been developed for decades, there are many research challenges on applying mathematical programming tools for shale gas process and supply chain optimization problems. These challenges include, but are not limited to: optimization of shale gas system under uncertainty, modeling and

optimization of shale gas supply chain from a non-cooperative perspective, and multi-scale optimization integrating shale gas supply chain and processing system.

First, uncertainty is ubiquitous in shale gas supply chains. Multiple uncertainties have been identified with significant influence on the overall performance of shale gas system, such as freshwater availability, EUR of shale wells, composition of shale gas, and price of natural gas. Despite the great importance of hedging against uncertainty, most of the existing literature either fail to capture such a key issue or only consider some types of uncertainties that are easy to address. With the recent development of optimization techniques, such as multi-stage robust optimization and data-driven stochastic programming, we expect more modeling frameworks and applications to be proposed in the optimization of shale gas system under various uncertainties.

Besides, a shale gas supply chain could consist of multiple stakeholders, such as the upstream shale producers, midstream processors, and downstream distributors and customers. In general, each of them pursues its own objective and makes decisions independently. Nevertheless, almost all of the existing studies rely on the cooperative shale gas supply chain model, in which a single decision maker is assumed to have full control of the entire system. Consequently, the optimal strategy obtained from a centralized model can be practically infeasible. In order to better capture the performance of a shale gas supply chain, it is necessary to properly address the non-cooperative relationship between multiple stakeholders in shale gas supply chain optimization problems.

Last, shale gas process design and supply chain optimization are both important problems with different spatial scales. Research addressing each of the systems has unveiled great economic and environmental benefits. However, there is no modeling framework addressing shale gas process design and supply chain management simultaneously for better economic and environmental performances. By developing a holistic systematic approach to integrate these two systems, it will lead to better overall performance and more potential design alternatives that can facilitate the development of shale gas industry.

2.6 Conclusion

This chapter provides a comprehensive introduction on the shale gas supply chain and processing system. Major research areas including the optimal design and operations of shale gas supply chain, water management and emission mitigation in shale gas development, and novel shale gas processing designs and integrations are identified. Corresponding literature is reviewed and discussed in details. Furthermore, potential research opportunities are summarized, including optimization under various uncertainties, modeling and optimization of non-cooperative supply chain, and simultaneous optimization of shale gas supply chain and process systems.

References

Acharya, H. R., Henderson, C., Matis, H., Kommepalli, H., & Wang, H. (2011). *Cost effective recovery of low TDS frac flowback water for re-use*. Niskayuna, NY 12309-1027: Department of Energy.

Allen, D. T. (2014a). Methane emissions from natural gas production and use: Reconciling bottom-up and top-down measurements. *Current Opinion in Chemical Engineering, 5*, 78–83. doi:10.1016/j.coche.2014.05.004.

Allen, D. T. (2014b). Atmospheric emissions and air quality impacts from natural gas production and use. *Annual Review of Chemical and Biomolecular Engineering, 5*(1), 55–75. doi:10.1146/annurev-chembioeng-060713-035938.

Alvarez, R. A., Pacala, S. W., Winebrake, J. J., Chameides, W. L., & Hamburg, S. P. (2012). Greater focus needed on methane leakage from natural gas infrastructure. *Proceedings of the National Academy of Sciences, 109*(17), 6435–6440. doi:10.1073/pnas.1202407109.

API. (2010). *Water management associated with hydraulic fracturing*. Institute AP.

Brandt, A. R., Heath, G. A., Kort, E. A., O'Sullivan, F., Pétron, G., Jordaan, S. M., et al. (2014). Methane leaks from North American natural gas systems. *Science, 343*(6172), 733–735. doi:10.1126/science.1247045.

Burnham, A., Han, J., Clark, C. E., Wang, M., Dunn, J. B., & Palou-Rivera, I. (2011). Life-cycle greenhouse gas emissions of shale gas, natural gas, coal, and petroleum. *Environmental Science and Technology, 46*(2), 619–627. doi:10.1021/es201942m.

Cafaro, D. C., & Grossmann, I. E. (2014). Strategic planning, design, and development of the shale gas supply chain network. *AIChE Journal, 60*(6), 21. doi:10.1002/aic.14405.

Chaudhri, M. M. (2012). Numerical modeling of multifracture horizontal well for uncertainty analysis and history matching: Case studies from Oklahoma and Texas shale gas wells.

Chima, C. M. (2011). Supply-chain management issues in the oil and gas industry. *Journal of Business & Economics Research (JBER), 5*(6).

Dale, A. T., Khanna, V., Vidic, R. D., & Bilec, M. M. (2013). Process based life-cycle assessment of natural gas from the Marcellus Shale. *Environmental Science and Technology, 47*(10), 5459–5466. doi:10.1021/es304414q.

DOE/NETL. (2011). Life cycle greenhouse gas inventory of natural gas extraction, delivery and electricity production.

Drouven, M. G., & Grossmann, I. E. (2016). Multi-period planning, design and strategic models for long-term, quality-sensitive shale gas development. *AIChE Journal*. doi:10.1002/aic.15174.

Duran, M., & Grossmann, I. (1986). A mixed-integer nonlinear programming algorithm for process systems synthesis. *AIChE Journal, 32*(4), 592–606.

Ehlinger, V. M., Gabriel, K. J., Noureldin, M. M. B., & El-Halwagi, M. M. (2014). Process design and integration of shale gas to methanol. *ACS Sustainable Chemistry & Engineering, 2*(1), 30–37. doi:10.1021/sc400185b.

EIA. (2011). *Review of emerging resources: U.S. shale gas and shale oil plays*. Washington, DC 20585: U.S. Energy Information Administration.

EIA. (2015). Annual Energy Outlook 2015 with projections to 2040. U.S. Energy Information Administration, Washington, DC 20585.

EIA Natural Gas Consumption by End Use. Retrieved June 16, 2015, from http://www.eia.gov/dnav/ng/ng_cons_sum_dcu_spa_a.htm.

EIA Underground Natural Gas Storage. Retrieved September 29, 2014, from http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/ngpipeline/undrgrnd_storage.html.

EPA. (2011). *Plan to study the potential impacts of hydraulic fracturing on drinking water resources*. EPA, Washington, D.C.: Office of Research and Development U.S.

Gao, J., & You, F. (2015a). Deciphering and handling uncertainty in shale gas supply chain design and optimization: Novel modeling framework and computationally efficient solution algorithm. *AIChE Journal, 61*(11), 3739–3755. doi:10.1002/aic.15032.

Gao, J., & You, F. (2015b). Optimal design and operations of supply chain networks for water management in shale gas production: MILFP model and algorithms for the water-energy nexus. *AIChE Journal, 61*(4), 1184–1208. doi:10.1002/aic.14705.

Gao, J., & You, F. (2015c). Shale gas supply chain design and operations toward better economic and life cycle environmental performance: MINLP model and global optimization algorithm. *ACS Sustainable Chemistry & Engineering, 3*(7), 1282–1291. doi:10.1021/acssuschemeng.5b00122.

Gaudlip, A.W., & Paugh, L. O. (2008). Marcellus Shale water management challenges in Pennsylvania. *Paper presented at the shale gas Production Conference*. Fort Worth, Texas, 16–18 November.

Goldstein, B. D. (2014). The importance of public health agency independence: Marcellus shale gas drilling in Pennsylvania. *American Journal of Public Health, 104*(2), e13–e15. doi:10.2105/ajph.2013.301755.

Gracceva, F., & Zeniewski, P. (2013). Exploring the uncertainty around potential shale gas development—A global energy system analysis based on TIAM (TIMES Integrated Assessment Model). *Energy, 57*, 443–457. doi:10.1016/j.energy.2013.06.006.

Gregory, K. B., Vidic, R. D., & Dzombak, D. A. (2011). Water management challenges associated with the production of shale gas by hydraulic fracturing. *Elements, 7*(3), 181–186. doi:10.2113/gselements.7.3.181.

Gupta, V., & Grossmann, I. E. (2012). An efficient multiperiod MINLP model for optimal planning of offshore oil and gas field infrastructure. *Industrial and Engineering Chemistry Research, 51*(19), 6823–6840. doi:10.1021/ie202959w.

Harding, N. R. (2008). Application of stochastic prospect analysis for shale gas reservoirs.

Harto, C. (2013). *Management of water from CCS: Life cycle water consumption for carbon capture and storage* (trans: Energy USDo). Argonne National Laboratory.

He, C., & You, F. (2014). Shale gas processing integrated with ethylene production: Novel process designs, exergy analysis, and techno-economic analysis. *Industrial and Engineering Chemistry Research, 53*(28), 11442–11459. doi:10.1021/ie5012245.

He, C., & You, F. (2015). Toward more cost-effective and greener chemicals production from shale gas by integrating with bioethanol dehydration: Novel process design and simulation-based optimization. *AIChE Journal, 61*(4), 1209–1232. doi:10.1002/aic.14713.

He, C., & You, F. (2016). Deciphering the true life cycle environmental impacts and costs of the mega-scale shale gas-to-olefins projects in the United States. *Energy & Environmental Science, 9*, 820–840. doi:10.1039/C5EE02365C.

Heath, G. A., O'Donoughue, P., Arent, D. J., & Bazilian, M. (2014). Harmonization of initial estimates of shale gas life cycle greenhouse gas emissions for electric power generation. *Proceedings of the National Academy of Sciences, 111*(31), E3167–E3176.

Horner, P., Halldorson, B., & Slutz, J. A. (2011). Shale gas water treatment value chain—A review of technologies including case studies. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Howarth, R. W. (2014). A bridge to nowhere: Methane emissions and the greenhouse gas footprint of natural gas. *Energy Science & Engineering, 2*(2), 47–60. doi:10.1002/ese3.35.

Hughes, J. D. (2013). A reality check on the shale revolution. *Nature, 494*(7437), 307–308.

Hughes, J. D. (2014). *Drilling deeper: A reality check on U.S. Government Forecasts for A Lasting Tight Oil & Shale Gas Boom*. California, Santa Rosa 95404: Post Carbon Institute.

Iyer, R. R., Grossmann, I. E., Vasantharajan, S., & Cullick, A. S. (1998). Optimal planning and scheduling of offshore oil field infrastructure investment and operations. *Industrial and Engineering Chemistry Research, 37*(4), 1380–1397. doi:10.1021/ie970532x.

Jayakumar, R., & Rai, R. R. (2012). Impact of uncertainty in estimation of shale gas reservoir and completion properties on EUR forecast and optimal development planning: A Marcellus case study.

Jiang, M., Hendrickson, C. T., & VanBriesen, J. M. (2014). Life cycle water consumption and wastewater generation impacts of a Marcellus shale gas well. *Environmental Science and Technology, 48*(3), 1911–1920. doi:10.1021/es4047654.

Julián-Durán, L. M., Ortiz-Espinoza, A. P., El-Halwagi, M. M., & Jiménez-Gutiérrez, A. (2014). Techno-economic assessment and environmental impact of shale gas alternatives to methanol. *ACS Sustainable Chemistry & Engineering, 2*(10), 2338–2344. doi:10.1021/sc500330g.

Karapataki, C. (2012). *Techno-economic analysis of water management options for unconventional natural gas developments in the Marcellus Shale*. Master thesis. Cambridge, MA 02139-4307: Massachusetts Institute of Technology.

Knudsen, B. R., & Foss, B. (2013). Shut-in based production optimization of shale-gas systems. *Computers & Chemical Engineering, 58*, 54–67. doi:10.1016/j.compchemeng.2013.05.022.

Knudsen, B. R., Grossmann, I. E., Foss, B., & Conn, A. R. (2014a). Lagrangian relaxation based decomposition for well scheduling in shale-gas systems. *Computers & Chemical Engineering, 63*, 234–249. doi:10.1016/j.compchemeng.2014.02.005.

Knudsen, B. R., Whitson, C. H., & Foss, B. (2014b). Shale-gas scheduling for natural-gas supply in electric power production. *Energy, 78*, 165–182. doi:10.1016/j.energy.2014.09.076.

Laurenzi, I. J., & Jersey, G. R. (2013). Life cycle greenhouse gas emissions and freshwater consumption of Marcellus Shale gas. *Environmental Science and Technology, 47*(9), 4896–4903. doi:10.1021/es305162w.

Lira-Barragán, L. F., Ponce-Ortega, J. M., Guillén-Gosálbez, G., & El-Halwagi, M. M. (2016a). Optimal water management under uncertainty for shale gas production. *Industrial and Engineering Chemistry Research, 55*(5), 1322–1335. doi:10.1021/acs.iecr.5b02748.

Lira-Barragán, L. F., Ponce-Ortega, J. M., Serna-González, M., & El-Halwagi, M. M. (2016). Optimal reuse of flowback wastewater in hydraulic fracturing including seasonal and environmental constraints. *AIChE Journal* n/a–n/a. doi:10.1002/aic.15167.

Martín, M., & Grossmann, I. E. (2013). Optimal use of hybrid feedstock, switchgrass and shale gas for the simultaneous production of hydrogen and liquid fuels. *Energy, 55*, 378–391. doi:10.1016/j.energy.2013.04.005.

Mauter, M., & Palmer, V. (2014). Expert elicitation of trends in Marcellus oil and gas wastewater management. *Journal of Environmental Engineering, 140*(5), B4014004. doi:10.1061/(ASCE)EE.1943-7870.0000811.

Mauter, M. S., Palmer, V. R., Tang, Y., & Behrer, A. P. (2013). The next frontier in United States shale gas and tight oil extraction: Strategic reduction of environmental impacts. Belfer Center for Science and International Affairs Discussion Paper Series.

Mauter, M. S., Alvarez, P. J. J., Burton, A., Cafaro, D. C., Chen, W., Gregory, K. B., et al. (2014). Regional variation in water-related impacts of shale gas development and implications for emerging international plays. *Environmental Science and Technology, 48*(15), 8298–8306. doi:10.1021/es405432k.

McHugh, T., Molofsky, L., Daus, A., & Connor, J. (2014). Comment on "an evaluation of water quality in private drinking water wells near natural gas extraction sites in the Barnett Shale formation". *Environmental Science and Technology, 48*(6), 3595–3596. doi:10.1021/es405772d.

Mitchell, S. F., & Shantz, D. F. (2015). Future feedstocks for the chemical industry—Where will the carbon come from? *AIChE Journal, 61*(8), 2374–2384. doi:10.1002/aic.14910.

Natural gas processing: The crucial link between natural gas production and its transportation to market (2006). Washington, DC: Energy Information Administration.

Nicot, J.-P., & Scanlon, B. R. (2012). Water use for shale-gas production in Texas, U.S. *Environmental Science and Technology, 46*(6), 3580–3586. doi:10.1021/es204602t.

Nicot, J. P., Scanlon, B. R., Reedy, R. C., & Costley, R. A. (2014). Source and fate of hydraulic fracturing water in the Barnett Shale: A historical perspective. *Environmental Science and Technology, 48*(4), 2464–2471. doi:10.1021/es404050r.

Noureldin, M. M. B., Elbashir, N. O., & El-Halwagi, M. M. (2014). Optimization and selection of reforming approaches for syngas generation from natural/Shale gas. *Industrial and Engineering Chemistry Research, 53*(5), 1841–1855. doi:10.1021/ie402382w.

Paper, W. (2008). *US shale gas—An unconventional resource, unconventional challenge*. Halliburton.

Parks, L. E., Perry, D., & Fedich, R. (2010). FLEXSORB ®SE A proven reliable acid gas enrichment solvent A2—Benyahia, Farid. In F. T. Eljack (Ed.), *Proceedings of the 2nd Annual Gas Processing Symposium* (Vol. 2, pp. 229–235). Amsterdam: Elsevier. doi:10.1016/S1876-0147(10)02025-2.

Puder, M. G., & Veil, J. A. (2006). *Offsite commercial disposal of oil and gas exploration and production waste: Availability, options, and costs* (trans: Division ES). Argonne National Laboratory for the U.S. Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory.

Rahm, B. G., & Riha, S. J. (2012). Toward strategic management of shale gas development: Regional, collective impacts on water resources. *Environmental Science & Policy, 17*, 12–23. doi:10.1016/j.envsci.2011.12.004.

Seydor, S. M., Clements, E., Pantelemonitis, S., & Deshpande, V. (2012). *Understanding the Marcellus Shale supply chain*. Pittsburgh, PA 15260: University of Pittsburgh, Katz Graduate School of Business.

Siirola, J. J. (2014). The impact of shale gas in the chemical industry. *AIChE Journal, 60*(3), 810–819. doi:10.1002/aic.14368.

Slutz, J. A., Anderson, J. A., Broderick, R., & Horner, P. H. (2012). Key Shale gas water management strategies: An economic assessment. In *International Conference on Health Safety and Environment in Oil and Gas Exploration and Production*. Perth, Austrailia: Society of Petroleum Engineers.

Small, M. J., Stern, P. C., Bomberg, E., Christopherson, S. M., Goldstein, B. D., Israel, A. L., et al. (2014). Risks and risk governance in unconventional shale gas development. *Environmental Science and Technology, 48*(15), 8289–8297. doi:10.1021/es502111u.

Soeder, D. J., & Kappel, W. M. (2009). *Water resources and natural gas production from the Marcellus Shale*. Virginia: US Department of the Interior, US Geological Survey Reston.

Stephenson, T., Valle, J. E., & Riera-Palou, X. (2011). Modeling the relative GHG emissions of conventional and shale gas production. *Environmental Science and Technology, 45*(24), 10757–10764. doi:10.1021/es2024115.

Swindell, G. S. (2014). Marcellus Shale in Pennsylvania: A 2,600 well study of estimated ultimate recovery. *Paper presented at the SPE Annual Meeting*. Dallas, TX.

van den Heever, S. A., & Grossmann, I. E. (2000). An iterative aggregation/disaggregation approach for the solution of a mixed-integer nonlinear oilfield infrastructure planning model. *Industrial and Engineering Chemistry Research, 39*(6), 1955–1971. doi:10.1021/ie9906619.

Veil, J. A. (2010). *Final report water management technologies used by Marcellus shale gas producers*. Argonne, IL: Oil & Natural Gas Technology, U.S. Department of Energy.

Vengosh, A., Warner, N., Jackson, R., & Darrah, T. (2013). The effects of shale gas exploration and hydraulic fracturing on the quality of water resources in the United States. In *Proceedings of the Fourteenth International Symposium on Water-Rock Interaction*, WRI (Vol. 14, No. 7, pp. 863–866). doi:10.1016/j.proeps.2013.03.213.

Vidic, R. D., Brantley, S. L., Vandenbossche, J. M., Yoxtheimer, D., & Abad, J. D. (2013). Impact of shale gas development on regional water quality. *Science, 340*(6134), 1235009. doi:10.1126/science.1235009.

Wang, M., Zhang, J., & Xu, Q. (2013). A novel conceptual design by integrating NGL recovery and LNG regasification processes for maximum energy savings. *AIChE Journal, 59*(12), 4673–4685. doi:10.1002/aic.14231.

Weber, C. L., & Clavin, C. (2012). Life cycle carbon footprint of Shale gas: Review of evidence and implications. *Environmental Science and Technology, 46*(11), 5688–5695. doi:10.1021/es300375n.

Yang, L., Manno, J., & Grossmann, I. E. (2014). Optimization models for shale gas water management. *AIChE Journal*. doi:10.1002/aic.14526.

Yang, L., Grossmann, I. E., & Manno, J. (2014b). Optimization models for shale gas water management. *AIChE Journal, 60*(10), 3490–3501. doi:10.1002/aic.14526.

Yang, L., Grossmann, I. E., Mauter, M. S., & Dilmore, R. M. (2015). Investment optimization model for freshwater acquisition and wastewater handling in shale gas production. *AIChE Journal, 61*(6), 1770–1782. doi:10.1002/aic.14804.

Y-Grade Product Specifications. (2012). ETC NGL Transport LLC. http://www.energytransfer.com/documents/UniformY-GradeSpecs-ETCV111612.pdf.

Zavala-Araiza, D., Allen, D. T., Harrison, M., George, F. C., & Jersey, G. R. (2015). Allocating methane emissions to natural gas and oil production from shale formations. *ACS Sustainable Chemistry and Engineering, 3*(3), 492–498. doi:10.1021/sc500730x.

Chapter 3
Strategic Design and Tactical Planning for Energy Supply Chain Systems

Omar J. Guerra, Andrés J. Calderón, Lazaros G. Papageorgiou and Gintaras V. Reklaitis

Abstract This chapter is intended to discuss the application of mathematical programing methodologies to assist in the design and planning of energy supply chains. Accordingly, a comprehensive literature review of various different energy systems models as well as optimization approaches which have been developed to support capacity planning and investment decisions is presented. Two mathematical programming-based frameworks, one for shale gas supply chains and another for interconnected power systems are briefly described and employed for relevant case studies. These case studies are used not only to illustrate the methodology but more importantly to draw some insights that are relevant to policy development.

Keywords Mathematical programming · Energy supply chains · Capacity planning · Shale gas · Power systems · Policy · Development

3.1 Introduction

In the current global primary energy mix, fossil fuels are the predominant source of energy. In 2014, 35.5 % of the energy demand was supplied by oil, followed by natural gas and coal with a share of 28.0 % and 18.3 %, respectively (BP 2015).

O.J. Guerra · G.V. Reklaitis (✉)
School of Chemical Engineering, Forney Hall
Purdue University, West Lafayette, IN 47907, USA
e-mail: reklaiti@purdue.edu

O.J. Guerra
e-mail: oguerra@purdue.edu

A.J. Calderón · L.G. Papageorgiou
Department of Chemical Engineering, Roberts Building
University College London, WC1E 7JE, London, UK
e-mail: andres.vergara.12@ucl.ac.uk

L.G. Papageorgiou
e-mail: l.papageorgiou@ucl.ac.uk

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_3

The nuclear power and renewable energies constituted only 8.5 % and 9.8 %, respectively, of the energy mix in the same year. Despite the remarkable advances in renewable technologies and the growing support of policy makers over the past 15 years, energy from renewable sources has increased only modestly from the level of 6.2 % of the total energy supply that was the case in 2000.

Over the last century (1900–2000), the primary energy consumption exhibited an almost 14-fold growth (Smil 2000). A more gradual increase is expected for the next 50 years in which fossil fuels will continue to be the major source of energy. The main driving forces for that increase are: population growth, urbanization and increasing living standards (ExxonMobil 2014). Furthermore, the energy sector has become the leading source of anthropogenic greenhouse gas (GHG) emissions, accounting for roughly two-thirds of global GHG emissions in recent years (International Energy Agency 2013a). CO_2 emissions represent around 90 % of the energy related GHG emissions, with NOx and SOx accounting for the remaining 10 %. Additionally, anthropogenic GHG emissions are recognized as one of the major causes of the global climate change, which will continue to aggravate if GHG emissions are not mitigated in the years to come. According to United Nations Framework Convention on Climate Change (UNFCCC), the global warming should be kept from exceeding 2 °C above the preindustrial temperature in order to avoid dangerous (irreversible) effects of climate change (http://unfccc.int/2860.php). The main challenge lies on how to design energy systems while considering three crucial aspects: environmental sustainability, energy security and energy equity. Certainly, this could be achieved through coordinated actions between governments, markets and industry. Energy security and energy equity involve the diversification of energy sources and the expansion of transmission and distribution networks, e.g. gas pipelines, oil pipelines, and power transmission networks, in order to not only guarantee a constant supply of energy but also to broaden the coverage to more regions. Furthermore, environmental sustainability, necessary to tackle the effects of climate change, entails not only the mitigation of GHG emissions from fossil sources but also the integration of renewable technologies into the conventional energy systems.

The integration of the aforementioned aspects in the development of future energy systems will inevitably result in more complex energy systems where a number of interdependent variables must be considered in the decision-making process. Mathematical modelling and optimization techniques are powerful tools that provide a systematic methodology to tackle decision problems of these kinds. A mathematical framework that incorporates important decision variables allows researchers, policy makers and shareholders to achieve a better understanding of the energy systems, identify key elements and trade-offs, and disclose synergies inherent to the nature of the problem.

Mixed integer linear programming (MILP) formulations, i.e. the optimization problem shown below in P_0, are powerful optimization approaches widely used to model and solve a variety of problems. Particularly, MILP formulations are very useful to assist in the long-term design and planning of energy supply chain systems. In planning problems, the binary variables, y, are typically used to represent strategic decisions that can be modeled as yes or no decisions. Examples of these

decisions include the selection of technologies, location, capacity, and construction period for energy conversion technologies as well as for transportation facilities. By contrast, the continuous variables x are mostly used to denote operational decisions. For instance, continuous variables can be used to represent mass and energy flows throughout the whole energy supply chain as well as to model stream properties such as composition, pressure, and temperature. The model constraints can be associated with the physical and thermodynamic laws governing each process in the system, i.e. capacity constraints, availability limits, and mass and energy constraints, as well as with the economics of the systems, i.e. operating cost, capital investment, and budget constraints. Commonly, the objective function is an economic metric used to quantify the performance of the system under different levels of the continuous and binary variables, i.e. net present value (NPV), return on investment (ROI) and the profit before taxes. A range of computational algorithms have been used to solve this type of optimization problems, including branch & bound (B&B) approaches and cutting plane methods. These methods are iterative and solve linear programming (LP) relaxations at each step and then fix some binary variables or add linear constraints (cuts) to the problem.

$$\begin{aligned} P_0\text{: } Min \quad & f(x,y) = c^T x + d^T y \\ & Ax + By \le b \\ & x \ge 0,\ x \in X \subseteq \Re^n \\ & y \in \{0,1\}^p \end{aligned} \tag{3.1}$$

This chapter is intended to demonstrate the relevance and power of mathematical modeling and optimization techniques in the design and planning of energy systems. Two different energy systems, shale gas supply chain and interconnected power systems, are discussed in the following sections including the associated MILP models. The solutions obtained by using these models on several case studies are used to draw some insights about important characteristics of these energy systems.

3.2 Shale Gas Supply Chain

During the last decade, the production of shale gas has progressively increased in importance in the energy sector, the economy of the United States, and more recently in the global economy. There is a strong indication that shale gas resources are and will continue to play an important role in the global energy market in the next decades.

Shale gas refers to natural gas trapped within sedimentary rocks formed by laminar deposits of fine-grained clay particles and organic matter. Geological processes of burial and compaction through continuous sedimentation on top provide the necessary temperature and pressure conditions to initiate the thermogenic

breakdown of the organic matter to generate oil and gas. Shale formations are characterized by extremely low matrix permeability, between 10^{-6} and 10^{-2} millidarcy (md) versus 100–10,000 md in conventional oil and gas reservoirs. Due to the low permeability, any migration processes of fossil resources towards upper formations are significant only over geologic time, i.e. millions of years. This has important repercussions from an industrial viewpoint. For example, the exploration stage is comparatively easier than in conventional reservoirs, however, the production stage is quite challenging. The exploitation of shale resources is highly dependent on artificial stimulation techniques, such as the hydraulic fracturing, which are used to increase the connectivity within the shale matrix, facilitating the flow of natural gas from the formation matrix to the well. Recent advances in horizontal well drilling and hydraulic fracturing technologies have made the development of shale gas reservoirs economically attractive. Nonetheless, the low matrix permeability limits the drainage area of the well mostly to the extent of the rock formation stimulated by the hydraulic fracturing. This results in recovery factors in the order of 20–30 % which are markedly lower than those for conventional natural gas resources, commonly between 80 and 90 % (Kaiser 2012a, b). Due to low recovery factors, companies involved in the exploitation of these resources opt for the implementation of intensive drilling schemes to benefit from economy of scales. Consequently, the feasibility and profitability of a shale gas reservoir are strongly determined by production costs and productivity, hence, minor changes in the market conditions can have substantial repercussions on the development of these resources. Moreover, a sustainable shale gas production is an essential aspect that can be handled through collaboration between policy makers, industry sector as well as engagement of neighboring communities. In particular, the depletion and degradation of water sources, as well as the potential for underground water contamination, are major drawbacks that could and do hinder the development of these resources (Clarkson 2013; Eaton 2013; Jenner and Lamadrid 2013; Siirola 2014; Vidic et al. 2013; Warner et al. 2013).

The development of shale reservoirs is a challenging task that requires the integration of different disciplines for an adequate evaluation. The geology of shale reservoirs has received great attention from the scientific community. Great efforts have been made on understanding geological aspects such as stratigraphy, sedimentation environments, geological structures, geochemistry and mineralogy, which is the basis for a comprehensive estimation of the potential for hydrocarbons in situ in a particular basin (Curtis 2002; Ross and Bustin 2008). Furthermore, the connectivity of naturally-fractured networks, which has major impact on the productivity of a reservoir, has also been addressed (Odling et al. 1999). Due to the complexity of shale reservoirs, which exhibit nonlinear adsorption/desorption processes, non-Darcy flows, complex flow geometry, and multi-scaled heterogeneity, theoretical models (King 1990), empirical models (Patzek et al. 2013), and numerical models (Wu et al. 2013) have been developed for the modeling and simulation of natural gas production from shale formations. In particular, the numerical models are of high interest in industry applications, since it is possible to link geologic characteristics of a reservoir with technical aspects of producing

wells. Economic models have been developed to study five main shale gas plays in Europe (Weijermars 2013). This study concludes that shale plays in Poland and Austria are strong candidates for future development; however, the author emphasizes the uncertainty associated with geological information and price variability. The effects of these parameters on the economic viability of a shale gas reservoir have been studied by Kaiser (2012a). Moreover, some studies present the impact of shale reservoirs on the economy on a regional scale (Kinnaman 2011; Munasib and Rickman 2015), national scale (Medlock 2012) and worldwide scale (Melikoglu 2014). These studies concluded that the development of these resources has a net positive effect on local economies and can reduce the dependence on importations and therefore strengthen the national energy security. Nonetheless, the sustainability of shale gas has been the subject of intense debate which has resulted in restrictions on the exploitation of these resources in some countries, such as Germany and France. Some authors state that shale gas is likely to have higher GHG footprint than conventional gas or even coal (Howarth et al. 2011a, b). However, it has been reported that GHG footprint of shale gas can be less than that of coal if fugitive methane emissions are controlled (Wang et al. 2011). Moreover, other works report that in power generation the GHG footprint of shale gas is less than half of the GHG footprint for coal, and roughly equal to conventional gas (Heath et al. 2014). By contrast, there seems to be a general consensus regarding the impact of shale gas in water resources. The consumption of local water resources and the high risk of underground drinking water and shallow aquifers contamination is one of the major concerns (Hou et al. 2012; Vengosh et al. 2014; Vidic et al. 2013).

Additional work has been focused on applications of optimization techniques in shale gas supply chains. A potential shale gas field is represented as a superstructure composed by a number of interconnected echelons such as well-pads, compressors, processing facilities and demand centers (Cafaro and Grossmann 2014). The trade-offs between environmental and economic performance of the shale gas supply chain have been addressed by means of multi-objective optimization techniques (Gao and You 2015a). As water management is considered a key issue in shale gas projects, some works have addressed the minimization of total water treatment related costs through MILP models with time-discrete representation (Yang et al. 2014). A different approach has been investigated in which the objective is defined as the maximization of profit per unit of fresh water consumption, the model is formulated as a mixed-integer linear fractional programming (MILFP) problem (Gao and You 2015b). In addition, a MILP model has been developed to address the optimal reuse of flowback water from hydraulic operations while considering fresh water seasonality, and economic and environmental aspects (Lira-Barragán et al. 2016b). Finally, the effects of uncertainty have been studied through two-stage stochastic MILFP models for the optimal design of shale gas supply chains (Gao and You 2015c). In addition, uncertainty has also been considered in the optimization of water management strategies (Lira-Barragán et al. 2016a).

3.2.1 Planning of Shale Gas Supply Chain: Overview

This section presents an overview of the issues that must be considered in the development of shale gas reservoirs, the design of the gas supply chain, and its integration with water management strategies. A generic shale gas supply chain superstructure is presented in Fig. 3.1.

Initially, fresh water is transported via trucks or pipelines from sources such as rivers and lakes to locations with high prospective production of shale gas. The transported water is mainly used to produce a fracturing fluid necessary for drilling and hydraulic fracturing operations, the latter activity being highly water-intensive as it accounts for about 86 % of the total water consumption over the life-cycle of a well-pad (Jiang et al. 2014). The fracturing fluid is composed mainly of water (≈90–95 vol.%), sand (≈4–9 vol.%), which is used as a proppant to maintain high permeability of the artificial fractures, and chemical additives (≈less than 1 vol.%). A well-pad can be defined as a cluster of single wells connected at the wellhead to a common point. In its most general form, the design of a well-pad can be described in terms of the number of wells, the horizontal length of each well, and the number of hydraulic fractures per well. The design of a well-pad is a critical aspect in the shale gas supply chain design and planning, since it affects the productivity of the well-pad and also the water requirements. During the early stage of the production of a well-pad, from 1 to 2 weeks, a fraction of the fracturing fluid returns to the wellhead. This fractions is highly variable, ranging between 10–40 % and it depends on geomechanical properties of the formation as well as on the composition of the

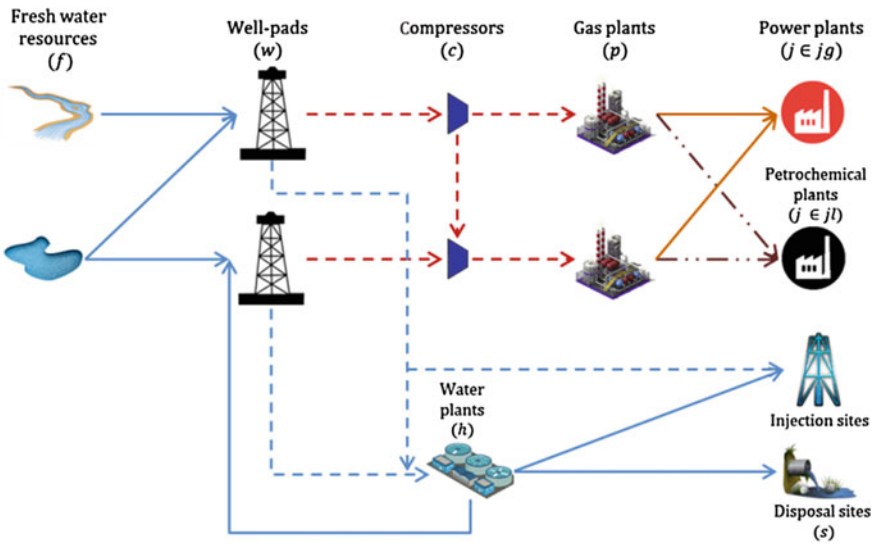

Fig. 3.1 Generic superstructure for shale gas supply chain (Reproduced from Guerra et al. (2016a))

fracturing fluid. This stream is known as flowback water and presents an average flow rate of 1,000 m^3/day. There can be additional production of water due to the presence of formation water. This stream is known as produced water and its flow rate is significantly lower than the flowback water, around 2–8 m^3/day. Total suspended solids (TSS) and total dissolved solids (TDS) are two important parameters for the characterization of the wastewater (flowback and produced water) associated with the production of shale gas (Vengosh et al. 2014). For flowback water, the TSS concentration varies from 1–500 mg/L and the concentration of TDSs ranges between 5,000 to 250,000 mg/L. The same ranges of TSS concentration apply to produced water, however, the TDS concentration varies between 10,000–336,000 mg/L (Fedotov et al. 2013). Concentration of TDS in flowback water increases with time, given that minerals and organic constituents present in the formation dissolve into the fracturing fluid (Gregory et al. 2011; Slutz et al. 2012). Given these characteristics, water management strategies clearly play an important role in dealing with the wastewater associated with shale gas production. According to the wastewater characterization and its final use, produced and flowback water can be sent either to water treatment plants, for primary and/or secondary treatment, or to deep-injection sites. Primary treatment processes only TSSs and re-uses the treated water in new well-pad locations provided that the concentration of TDS is low. Secondary treatment is required if the TSD's are higher than the specifications required for drilling and fracturing. In this case, the treated wastewater can be recycled to new well-pad locations or discharged into rivers. Finally, if the technology is available on-site, deep-injection is the most-preferred option as it avoids water treatment costs. However, if the injection point is located far from the reservoir, the trucking costs can be high enough to consider water treatment technologies instead. The concentration of TDS is one of the most important evaluation parameters for wastewater treatment economics and management strategy.

The composition of the produced shale gas depends on the geochemical characterization of the shale formation. Shale gas can be classified as dry gas (methane >90 %, with the rest largely CO_2 and N_2) or wet gas (methane, ethane, condensable fractions of propane, butane, iso-butane, CO_2, N_2). Usually, the composition of shale gas varies not only with location but also as the production progresses. The produced shale gas is sent to gas treatment facilities via pipelines either directly or through compressor stations. The gas is separated into different fractions and then the final products are sent to final customers, i.e. petrochemical plants, power stations, national gas pipeline network, etc.

3.2.2 Strategic Design and Planning of Shale Gas Supply Chains: Case Studies

The development of shale gas resources involve complex strategic and operational decisions, including the selection of sources of water for fracking processes,

selection of well-pads location and configuration, the design of gas and liquid pipelines network, technology, location, and capacity for wastewater treatment plants, and the location and design of gas processing plants. Accordingly, Guerra et al. (2016a) developed an optimization framework for the assessment of shale gas resources with water management from a supply chain perspective. The proposed framework can be simplified to a MILP model if gas composition is considered to be constant across the reservoir and constraints in water quality are simplified; the interested reader is referred to the original publication (Guerra et al. 2016a). The proposed framework takes into account different alternatives regarding fresh water supply and wastewater management strategies, as well as well-pad designs. The decisions tackled by the proposed model are summarized as follows:

Strategic decisions

- Well-pad designs
- Capacity and location for gas treatment plants and water treatment plants
- Water management strategies
- Capital investment decisions

Operational decisions

- Location and timing for drilling and fracturing a well-pad
- Water trucking
- Capacity of compressor stations and location
- Design of the gas pipeline network for transporting raw gas to processing facilities and final products to demand centers
- Amount of gas processed in each treatment facility

The mathematical framework presented in Guerra et al. (2016a) was used to study a prospective shale play in Colombia. Two case studies, Case A and Case B, will be presented in order to evaluate different aspects of the shale gas supply chain. Case A focuses on the optimal decisions related to the development of a shale reservoir, whereas the Case B is intended to reveal the repercussions of limitations on the availability of water resources on the design and economics of the shale gas supply chain. Both case studies are based on the same potential infrastructure for gas and water transportation and processing. The potential transport and processing infrastructures for gas and water supply chains (see Fig. 3.2) was generated using ArcGIS® 10.2 (ESRI 2014). Two well-pad designs were selected according to economic and environmental criteria. As an economic attractive well-pad design, labeled as "MaxNPV", the design is composed by 14 wells, with a horizontal length of 9,000 ft and fracture stages spaced every 200 ft. This design has high production levels of shale gas, but the demand of water for drilling and fracturing is also high resulting in high production of wastewater due to high production of flowback water. The second well-pad design is chosen based on environmental criteria in terms of consumption of water per total gas production. This design, labeled as "MinWI", is composed of 6 wells, with a horizontal length of 5,000 ft and fracture stages spaced every 200 ft. This design has less demand for water, as well as low

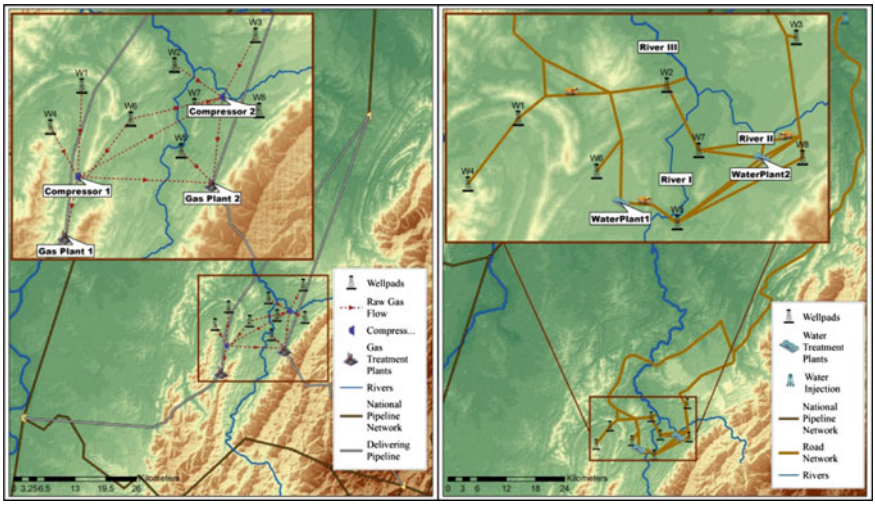

Fig. 3.2 Gas supply chain (*Left-hand side*) and water supply chain (*Right-hand side*) for a case study with 8 potential well-pads (Reproduced from Guerra et al. (2016a))

shale gas and wastewater production. The interested reader is referred to (Calderón et al. 2015) for further details on alternative well-pad designs and their relative merits. The methane fraction is delivered to gas-based power plants through the national gas pipeline system. One connection point in the north of the shale play is included for ethane injection. This point is indirectly connected to a petrochemical plant.

The potential infrastructure of the water supply chain was based on a road network connecting the different water sources with the demand points and the water treatment facility locations. The fresh water availability is subject to unimodal rainfall patterns, therefore a seasonal behaviour was taken into account. This was estimated based on hydrological balances carried out in ArcGIS. The hydrological balances incorporate historical data about precipitation, evapotranspiration, infiltration, and downstream water demand. There are two potential water treatment plants to be installed. Additionally, it is considered that one treatment plant operates with primary treatment and can process wastewater with maximum TDSs of 50,000 mg/l, whereas the other facility operates with secondary treatment and can process water with maximum TDSs of 120,000 mg/l. The infrastructure for both case studies is summarized as follows:

- 8 potential well-pad locations with 2 well-pad designs
- 2 compressor stations with 2 optional capacities
- 2 gas treatment plants with 2 optional capacities
- 3 demand centers located along the national pipeline network in Colombia
- 3 rivers supply fresh water for drilling and fracturing the well-pads

- The TDSs of the wastewater (produced and flowback water) were fixed between 34,300 and 106,700 mg/l
- 2 water treatment plants with 2 optional capacities.
- 2 disposal sites (discharge into rivers)
- 1 site for deep-injection of waste water
- Water trucking is the only transportation mode considered
- Gas methane prices vary from 4.1 up to 8.3 $/SCF along the planning horizon
- Gas composition was considered to be constant across the reservoir. Therefore, the optimization model can be simplified to an MILP
- 10 years as time horizon divided into 40 quarters.

The optimization problems were solved using GAMS 24.4.1. The MILP problem was solved with CPLEX 12.6.1 on a Dell OptiPlex 7010 with Intel® Core™ i7-3770 CPU @3.40 GHz and 16 GB RAM running Windows 7® Enterprise (64-bit operating system). The optimality gap was set to less than or equal to 1 %.

3.2.3 *Case Study Results and Discussion*

In this section we present the results of the optimization framework applied to the infrastructure described previously. In Case A, the total gas production was 473.9 billion standard cubic feet [BSCF] and the optimal NPV was $333.4 million, which corresponds to a net profit of 0.70 $/MMBtu. This margin could increase if more potential well-pads are considered for the exploitation of the play. The values for Capex, Opex, royalties and taxes are discounted to the first period and the total cost breakdown is presented in Fig. 3.3. A share of 66.3 % of the total costs corresponds to capital investments, followed by taxes with a share of 18.8 % and finally royalties and Opex with 9.1 % and 5.8 %, respectively. These results reflect a well-known fact of the shale gas industry, namely that the finances are dominated by the capital investment component rather than the operating costs. The breakeven

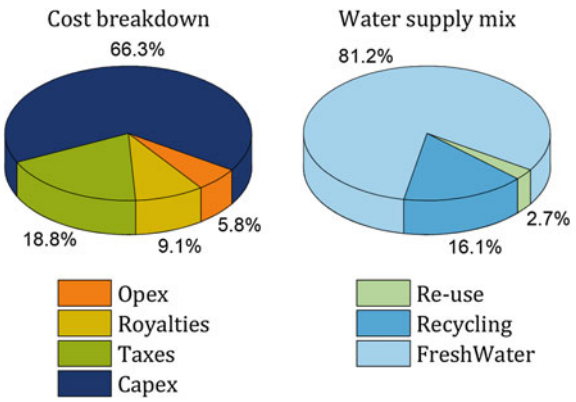

Fig. 3.3 Cost breakdown and water supply mix for case study A

gas price, defined here as the ratio between total expenditures (Capex plus Opex including water transportation cost) and total gas production, was found to be 3.86 $/MMBtu.

The drilling strategy consisted of 5 well-pads drilled and fractured using the MaxNPV design (well-pads W2, W3, W5, W7, and W8), and 3 well-pads were put in operation using the MinWI design (well-pads W1, W4 and W6). In total, 88 wells were drilled and fractured during the planning horizon. The wastewater from the well-pads W2 and W3 has low TDS concentration, below 50,000 mg/l. This situation allows higher water production, and therefore higher gas production, which explains the selection of MaxNPV designs, without affecting the technology selected for the wastewater treatment. The well-pad W4 produces wastewater with TDS of 104,000 mg/L, which explains the selection of a MinWI design in order to reduce the amount of produced water. However, the concentration of TDSs is not the only factor that influences the decision of the well-pad designs. The well-pads W5, W7 and W8, which use the MaxNPV design, produce wastewater with TDSs around 80,000 mg/L, 73,000 mg/L, and 52,000 mg/L, respectively. On the other hand, the well-pads W1 and W6, which use the MinWI design, produce wastewater with 53,000 and 61,000 mg/L, respectively. One would expect W5 and W7 to be selected to be MinWI designs, whereas W1 and W6 would be selected to be MaxNPV designs, in order to reduce the amount of wastewater that requires secondary treatment, and therefore reduce processing costs. The key to understand this lies in the relative position of the well-pads within the shale reservoir. From Fig. 3.2, it is evident that well-pads W5, W7, and W8 are comparatively closer not only to the delivery points of the produced gas but also to the fresh water sources than the well-pads W1 and W6. This is directly related to investment of pipelines and operating cost of gas and water transportation, which in this case is a determining factor.

The drilling schedule obtained has well-pad W5 drilled first in period 7; then well-pad W8 in period 8, W7 in period 10, W3 in period 11 and W2 in period 13. It is worth to note that the well-pads with design MaxNPV are installed first and then the well-pads W4, W1, and W6, with design MinWI, are installed consecutively in periods 14, 15, and 16. In this case, the results suggest that the east part of the reservoir should be fully developed before starting the development of the west part. The gas production initiates in period 8 with a production of 99.8 MMSCFD and quickly reaches a peak of 335.6 MMSCFD in period 14. After period 17, the gas production decreases steadily until the end of the planning horizon with a final production of 42.9 MMSCFD. The cumulative production indicates that at the end of the planning horizon, a total of 473.9 Bcf of raw gas were produced. The compressor 2, receiving the production from the well-pads with MaxNPV design, was installed with a capacity of 300 MMSCFD, whereas the compressor 1, receiving the production from well-pads with MinWI design, was installed with capacity of 150 MMSCFD. The design of the well-pads influences the selection of the capacity of the compressors. Only one gas treatment plant, gas plant 2, was installed with a capacity for processing 350 MMSCFD.

Regarding the water management decisions, both water treatment facilities were installed; water treatment plant 2, operating with secondary technology, was installed with a total capacity of ~1.8 million gallon/day while the water treatment plant 1, operating with primary technology, was installed with capacity for processing ~0.4 million gallon/day. In total, 2,420.2 million gallons are required to drill and fracture 8 well-pads. 1,964.5 million gallons (81.2 %) were supplied with fresh water, 65.1 million gallons (2.7 %) were supplied with wastewater after being processed with primary treatment, and 390.5 million gallon (16.1 %) were supplied with wastewater processed with second treatment. The water treatment facilities supply in total 455.6 million gallons of treated water for drilling and fracturing operations. The total production of wastewater associated with the shale gas production is around 949.4 million gallons. From the wastewater, 376.4 million gallons (39.6 %) are processed through primary treatment in water plant 1, 509.5 million gallons (53.7 %) are processed with secondary treatment in water plant 2, and only 63.5 million gallons (6.7 %) are sent to deep-injection. The global water intensity, defined here as the total water demand per total production, was estimated in 5.11 gallons/MMBtu.

A second case study, Case B, was undertaken based on the same infrastructure to evaluate the effects of limited availability of water resources on the economy of the supply chain and, thus, on the decisions regarding water management. In this case, the water availability was reduced by 30 % for all the rivers along the planning horizon. A summary of the results for Case B and a comparison with the base case study (Case A) are provided in Table 3.1. A reduction of 30 % in the availability of fresh water significantly impacts the decisions regarding the design and number of installed well-pads. In Case B, only 3 well-pads (W5, W7, and W8) were installed with MaxNPV configuration and 4 well-pads (W1, W2, W3, W4, and W6) were installed with MinWI configuration. This is reflected in the total gas production which was 319.0 BSCF, 32.6 % lower than in the previous case. The well-pads installed with MaxNPV are close to a treatment facility, which reduces capital and operating costs. The remarkable drop in production inevitably impacts the final NPV which for Case B is $117.7 million; corresponding to a reduction of 64.7 % in comparison to the Case A. The capital investment continues to be the most important component of the total costs, followed by taxes, royalties, and finally operating costs. The final breakeven costs is 4.05 $/MM Btu, 4.9 % higher than in Case A.

The decisions regarding the water management were also altered by the lower availability of fresh water. The total water consumption in Case B was 1,675.2 million gallon, 30.8 % lower than the consumption in Case A. The fresh water is still the main component to supply the water demand. In this case, 77.3 % of the total demand is supplied by rivers, 8.0 % is supplied by primary treatment facilities, contrasting with 2.7 % for Case A, and 14.7 % is supplied by secondary treatment facilities. The amount of water sent for deep injection dropped dramatically to 50,000 gallons/day. Finally, the water intensity index increased 2.8 % from 5.11 to 5.25 gallons/MMBtu. This increment is due to a reduction of productivity of the well-pads caused by limited operational times which are specified by the drilling

Table 3.1 Optimization results for case study B and comparison with base case

	70 % Water availability (Case B)	Base case (Case A)
MaxNPV	3	5
MinWI	4	3
NPV [Million $]	117.7	333.4
Capex [Million $]	1218.0	1682.5
Opex [Million $]	72.8	148.4
Royalties [Million $]	147.7	229.7
Taxes [Million $]	289.6	477.1
Total gas production [BSCF]	319.0	473.9
Breakeven cost [$/MM Btu]	4.05	3.86
Total water consumption [Million gallon]	1675.2	2420.5
Fresh water [Million gallon]	1294.4	1964.5
Re-use water [Million gallon]	134.8	65.1
Recycled water [Million gallon]	246.0	390.5
Injected water [Million gallon]	0.05	63.5
Water intensity [gal/MM Btu]	5.25	5.11

scheme. It is important to note that in scenarios with restricted access to water resources, the optimization framework opts for adjusting the water consumption by implementing lower water-intensive designs and decreasing the total number of well-pads that are drilled and fractured rather than increasing the capacity of water treatment facilities.

3.3 Interconnected Power Systems

Electricity, which is undoubtedly the most flexible and manageable energy form, has become one of the most important energy sources in recent decades. For instance, electricity is used for operating electronic devices and electric vehicles, as well as for heating, cooling, and lighting. Because the development and commercialization of the aforementioned devices and technologies have accelerated, electricity demand is currently growing even faster than overall primary energy supply (International Energy Agency 2013b; World Energy Council 2013). Moreover, given the early development of fossil power generation technologies, the power sector has become the leading source of energy-related CO_2 emissions, i.e. it accounts for roughly 42 % of total CO_2 emissions associated with the energy sector (International Energy Agency 2015). This has motivated the development and integration of renewable power generation technologies, namely, wind turbines and solar panels, within power systems. However, the intermittency of wind and solar sources imposes new challenges to the design and operation of power systems. Specifically, the integration of the aforementioned renewable power sources has

negative implications on the reliability of the power system. Consequently, the design and planning of power systems—arrangement of electrical devices, i.e. power plants, transformers, and transmission lines—used to generate, transmit, and supply electricity—have captured the attention of the research community.

An appropriate electricity demand forecast is critical for the adequate design and planning of power systems. Electricity demand can be described, from a nationwide, regional, or statewide viewpoint, as function of economic variables, i.e. gross domestic product (GDP), weather and climate conditions, i.e. ambient temperature, and demographic variables, i.e. population growth. A review of forecasting approaches for electricity demand including the econometric, end-use, and time series, is presented by Suganthi and Samuel (2012). Regarding the power generation expansion problem, different optimization models have been developed and applied for the systematic selection of capacity, technology, location, and time for construction for new power generation plants. These models are mostly MILP (Koltsaklis et al. 2014; Unsihuay-Vila et al. 2011) and some of them include operational constraints (Bakirtzis et al. 2012). Similarly, based either on the alternating current (AC) or the direct current (DC) formulation, mixed integer nonlinear programming (MINLP) and MILP models have been developed to support decisions on selective investment in transmission assets in power systems (Alguacil et al. 2003; Lumbreras et al. 2014; Rider et al. 2007; Vinasco et al. 2014). The MINLP model for transmission planning (Rider et al. 2007), which is based on the AC formulation, is more accurate, i.e. it takes into account both the active and reactive part of a power system. However, this model also imposes significant computational challenges even for finding local solutions to the optimization problem. By contrast to an MINLP model, MILP models are more efficient from a computational viewpoint but at the cost of an approximate representation of the original problem which may give rise to less accurate solutions. A number of authors have been focused on the integration of power generation and transmission planning models (Guerra, et al. 2016b; Roh et al. 2007; Sharan and Balasubramanian 2012). Given the complexity of the problem, the approaches proposed for the integrated planning of power generation and transmission are mainly based on the DC formulation of the transmission expansion planning problem.

Besides the power generation and transmission planning problems, the reliability of power systems and their role in global warming are important aspects that have also received attention (Koltsaklis et al. 2014; Unsihuay-Vila et al. 2011; Xiao et al. 2011). Reliability can be defined as the system's availability to match the electricity demand under both normal and abnormal circumstances, including variations on demand, losses of transmission lines, and device outage that could generate negative impacts on the performance of the system. This aspect is particularly important from an operational point of view and especially in systems with a relatively high penetration of intermittent renewable energy sources. Regarding the environmental impacts, the scientific community has studied different strategies for mitigation of CO_2 emissions for power systems. These strategies can be classified into three different groups, as follows: (1) integration of renewable energy sources, where energy storage seems to be the key for a higher penetration of renewable

energy sources (Dunn et al. 2011; Gençer et al. 2015; Hart and Jacobson 2012; Luo et al. 2015). Regarding energy storage, high storage capacity as well as low cost remain as major barriers. (2) use of carbon capture and sequestration (CCS) and carbon capture and utilization (CCU) technologies for fossil power plants (Cuéllar-Franca and Azapagic 2014; De Coninck and Benson 2014; Guerra et al. 2015; Jenkins et al. 2012; Szulczewski et al. 2012). The success of these technologies could be determined by low coal prices and carbon taxes. (3) implementation of demand side management (DSM) strategies, i.e. demand response and energy efficiency (Choi and Thomas 2012; Kirschen 2003; Mohsenian-Rad et al. 2010; Strbac 2008). Undoubtedly, this strategy is the most cost-effective alternative for CO_2 emissions mitigation in power systems. It offers environmental benefits as well as economic and reliability gains. Moreover, the relationship of climate change and abnormal weather patterns, i.e. El Niño and La Niña, with the design and planning of power systems has also been studied (Bartos and Chester 2015; Cai et al. 2014, 2015). For instance, (Bartos and Chester 2015) analyzed the impacts of climate change on the electricity supply in the Western United States. They found that climate change may reduce summertime power capacity for vulnerable power stations in a significant way. Lastly, some studies have focused on the development and application of optimization under uncertainty techniques to the design and planning of power systems, including two-stage stochastic optimization approaches (Ahmed et al. 2014; Feng and Ryan 2013; Jin et al. 2011) and dynamic stochastic fractional programming (DSFP) models (Zhu and Huang 2013).

In the following section, an application of a novel optimization framework developed by Guerra et al. (2016b) for the integrated design and planning of power generation and transmission expansion in interconnected power systems is presented. First, based on a general view of interconnected power systems, the general optimization problem is presented and a brief description of the aforementioned framework is provided. Then, two case studies are addressed in order to illustrate the applicability of mathematical programing tools at dealing with two typical situations than can arise when designing and planning power systems.

3.3.1 Planning of Integrated Power Generation and Transmission: Problem Statement and Solution Approach

Interconnected power systems are networks of power generation and transmission assets that convert primary energy sources into electricity to supply demand in different interconnected regions (see Fig. 3.4). From a modeling view, each region can be represented by a subset of existing, planned, and potential power plants, as well as a specific availability and cost of primary energy sources. Additionally, each region has a specific electricity demand and interregional electricity import/export capacity. Therefore, regional electricity demand can be met either by local power

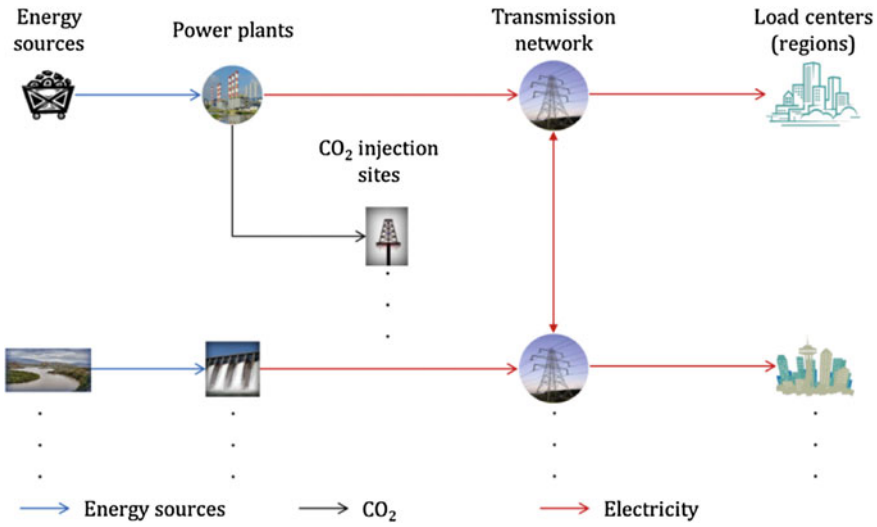

Fig. 3.4 Generic superstructure for interconnected power systems (Reproduced from (Guerra et al. 2016b))

generation facilities or by a combination of local generation and imports from other regions. Furthermore, each power plant can be characterized by its location, capacity, capital cost, operating cost, heat rate (fossil power plants only), fuel cost, availability and capacity factors, and CO_2 emission factors. Reserve margin and spinning reserve constraints are taken into account in order to guarantee the safe and reliable operation of the power system at any time. The strategic design and tactical planning of such systems (i.e. on a nationwide scale) deals with the simultaneous optimization of power generation and transmission capacity expansion for a given electricity demand projection in different regions over a long-term planning horizon (i.e. 15 years). Indeed, the optimization framework developed recently by Guerra et al. (2016b) was demonstrated to be suitable for supporting the strategic and tactical decisions associated with the long-term planning of integrated power systems. This framework addresses different type of decisions, as follows:

Strategic decisions

- Technology, capacity and location for power generation expansion
- Timing for the power generation and transmission expansions
- Investment in CO_2 mitigation strategies (if CO_2 emission constraints are considered)

Operational decisions

- Electricity allocation
- Non-Spinning and spinning reserve allocation
- Power transmission flows

Additionally, the framework minimizes the total cost associated with the operation of interconnected power systems over a given long-term horizon from the governmental perspective, taking into account the investments required to balance future electricity demand. The framework also includes the following aspects: (1) reserve constraints, (2) CO_2 emission constraints, (3) CO_2 emission mitigation strategies, i.e. the penetration of renewable energies sources, the integration of CCS technologies, and the adoption of DSM strategies. Details of the mathematical formulation of the framework are omitted here, the interested reader is referred to the original publication (Guerra et al. 2016b). The framework was applied to study two different scenarios, "business as usual" and "CO_2 mitigation policy". In the "business as usual" scenario, CO_2 mitigation strategies are not contemplated. By contrast, in the "CO_2 mitigation policy" scenario, constraints for CO_2 emissions are considered and thus different mitigation strategies are considered. The mitigation strategies include the integration of renewable energies, CCS retrofitting for fossil power plants and DSM strategies. This study is based on the "business as usual" scenario, which was modified in order to consider penalties for unserved energy that were not considered in the original formulation. The implementation of the framework to a real world power system is described in the following section.

3.3.2 Integrated Planning of Generation and Transmission Planning: Outline of Case Studies

Two case studies are presented in this section in order to illustrate the application of aforementioned optimization framework to assist the decision making processes in the long-term planning of power systems. In the first case study (Case A) abnormal weather patterns, i.e. El Niño and La Niña, are not considered and thus annual variability of capacity factors for hydroelectric power plants is not taken into account. By contrast, in the second case study (Case B) both El Niño and La Niña patterns are included and capacity factors for hydropower plants are varied accordingly. It is important to note that El Niño causes droughts in Colombia, while La Niña generates intense and abundant rainfalls. Therefore, a variation of -5 % in the capacity factors for hydropower was assigned to El Niño phenomenon, while a change of +5 % was considered for La Niña events. Furthermore, the frequency of each phenomenon was set to 5 years with 2 years of span between each event. The El Niño event takes place in 2016, 2021, and 2026, whereas La Niña appears in 2018, 2023, and 2028. The case studies are based on the infrastructure of the Colombian interconnected power system (see Fig. 3.5), which is described in details in Guerra et al. (2016b). Additionally, a 15-year planning horizon is considered, from 2015 to 2029. The main features of this power system are summarized below.

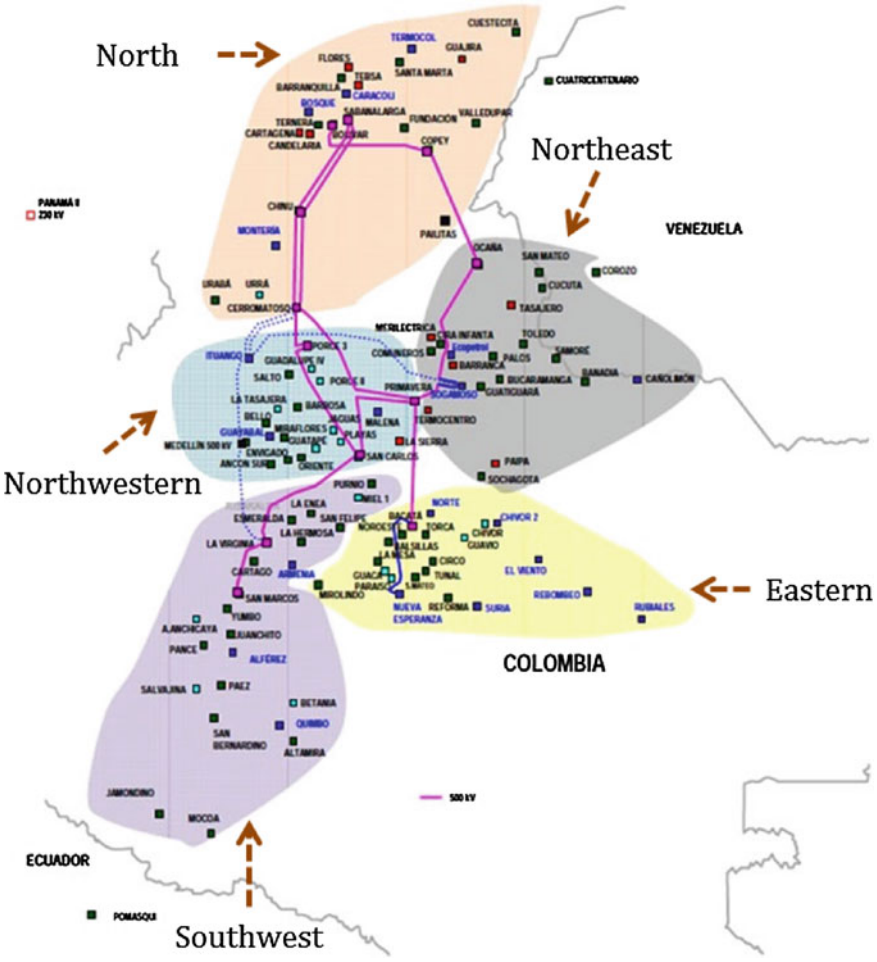

Fig. 3.5 Colombian interconnected power system (Reproduced from (Guerra et al. 2016b))

- 5 interconnected regions (See Fig. 3.5)
- Transmission network: ~2,646 km of 500 kV transmission lines (continuous pink lines in Fig. 3.5) and ~11,654 km of 220 kV transmission lines.
- Installed power plants: 45 installed (26 hydro, 15 natural gas, and 4 coal)
- Installed capacity: ~15,680 MW
- Planned power plants: 14 power plants that represent about ~4,054 MW of additional capacity
- Total electricity demand: ~64.44 TWh for 2015, average increase ~2.82 % per year.
- Demand curve: 2 h/day peak load block and 22 h/day base load block

- Maximum power demand: ~9,885 MW for 2015 and increase almost linearly to 12,953 MW in 2029
- Reserve requirements: non-spinning reserves ~15 % of the peak power demand and spinning reserve margin ~5 % of current electricity demand
- Fuel cost: coal between 1.86 $/MM Btu and 4.99$/MM Btu and natural gas between 3.97 $/MM Btu to 14.47$/MM Btu

Besides the power generation capacity associated with existing and already planned power plants, additional expansions of the generation and transmission capacity will be necessary in order to offset the growth in electricity demand along the planning horizon. Thus, potential new power plants, which have a capacity that ranges from 200 MW to 500 MW, are considered in both case studies. The new power plants include hydroelectric, natural gas, and coal-based power generation plants. Each power plant has a given generation capacity and geographic location. The capital cost for new power plants was estimated to be around $2,242/kWh, $2,058/kWh, and $1,023/kWh for hydroelectric, coal and natural gas power plants, respectively. These capital costs were based on information from the IEA (International Energy Agency/Nuclear Energy Agency 2010, 2015) and adjusted to the Colombian context. Additionally, capacity and availability factors, heat rate, and emission factors were based on information from UPME (Unidad de Planeación Minero Energética—UPME 2014) and XM (2015a, b). Furthermore, it is considered that transmission capacity can be expanded either by installing new transmission circuits and repowering existing transmission connection. The capital cost for adding new circuits was estimated to be around $489,737/km, whereas the capital cost for repowering was defined as a 50 % of the cost for adding new circuits. Additional details regarding the input information required by the optimization framework can be found in Guerra et al. (2016b).

The implementation of the case studies resulted into two MILP optimization models that were solved using GAMS 24.4.1 with the solver CPLEX 12.6.1 on a Dell OptiPlex 7010 with Intel® Core™ i7-3770 CPU @3.40 GHz and 16 GB RAM running Windows 7® Enterprise (64-bit operating system). The relative gap (in GAMS notation OPTCR) was fixed at 0 % for both case studies.

3.3.3 Results and Discussion

This section summarizes the results and the corresponding analysis for the two case studies described in the previous section. The minimum total cost for Case A was about $17.47 billion, with investments in the expansion of power generation and transmission accounting for the highest share, i.e. 38.3 % of total cost. It is important to clarify that this cost share takes into account capital investments associated with already planned power plants and that the optimization model incorporates only the operational decisions associated with this type of power plant. Fuel cost, associated with fossil power plants, also represents a significant share of

total cost, i.e. 33.2 %. The remaining 28.5 % of total cost corresponds to 11.4 %, 9.1 %, and 8.1 % for transmission cost, fixed O&M (operation and maintenance) cost, and variable O&M cost, respectively. The total investment in the expansion of the power generation capacity results in around 5,654.4 MW of additional capacity, with investments in hydroelectric, coal and, natural gas power plants accounting for a share of 73.9 %, 21.1 %, and 5.0 % of the total investment, respectively. The schedule of the capacity expansions for power generation indicated that a new 200 MW gas power plant located in the North region (see Fig. 3.5) should be installed and operated from 2018 to 2029. This early expansion seems to be driven by the rise of electricity demand in the North region and the transmission capacity constraints associated with electricity imports from the two neighboring regions, i.e. Northwestern and Northeast regions. A hydropower plant with 500 MW of capacity and located in the Northwestern region needs to be installed and operated in order to provide electricity from 2026 to 2029. Subsequently, farther expansions based on coal power generation are required. For instance, a new coal power plant with 250 MW of capacity should be installed in the Eastern region and used to generate electricity from 2027 to 2029. An additional 350 MW coal power plant needs to be installed in the North region and operated in 2028 and 2029. Moreover, a natural gas power plant should be installed in the Southwest region and used to offset electricity demand by 2029. Transmission capacity expansion includes two new transmission lines that must be installed and one existing transmission line that must be repowered. A new transmission line, connecting Northwestern and Northeast regions, should be installed in 2019 and used from 2020 to 2029. An additional transmission line needs to be built in 2022 and used from 2023 to 2029. This new transmission line should be installed between the Northwestern and Southwest regions. Furthermore, the existing transmission line between the Northwestern and Eastern regions should be repowered in 2025 in order to increase the capacity for power transmission between the corresponding regions from 2026 to 2029. It is important to note that the transmission expansion plan is focused on increasing power transmission capacity from the region with the highest installed power generation capacity, (the Northwestern region) to other regions. As a reference, the Northwestern region accounts for about 35.1 % of the total installed power generation capacity in 2015.

The optimal electricity generation plan for Case A is presented in Fig. 3.6. The power system relies mostly on power generation from hydropower plants, with roughly 74.9 % of total electricity demand in the entire planning horizon provided by hydroelectric power plants. The remaining 25.1 % is provided by fossil power plants. Electricity supply from hydropower plants reaches a peak of 64.9 TWh in 2026. Similarly, electricity supply from gas and coal power plants reaches a peak of 13.0 and 14.6 TWh by the end of the time horizon, respectively. Total cumulative CO_2 emissions associated with power generation from fossil power plants was found to be about 237.6 million t CO_2. Finally, the average electricity generation cost for planned projects and new power plants was roughly 40.9 $/MWh.

For Case B, in which a penalty for unserved electricity demand is considered, the minimum total cost was almost the same as for Case A, i.e. $17.47 billion.

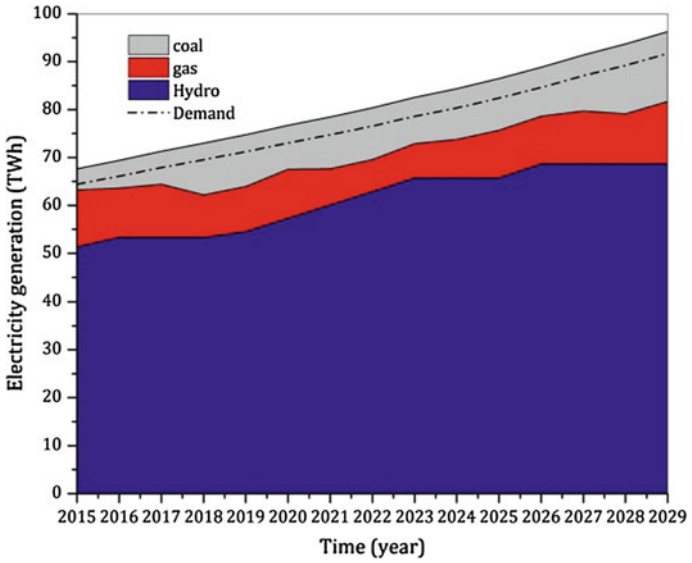

Fig. 3.6 Electricity generation for Case A

However, the total cost associated with capital investment and operational expenditures was about $17.09 billion and the total cost for unserved demand was around $0.38 billion. The decrease in total cost, without taking into account the penalty for unserved demand, was mainly associated with a decrease in capital investment in the expansion of power generation and transmission capacity. For instance, the share of capital investment in the total cost was found to be 35.8 % for Case B, which represents a decrease of about 2.5 % when compare with Case A. By contrast, the share of fuel cost for Case B was about 35.4 %, an increase of roughly 2.2 %. Additionally, the share of transmission cost, fixed O&M cost, and variable O&M cost was found to be 11.3 %, 9.3 %, and 8.3 %, respectively. For Case B, a total of 4,904.4 MW of additional power generation capacity is required in order to offset the electricity demand and provide the reserve requirements. This expansion represents a decrease of about 13.3 % in the expansion of the power generation capacity when compare with Case A. Investments in hydroelectric power plants account for a share of 68.4 % of the total investment in power generation assets, which represents a decrease of about 5.5 % when compared with Case A. Additionally, investments in coal and natural gas power plants represent a share of 25.6 % and 6.0 %, respectively. The schedule of the capacity expansion for power generation indicates that a re-schedule is needed for Case B. For instance, the 200 MW gas power plant located in the North region is now installed in 2017 and operated from 2019 to 2029. In other words, this investment is postponed for 1 year. Moreover, the 500 MW hydropower plant located in the Northwestern region and the 250 MW coal power plant located in the Eastern region are not selected in Case B. On the other hand, the coal power plant with 350 MW of

capacity and the 300 MW gas power plant located in the North and the Southwest regions, respectively, are both installed in 2024 and operated from 2026 to 2029. The plan for the expansion of the transmission capacity is the same except that the existing transmission line between Northwestern and Eastern regions is not repowered.

The expansion plan for power generation and transmission in Case B has a direct impact on the optimal electricity generation plan, which is shown in Fig. 3.7. For instance, hydropower plants supply about 70.1 % of total electricity demand, which represents a reduction of about 4.8 % when compared with Case A. Moreover, the power system is not able to offset the electricity demand in three periods, 2016, 2026, and 2029. Additionally, electricity supply from hydropower plants reaches a peak of 64.3 TWh in 2023 instead of 2026, as in Case A. Similarly to Case A, electricity supply from gas and coal power plants reaches a peak of 15.2 TWh and 13.4 TWh, respectively, by the end of the planning horizon. It is important to note that the electricity supply peak for gas increases by 2.2 TWh, while the peak for coal decreases by 1.6 TWh when Case B is compared with Case A. Given the net increase in electricity supply from fossil power plants, total CO_2 emissions increase by 2.1 % (roughly 5 million t of CO_2), from Case A to Case B. Finally, based on the two cases described above, it can be concluded that abnormal weather patterns, i.e. El Niño and La Niña, could have an important impact on the decision-making process in power systems that rely mostly on hydropower generation. Specifically, it was demonstrated that under El Niño/La Niña events, investment in hydropower plants are reduced and generation from fossil power plant is increased in order to balance the electricity demand. Therefore, CO_2 emissions also increase as a

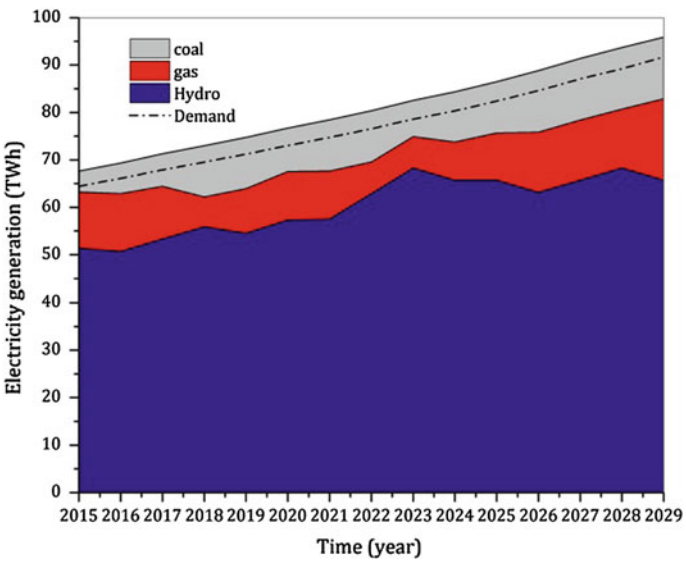

Fig. 3.7 Electricity generation for Case B

consequence of the higher variability of hydropower generation and the increase in power generation from fossil power plants. These results have also important implications from a policy perspective. For instance, under El Niño/La Niña events, hydropower generation may not be a cost-effective alternative to curb carbon emissions.

3.4 Concluding Remarks

Energy systems involve high complexity since they encompass the integration of multiple components and their operation has both environmental and societal impacts. Decision-making processes in the energy sector are complex tasks that require the support of modeling, simulation, and optimization tools. The case studies presented in this chapter illustrate how optimization tools, e.g. MILP models, can be used for supporting the strategic design and the tactical planning of energy systems. Moreover, it was demonstrated that the use of these tools can reveal synergies and trade-offs within the system, that otherwise will remain obscured. For example, it was found that TDS concentration in wastewater is an important decision factor for the selection of the drilling and fracturing scheme in shale gas supply chains. Additionally, it was shown that in scenarios under water scarcity, it is preferable to re-adjust the configuration of the well-pad in order to reduce total water consumption instead of investing in larger wastewater treatment infrastructures. Thus, decisions regarding the design of the well-pad must be taken into consideration when designing shale gas supply chains. Similarly, it was found that El Niño and La Niña events can have significant and apparently counterintuitive consequences for the design and planning of power systems. For instance, under the aforementioned abnormal weather patterns, investments in hydropower plants are decreased given the higher variability of hydropower resources. Consequently, electricity supply from fossil power plants is increased and so the CO_2 emissions increase significantly. Similarly, investments in transmission assets are also reduced. This situation reinforces the importance of addressing simultaneously the power generation and transmission expansion problems and thus the relevance of the proposed integrated approach.

From a technical point of view, there are challenges and opportunities in the application of mathematical programing tools to assist in the strategic design and tactical planning of energy supply chains. These include the development of high-fidelity optimization models while maintaining the computational tractability of these models. In general, MILP formulations are preferable since their solution usually require reasonable computational effort. Nonetheless, MILP problems are not always the most adequate approach to model certain phenomena. In shale gas supply chains, the main challenges stem from handling spatial and temporal variations in both gas composition and TDS concentration in wastewater streams, and addressing uncertainty in a computational efficient manner. The consideration of spatial and temporal variations in shale gas supply chain problems lead

automatically to MINLP formulations, which are more challenging optimization problems from a computational viewpoint. Likewise, the accuracy of the optimization model for the integrated planning of power generation and transmission can be improved by using the AC form of the transmission constraints at the expense of using an MINLP formulation. Also, effectively handing uncertainties in the design and planning of power systems remains an open field.

Given the complexity of the optimization models used to assist the design and planning of energy supply chain problems, global sensitivity analysis can be used to identify the key uncertain parameters -the parameters with an important impact on the optimal solution—which could allow a more efficient modeling of the uncertainty space. Moreover, given the fact that the aforementioned problems involve strategic and operational decisions taken in different stages of the planning horizon, multi-stage stochastic approaches could more realistically capture not only the impact of uncertainty but also the additional degree of freedom provided by recourse actions taken once uncertainties are realized. For instance, as the next step beyond deterministic models, two-stage stochastic formulations can be useful in handling parameter uncertainty in design and planning problems associated with energy supply chain systems. Finally, the integration of the development of shale gas resources along with the planning and design of power systems is an open area that offers interesting challenges and opportunities to the mathematical programming community.

References

Ahmed, S., Elsholkami, M., Elkamel, A., Du, J., Ydstie, E. B., & Douglas, P. L. (2014). Financial risk management for new technology integration in energy planning under uncertainty. *Applied Energy, 128*, 75–81. doi:10.1016/j.apenergy.2014.03.058.

Alguacil, N., Motto, A. L., & Conejo, A. J. (2003). Transmission expansion planning: A mixed-integer LP approach. *IEEE Transactions on Power Systems, 18*, 1070–1077. doi:10.1109/TPWRS.2003.814891.

Bakirtzis, G. A., Biskas, P. N., & Chatziathanasiou, V. (2012). Generation expansion planning by MILP considering mid-term scheduling decisions. *Electric Power Systems Research*, *86*, 98–112. doi:10.1016/j.epsr.2011.12.008.

Bartos, M. D., & Chester, M. V. (2015). Impacts of climate change on electric power supply in the Western United States. *Nature Climate Change*, 1–5. doi:10.1038/nclimate2648.

BP. (2015). BP Energy Outlook 2035.

Cafaro, D., & Grossmann, I. (2014). Strategic planning, design and development of the shale gas supply chain network. *AIChE, 60*, 2122–2142. doi:10.1002/aic.

Cai, W., Borlace, S., Lengaigne, M., van Rensch, P., Collins, M., Vecchi, G., et al. (2014). Increasing frequency of extreme El Niño events due to greenhouse warming. *Nature Climate Change, 5*, 1–6. doi:10.1038/nclimate2100.

Cai, W., Wang, G., Santoso, A., McPhaden, M. J., Wu, L., Jin, F.-F., et al. (2015). Increased frequency of extreme LaNiña events under greenhouse warming. *Nature Climate Change, 5*, 132–137. doi:10.1038/nclimate2492.

Calderón, A. J., Guerra, O. J., Papageorgiou, L. G., Siirola, J. J., & Reklaitis, G. V. (2015). Preliminary evaluation of shale gas reservoirs: Appraisal of different well-pad designs via

performance metrics. *Industrial and Engineering Chemistry Research, 54*, 10334–10349. doi:10.1021/acs.iecr.5b01590.

Choi, D. G., & Thomas, V. M. (2012). An electricity generation planning model incorporating demand response. *Energy Policy, 42*, 429–441. doi:10.1016/j.enpol.2011.12.008.

Clarkson, C. R. (2013). Production data analysis of unconventional gas wells: Workflow. *International Journal of Coal Geology, 109–110*, 147–157. doi:10.1016/j.coal.2012.11.016.

Cuéllar-Franca, R. M., & Azapagic, A. (2014). Carbon capture, storage and utilisation technologies: A critical analysis and comparison of their life cycle environmental impacts. *Journal of CO_2 Utilization*, *9*, 82–102. doi:10.1016/j.jcou.2014.12.001.

Curtis, J. B. (2002). Fractured shale-gas systems. *AAPG Bulletin, 86*, 1921–1938. doi:10.1306/61EEDDBE-173E-11D7-8645000102C1865D.

De Coninck, H., & Benson, S. M. (2014). Carbon dioxide capture and storage: Issues and prospects. *Annual Review of Environment and Resources, 39*, 243–270. doi:10.1146/annurev-environ-032112-095222.

Dunn, B., Kamath, H., & Tarascon, J.-M. (2011). Electrical energy storage for the grid: A battery of choices. *Science, 334*, 928–935. doi:10.1126/science.1212741.

Eaton, T. T. (2013). Science-based decision-making on complex issues: Marcellus shale gas hydrofracking and New York City water supply. *The Science of the Total Environment, 461–462*, 158–169. doi:10.1016/j.scitotenv.2013.04.093.

ESRI. (2014). ArcGIS Desktop: Release 10.2.2. Redlands, CA: Environmental Systems Research Institute.

ExxonMobil. (2014). The Outlook for Energy: A View to 2040.

Fedotov, V., Gallo, D., Hagemeijer, P., & Kuijvenhoven, C. (2013). Water management approach for shale operations in North America. *SPE Unconventional Resources Conference and Exhibition-Asia Pacific*. Society of Petroleum Engineers. doi:10.2118/167057-MS.

Feng, Y., & Ryan, S. M. (2013). Scenario construction and reduction applied to stochastic power generation expansion planning. *Computers & Operations Research, 40*, 9–23. doi:10.1016/j.cor.2012.05.005.

Gao, J., & You, F. (2015a). Shale gas supply chain design and operations toward better economic and life cycle environmental performance: MINLP model and global optimization algorithm. *ACS Sustainable Chemistry & Engineering, 3*, 1282–1291. doi:10.1021/acssuschemeng.5b00122.

Gao, J., & You, F. (2015b). Optimal design and operations of supply chain networks for water management in shale gas production: MILFP model and algorithms for the water-energy nexus. *AIChE Journal, 61*, 1184–1208. doi:10.1002/aic.14705.

Gao, J., & You, F. (2015c). Deciphering and handling uncertainty in shale gas supply chain design and optimization: Novel modeling framework and computationally efficient solution algorithm. *AIChE Journal, 61*, 3739–3755. doi:10.1002/aic.15032.

Gençer, E., Mallapragada, D. S., Maréchal, F., Tawarmalani, M., & Agrawal, R. (2015). Round-the-clock power supply and a sustainable economy via synergistic integration of solar thermal power and hydrogen processes. *Proceedings of the National Academy of Sciences 201513488*. doi:10.1073/pnas.1513488112.

Gregory, K. B., Vidic, R. D., & Dzombak, D. A. (2011). Water management challenges associated with the production of shale gas by hydraulic fracturing. *Elements*, *7*, 181–186. doi:10.2113/gselements.7.3.181.

Guerra, O. J., Calderón, A. J., Papageorgiou, L. G., Siirola, J. J., & Reklaitis, G. V. (2016a). An optimization framework for the integration of water management and shale gas supply chain design. *Computers & Chemical Engineering, 92*, 230–255. doi: 10.1016/j.compchemeng.2016.03.025.

Guerra, O. J., Tejada, D. A., & Reklaitis, G. V. (2016b). An optimization framework for the integrated planning of generation and transmission expansion in interconnected power systems. *Applied Energy, 170*, 1–21. doi:10.1016/j.apenergy.2016.02.014.

Guerra, O. J., Tejada, D. A., Rodríguez, R., & Reklaitis, G. V. (2015). A spatial multi-period Mixed Integer Linear Programming (MILP) model for optimal power planning: CO_2 emissions

mitigation. In *12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering* (pp. 2345–2350). doi:10.1016/B978-0-444-63576-1.50085-6.

Hart, E. K., & Jacobson, M. Z. (2012). The carbon abatement potential of high penetration intermittent renewables. *Energy & Environmental Science,*. doi:10.1039/c2ee03490e.

Heath, G. A., O'Donoughue, P., Arent, D. J., & Bazilian, M. (2014). Harmonization of initial estimates of shale gas life cycle greenhouse gas emissions for electric power generation. *Proceedings of the National Academy of Sciences* 111, E3167–E3176. doi:10.1073/pnas.1309334111.

Hou, D., Luo, J., & Al-Tabbaa, A. (2012). Shale gas can be a double-edged sword for climate change. *Nature Climate Change, 2*, 385–387. doi:10.1038/nclimate1500.

Howarth, R. W., Ingraffea, A., & Engelder, T. (2011a). Natural gas: Should fracking stop? *Nature, 477*, 271–275. doi:10.1038/477271a.

Howarth, R. W., Santoro, R., & Ingraffea, A. (2011b). Methane and the greenhouse-gas footprint of natural gas from shale formations. *Climatic Change, 106*, 679–690. doi:10.1007/s10584-011-0061-5.

International Energy Agency. (2013a). CO_2 Emissions From Fuel Combustion: Highlights 2014. Paris.

International Energy Agency. (2013b). World Energy Outlook 2014: Executive summary. Paris.

International Energy Agency. (2015). World Energy Outlook Special Report 2015: Energy and Climate Change. Paris.

International Energy Agency/Nuclear Energy Agency. (2010). Projected Costs of Generating Electricity—2010 Edition. Paris.

International Energy Agency/Nuclear Energy Agency. (2015). Projected Costs of Generating Electricity—2015 Edition. Paris.

Jenkins, C. R., Cook, P. J., Ennis-King, J., Undershultz, J., Boreham, C., Dance, T., et al. (2012). Safe storage and effective monitoring of CO_2 in depleted gas fields. *Proceedings of the National Academy of Sciences,*. doi:10.1073/pnas.1107255108.

Jenner, S., & Lamadrid, A. J. (2013). Shale gas vs. coal: Policy implications from environmental impact comparisons of shale gas, conventional gas, and coal on air, water, and land in the United States. *Energy Policy, 53*, 442–453. doi:10.1016/j.enpol.2012.11.010.

Jiang, M., Hendrickson, C. T., & VanBriesen, J. M. (2014). Life cycle water consumption and wastewater generation impacts of a marcellus shale gas well. *Environmental Science and Technology, 48*, 1911–1920. doi:10.1021/es4047654.

Jin, S., Ryan, S. M., Watson, J.-P., & Woodruff, D. L. (2011). Modeling and solving a large-scale generation expansion planning problem under uncertainty. *Energy Systems,*. doi:10.1007/s12667-011-0042-9.

Kaiser, M. J. (2012a). Profitability assessment of Haynesville shale gas wells. *Energy, 38*, 315–330. doi:10.1016/j.energy.2011.11.057.

Kaiser, M. J. (2012b). Haynesville shale play economic analysis. *Journal of Petroleum Science and Engineering, 82–83*, 75–89. doi:10.1016/j.petrol.2011.12.029.

King, G. R. (1990). Material balance techniques for coal seam and devonian shale gas reservoirs. *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers (pp. 181–192). doi:10.2118/20730-MS.

Kinnaman, T. C. (2011). The economic impact of shale gas extraction: A review of existing studies. *Ecological Economics, 70*, 1243–1249. doi:10.1016/j.ecolecon.2011.02.005.

Kirschen, D. S. (2003). Demand-side view of electricity markets. *IEEE Transactions on Power Systems, 18*, 520–527. doi:10.1109/TPWRS.2003.810692.

Koltsaklis, N. E., Dagoumas, A. S., Kopanos, G. M., Pistikopoulos, E. N., & Georgiadis, M. C. (2014). A spatial multi-period long-term energy planning model: A case study of the Greek power system. *Applied Energy, 115*, 456–482. doi:10.1016/j.apenergy.2013.10.042.

Lira-Barragán, L. F., Ponce-Ortega, J. M., Guillén-Gosálbez, G., & El-Halwagi, M. M. (2016a). Optimal water management under uncertainty for shale gas production. *Industrial and Engineering Chemistry Research, 55*, 1322–1335. doi:10.1021/acs.iecr.5b02748.

Lira-Barragán, L. F., Ponce-Ortega, J. M., Serna-González, M., & El-Halwagi, M. M. (2016b). Optimal reuse of flowback wastewater in hydraulic fracturing including seasonal and environmental constraints. *AIChE Journal*, *7*. doi:10.1002/aic.15167.

Lumbreras, S., Ramos, A., & Sánchez, P. (2014). Automatic selection of candidate investments for Transmission Expansion Planning. *International Journal of Electrical Power & Energy Systems*, *59*, 130–140. doi:10.1016/j.ijepes.2014.02.016.

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. *Applied Energy, 137*, 511–536. doi:10.1016/j.apenergy.2014.09.081.

Medlock, K. B. (2012). Modeling the implications of expanded US shale gas production. *Energy Strategy Reviews, 1*, 33–41. doi:10.1016/j.esr.2011.12.002.

Melikoglu, M. (2014). Shale gas: Analysis of its role in the global energy market. *Renewable and Sustainable Energy Reviews, 37*, 460–468. doi:10.1016/j.rser.2014.05.002.

Mohsenian-Rad, A. H., Wong, V. W. S., Jatskevich, J., Schober, R., & Leon-Garcia, A. (2010). Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Transactions on Smart Grid, 1*, 320–331. doi:10.1109/TSG.2010.2089069.

Munasib, A., & Rickman, D. S. (2015). Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis. *Regional Science and Urban Economics, 50*, 1–17. doi:10.1016/j.regsciurbeco.2014.10.006.

Odling, N. E., Gillespie, P., Bourgine, B., Castaing, C., Chiles, J. P., Christensen, N. P., et al. (1999). Variations in fracture system geometry and their implications for fluid flow in fractures hydrocarbon reservoirs. *Petroleum Geoscience, 5*, 373–384. doi:10.1144/petgeo.5.4.373.

Patzek, T. W., Male, F., & Marder, M. (2013). Gas production in the Barnett Shale obeys a simple scaling theory. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 19731–19736. doi:10.1073/pnas.1313380110.

Rider, M. J., Garcia, A. V., & Romero, R. (2007). Power system transmission network expansion planning using AC model. *IET Generation, Transmission and Distribution,*. doi:10.1049/iet-gtd:20060465.

Roh, J. H., Shahidehpour, M., & Fu, Y. (2007). Market-based coordination of transmission and generation capacity planning. *IEEE Transactions on Power Systems, 22*, 1406–1419. doi:10.1109/TPWRS.2007.907894.

Ross, D. J. K., & Bustin, R. M. (2008). Characterizing the shale gas resource potential of Devonian-Mississippian strata in the Western Canada sedimentary basin: Application of an integrated formation evaluation. *AAPG Bulletin, 92*, 87–125. doi:10.1306/09040707048.

Sharan, I., & Balasubramanian, R. (2012). Integrated generation and transmission expansion planning including power and fuel transportation constraints. *Energy Policy, 43*, 275–284. doi:10.1016/j.enpol.2012.01.004.

Siirola, J. J. (2014). The impact of shale gas in the chemical industry. *AIChE Journal, 60*, 810–819. doi:10.1002/aic.14368.

Slutz, J., Anderson, J., Broderick, R., & Horner, P. (2012). Key Shale Gas Water Management Strategies: An Economic Assessment Tool. *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Smil, V. (2000). Energy in the twentieth century: resources, conversions, costs, uses, and consequences. *Annual Review of Energy and the Environment,*. doi:10.1146/annurev.energy.25.1.21.

Strbac, G. (2008). Demand side management: Benefits and challenges. *Energy Policy, 36*, 4419–4426. doi:10.1016/j.enpol.2008.09.030.

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*. doi:10.1016/j.rser.2011.08.014.

Szulczewski, M. L., MacMinn, C. W., Herzog, H. J., & Juanes, R. (2012). Lifetime of carbon capture and storage as a climate-change mitigation technology. *Proceedings of the National Academy of Sciences,*. doi:10.1073/pnas.1115347109.

Unidad de Planeación Minero Energética—UPME. (2014). Plan de Expansion de Referencia Generacion—Transmisión 2014-2028. Bogotá/Colombia.

Unsihuay-Vila, C., Marangon-Lima, J. W., Zambroni De Souza, A. C., & Perez-Arriaga, I. J. (2011). Multistage expansion planning of generation and interconnections with sustainable energy development criteria: A multiobjective model. *International Journal of Electrical Power & Energy Systems, 33*, 258–270. doi:10.1016/j.ijepes.2010.08.021.

Vengosh, A., Jackson, R. B., Warner, N., Darrah, T. H., & Kondash, A. (2014). A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environmental Science and Technology, 48*, 8334–8348. doi:10.1021/es405118y.

Vidic, R. D., Brantley, S. L., Vandenbossche, J. M., Yoxtheimer, D., & Abad, J. D. (2013). Impact of shale gas development on regional water quality. *Science, 340*, 1235009. doi:10.1126/science.1235009.

Vinasco, G., Tejada, D., Da Silva, E. F., & Rider, M. J. (2014). Transmission network expansion planning for the Colombian electrical system: Connecting the Ituango hydroelectric power plant. *Electric Power Systems Research, 110*, 94–103. doi:10.1016/j.epsr.2013.12.016.

Wang, J., Ryan, D., & Anthony, E. J. (2011). Reducing the greenhouse gas footprint of shale gas. *Energy Policy, 39*, 8196–8199. doi:10.1016/j.enpol.2011.10.013.

Warner, N. R., Christie, C. A., Jackson, R. B., & Vengosh, A. (2013). Impacts of shale gas wastewater disposal on water quality in western Pennsylvania. *Environmental science & technology*, *47*, 11849–11857. doi:10.1021/es402165b.

Weijermars, R. (2013). Economic appraisal of shale gas plays in Continental Europe. *Applied Energy, 106*, 100–115. doi:10.1016/j.apenergy.2013.01.025.

World Energy Council. (2013). World Energy Resources: 2013 Survey.

Wu, Y., Li, J., Ding, D., Wang, C., & Di, Y. (2013). A generalized framework model for simulation of gas production in unconventional gas reservoirs, *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers. doi:10.2118/163609-MS.

Xiao, J., Hodge, B. M. S., Pekny, J. F., & Reklaitis, G. V. (2011). Operating reserve policies with high wind power penetration. *Computers & Chemical Engineering, 35*, 1876–1885. doi:10.1016/j.compchemeng.2011.03.004.

XM. (2015a). Descripción del Sistema Eléctrico Colombiano [WWW Document]. http://www.xm.com.co/Pages/DescripciondelSistemaElectricoColombiano.aspx.

XM. (2015b). Informes Anuales de Operación [WWW Document]. http://informesanuales.xm.com.co/SitePages/Default.aspx.

Yang, L., Grossmann, I. E., & Manno, J. (2014). Optimization models for shale gas water management. *AIChE Journal, 60*, 3490–3501. doi:10.1002/aic.14526.

Zhu, H., & Huang, G. H. (2013). Dynamic stochastic fractional programming for sustainable management of electric power systems. *International Journal of Electrical Power & Energy Systems, 53*, 553–563. doi:10.1016/j.ijepes.2013.05.022.

Chapter 4
CFD Modeling of a Pilot-Scale Steam Methane Reforming Furnace

Andres Aguirre, Anh Tran, Liangfeng Lao, Helen Durand, Marquis Crose and Panagiotis D. Christofides

Abstract Hydrogen is a required key material for petroleum refineries that convert crude oil into a variety of products with higher economic value, e.g., gasoline. In chemical process plants and petroleum refineries, hydrogen is produced primarily by the steam methane reforming (SMR) process synthesizing hydrogen and carbon oxides from methane and superheated steam in the presence of a nickel-based catalyst network in a steam methane reformer. Traditionally, the optimized and profitable operating conditions of a steam methane reformer are analyzed and determined by on-site parametric study at industrial-scale plants or pilot-scale units, which is an experimental approach, and therefore, it must be conducted by changing process parameters in small increments over a long time period in order to prevent significant production and capital loss. Motivated by the above considerations, the present work focuses on developing a computational fluid dynamics (CFD) model of a pilot-scale steam methane reformer comprised of four industrial-scale reforming reactors, three industrial-scale burners and three flue gas tunnels. The pilot-scale reformer CFD model is developed by analyzing well-established physical phenomena, i.e., the transport of momentum, material and energy, and chemical reactions, i.e., combustion and the SMR process, that take place inside the steam methane reformer. Specifically, the $P-1$ radiation model, standard $k-\epsilon$ turbulence model, compressible ideal gas equation of state and finite rate/eddy dissipation (FR/ED) turbulence-chemistry interaction model are adopted to simulate the macroscopic and microscopic events in the reformer. The conditions for the tube-side feed, burner feed and combustion chamber refractory walls are consistent with typical reformer plant data Latham (2008) so that the simulation results generated by the pilot-scale reformer can be

A. Aguirre · A. Tran · L. Lao · H. Durand · M. Crose · P.D. Christofides (✉)
Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA
e-mail: pdc@seas.ucla.edu

P.D. Christofides
Department of Electrical Engineering, University of California, Los Angeles, CA 90095-1592, USA

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_4

validated by the plant data. The simulation results are shown to be in agreement with publicly available plant data reported in the literature and also with the simulation data generated by a well-developed single reforming tube CFD model. Subsequently, the proposed pilot-scale reformer CFD model is employed for a parametric study of the mass flow rate of the burner feed, i.e., a 20 % increase from its nominal value. The corresponding simulation results demonstrate the advantages offered by this CFD model for parametric study by showing that with the increased burner feed, the outer reforming tube wall temperature exceeds the maximum allowable temperature; these results were developed quickly with the aid of a CFD model, compared to the timescale on which parametric studies are performed on-site and without the potential for rupture of the reforming tubes during the study.

4.1 Introduction

Hydrogen is one of the most important raw materials for petroleum refineries that convert crude oil into a variety of products with higher economic value, e.g., gasoline, jet fuel and diesel, and its unavailability can limit the production rates of these petroleum products. The environmental requirement for low-sulfur-content fuels results in an increasing amount of hydrogen required by hydrotreating processes, and the attempt to process heavier components of the crude oil known as bottom-of-the-barrel processing also increases the demand for hydrogen in hydrocracking processes Udengaard (2004), Zamaniyan et al. (2010). Additionally, hydrogen can also be used as a fuel. Hydrogen is an efficient energy carrier, e.g., it can be converted into electrical energy by electrochemical reactions taking place inside fuel cells, which consist of three primary components, i.e., a cathode, an anode and an ion exchange membrane separating the electrodes. Particularly, hydrogen molecules are split into electrons and positively charged hydrogen ions at the anode by electrochemical reactions inside the fuel cell, and then the newly generated ions pass through the ion exchange membrane, which generates an electric potential between the electrodes.

Hydrogen is industrially produced primarily by the steam reforming process and partial oxidation of hydrocarbons Amirshaghaghi et al. (2010). Specifically, the steam reforming process is an overall endothermic process in which raw natural gas, e.g., methane, reacts with high-pressure, high-temperature steam (superheated steam) in the presence of a nickel-based catalyst to produce hydrogen, carbon dioxide and carbon monoxide. The partial oxidation of hydrocarbons, e.g., natural gas, petroleum coke and light naphtha, is an exothermic process in which hydrocarbons react with a limited amount of oxygen, typically less than the stoichiometric amount required for complete oxidation of the hydrocarbons to carbon dioxide and water, to produce hydrogen and carbon oxides (predominantly carbon monoxide). It is worth noting that the amount of hydrogen produced by the partial oxidation process is always less than that produced by the steam reforming process for a fixed amount of raw material. As a result, the steam reforming process, specifically, the steam

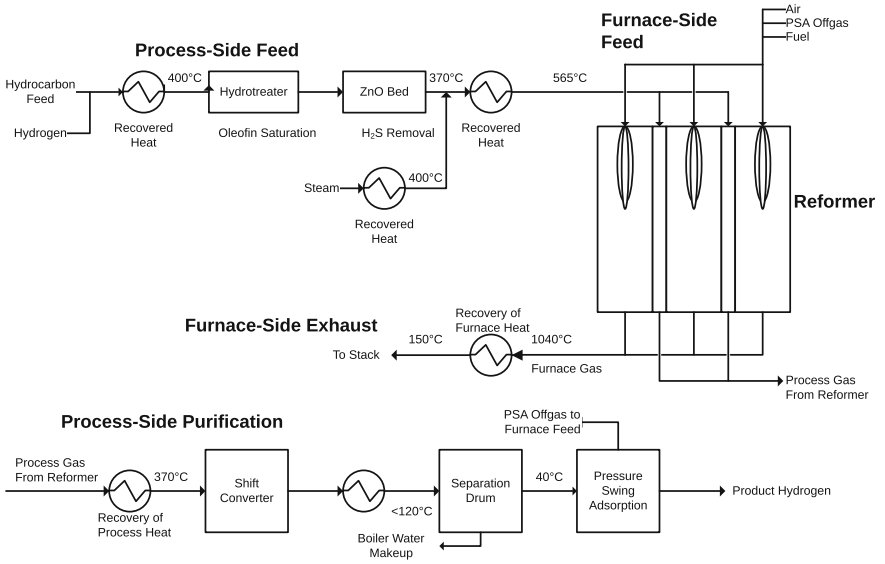

Fig. 4.1 Complete steam methane reforming process diagram, which includes the pretreating process, SMR process and purification process Kroschwitz and Howe-Grant (1999)

methane reforming (SMR) process (shown in Fig. 4.1), remains the most economic and common commercial method for industrial hydrogen production Sadooghi and Rauch (2013).

A steam methane reformer (it will be referred to as "reformer" in the following text) is the core unit in a SMR process, and it comes in four typical configurations, i.e., top-fired, side-fired, bottom-fired and terrace wall-fired reformers, which are categorized based on the burner arrangements. It is important to note that the terrace wall-fired configuration is a modification of the bottom-fired reformer, and therefore, they possess similar characteristics. Specifically, they are known to produce an approximately constant heat flux profile across the heated reforming tube length de Lasa et al. (1992). The top-fired reformers are characterized by a temperature field inside the combustion chamber for which the maximum temperature value and maximum heat flux are located in the upper part of the reformer de Lasa et al. (1992). Lastly, the characteristic of side-fired reformers that differentiates them from the other three configurations is that they allow control of the reforming tube wall temperature, which in turn offers more flexible design and operation to generate a desired temperature distribution throughout the reformer. Nevertheless, the number of burners in side-fired reformers is ∼4 times as many as that in top-fired units with a similar amount of thermal energy transferred to the reforming tubes de Lasa et al. (1992). Therefore, a top-fired configuration, which includes burners located at the reformer ceiling, flue gas tunnels located at the reformer floor allowing the furnace-side mixture to exit the reformer, and reforming tubes, is employed most

frequently in hydrogen plants and the petroleum refinery industry, and is the subject of this work.

The top-fired reformer is composed of two closed domains including a tube side, which consists of tubular reforming reactors (they will be referred to as the "reforming tubes") and a furnace side, which is a combustion chamber. The combustion chamber encapsulates the thermal energy released by the oxidation of the furnace-side feed composed of methane, hydrogen, carbon oxides and air, to prevent it from escaping to the surrounding environment through the combustion chamber refractory walls. Typically, only ∼2–5 % of the total fired duty is lost Latham (2008), while a significant percentage of trapped thermal energy, e.g., ∼50 %, is transferred to the reforming tubes, primarily by radiative heat transfer de Lasa et al. (1992). The remaining energy leaves the reformer by species transport through the chamber outlet. In the tube side, prior to entering the reforming tubes, raw natural gas undergoes a hydro-treating process to remove unsaturated hydrocarbon compounds. Then, this hydro-treated natural gas also undergoes a desulfurization process to remove sulfur compounds to prevent the nickel-based catalyst of the steam methane reforming process from being poisoned and deactivated. As a result, the major component of pre-treated natural gas is methane. Subsequently, the pre-treated natural gas is mixed with superheated steam prior to be fed into the reforming tubes where an overall endothermic SMR process driven by the thermal energy absorbed from the furnace side and facilitated by the nickel-based catalyst network converts steam and methane into the desired product, i.e., hydrogen, and byproducts, i.e., carbon oxides (including CO and CO_2).

In 2004, the hydrogen production rate and annual growth rate of hydrogen production in North America were estimated to be 6,700,000 Nm^3/h and above 4 %, respectively Latham (2008), and the largest plant could produce up to 300,000 Nm^3/h Latham (2008). Due to the aforementioned hydrogen production rate of plants, the annual cost of raw materials alone would be on the order of millions of dollars. For instance, a hydrogen plant with a production rate of 112,000 Nm^3/h can spend up to 62 million dollars on raw natural gas annually Latham (2008). Therefore, a small improvement in process efficiency results in a great gain in profit margin of a plant. This became the driving force for the development of reformer study, and specifically, reformer parametric study has become a highly researched topic. Since the mid-1900s, extensive work has been conducted on the development of reformer first-principles modeling, and in the 1960s McGreavy and Newmann (1969), the first mathematical model of a complete reformer was developed and proposed. This mathematical model of reformers gradually became more sophisticated and highly complex in order to account for physical phenomena, i.e., the transport of momentum, materials and energy, and chemical phenomena, i.e., combustion processes and the SMR process, taking place inside the unit. As a result, the mathematical model of the complete reformer composed of two governing equation sets, i.e., the continuity equation, and momentum, species and energy conservation equations, is a set of highly non-linear coupled differential equations, which are solved simultaneously to characterize the reformer. Unfortunately, solving the mathematical model of the complete reformer is a formidable task.

With the onset of technology evolution, computational fluid dynamics (CFD) modeling became a powerful tool for predicting fluid behavior with a high level of accuracy, and also one of the most effective tools employed in reformer parametric study because the parametric effect can be visualized and quantified through CFD simulations. Additionally, CFD modeling can precisely capture all geometry characteristics of a given reformer through computer-aided design software, which in turn allows CFD models to generate simulation results that can be expected to serve as reasonable substitutes for experimental data. Therefore, CFD modeling has gained popularity and become the primary method for reformer parametric study to replace the traditional on-site experimental approach for optimizing the SMR process, which is time-consuming, expensive and potentially risky.

For example, the SMR process converting superheated steam and methane into carbon oxides and hydrogen is an overall endothermic and reversible process, and therefore, in order to maintain or increase the hydrogen production rate, thermal energy has to be continuously supplied to replenish that consumed by the SMR process. It is important to note that hydrogen conversion for the SMR process depends primarily on the temperature field inside the combustion chamber, and more specifically, the temperature of the outer reforming tube wall, i.e., a higher outer reforming tube wall temperature theoretically results in a higher hydrogen conversion. Nevertheless, operating at excessively high temperature decreases the reforming tube expected life span, i.e., an increase in outer reforming tube wall temperature of 20 K from the nominal operating temperature can reduce the reforming tube lifetime by half Pantoleontos et al. (2012), Latham (2008). In addition, it might cause the reforming tube wall to lose its integrity and to rupture, which results in production and capital losses. Specifically, the required total capital investment to replace a typical industrial-scale reformer is estimated to be 5–8 million USD Pantoleontos et al. (2012). Consequently, on-site parametric study to determine the optimized furnace-side feed flow rate to the burners must be carried out with small changes in the parameters, while the outer reforming tube walls are closely monitored in order to avoid the aforementioned production and capital losses. Therefore, this high-risk study usually is conducted over a long time period, and is costly. On the contrary, simulation results generated by a well-developed reformer CFD model are expected to be consistent with experimental data collected from industrial-scale plants because CFD models are capable of capturing not only physical and chemical phenomena, but also the geometry of a reformer. Therefore, CFD modeling can be used as an alternate method for parametric study. Furthermore, a well-developed CFD model of a reformer can provide insights into the system which cannot be captured in experimental data generated by on-site parametric study, e.g., the species distributions inside the combustion chamber.

Motivated by the above considerations, we initially employ *ANSYS Fluent* CFD software to develop a pilot-scale reformer CFD model that is composed of four industrial-scale reforming tubes, three industrial-scale burners (details given in Sect. 4.2) and three flue gas tunnels. It is important to note that the industrial-scale reforming tube modeling strategy is adopted from our previous work Lao et al. (2016), which has an approximate representation of the catalyst network that can

account for the presence of catalyst particles inside the reforming tube, and the effect of internal and external diffusion limitations on the observed reaction rates of the SMR process (details given in Sect. 4.4.2). The boundary conditions for the reforming tube inlet (referred to in the following text as "tube-side feed"), burner inlet (referred to in the following text as "furnace-side feed"), and combustion chamber refractory walls are consistent with typical plant data, so that simulation data generated by the pilot-scale reformer is in agreement with the former data Latham (2008). Based on the conditions of the tube-side feed and furnace-side feed, and the expected physical and chemical phenomena taking place inside the reformer, the standard $k - \epsilon$ turbulence model, finite rate/eddy dissipation turbulence chemistry interaction model (details given in Sect. 4.6) and global kinetic models of combustion processes Bane et al. (2010), Nicol (1995) and the SMR process Xu and Froment (1989) are implemented to capture the individual species reaction rates in the turbulent reacting flow. Next, the simulation results generated by the pilot-scale reformer CFD model are rigorously validated by comparing them with available data in the literature and simulation results generated by a well-developed single reforming tube CFD model. Finally, we conduct a parametric study on furnace-side feed flow rate variation, i.e., a 20 % increase in the furnace-side inlet mass flow rate is introduced, in order to demonstrate the importance of the pilot-scale CFD model for evaluating conditions that may not be safely accessed experimentally.

4.2 Industrial Steam Methane Reformer Geometry

The pilot-scale reformer investigated in this work is developed based on an industrial-scale top-fired, co-current reformer designed by Selas Fluid Processing Corporation (Figs. 4.2 and 4.3). This industrial-scale reformer produces 2,830,000 Nm^3 of high-purity hydrogen along with 1,708,800 kg of superheated steam (i.e., 663.15 K and 4,580 kPa) per day Latham (2008). The combustion chamber of this industrial-scale reformer contains seven rows of forty-eight reforming tubes. The external diameter, internal diameter and exposed length of these reforming tubes are 14.6 cm, 12.6 cm and 12.5 m, respectively. Inside these reforming tubes, nickel-based catalyst pellets, specifically, alpha-alumina-supported nickel oxide denoted as $NiO - \alpha Al_2O_3$, are used as packing material to facilitate the formation of hydrogen from steam and methane through the endothermic SMR process, and to act as an intermediate medium to enhance the rate of convective heat transfer to the tube-side gas mixture. At the combustion chamber ceiling, these rows of reforming tubes are separated by eight rows of twelve burners which are fed with a furnace-side feed composed of two separate streams, i.e., a fuel stream composed of natural gas and tail gas (CO and H_2), and an oxidizer stream composed of combustion air (Ar, N_2 and O_2). It is important to mention that the rows of burners adjacent to the combustion chamber refractory walls and a single row of reforming tubes (for brevity, these burners are denoted as "outer-lane burners") are fed with a lower furnace-side feed flow rate than the rows of burners adjacent to two rows of reforming tubes (for brevity, these burners are

Fig. 4.2 The top view of an industrial-scale, top-fired, co-current reformer with 336 reforming tubes, which are symbolized by 336 smaller circles, and 96 burners, which are denoted by 96 larger circles. The outer-lane burners are shown as the circles on the right and left boundaries of the figure, while the inner-lane burners are shown as slightly larger circles within the figure

denoted as "inner-lane burners"). Specifically, the furnace-side feed flow rate of the outer-lane burners is 60 % of that of the inner-lane burners to avoid causing "over-firing" in the outer lanes and "under-firing" in the inner lanes, which would occur if the same furnace-side feed flow rate were fed to all burners. Based on typical flow rates and species compositions of the fuel and oxidizer streams of an inner-lane burner, the fuel stream of the furnace-side feed is completely oxidized over a flame length of 4.5–6 m releasing the thermal energy needed to drive the SMR process Latham (2008). The thermal energy released by combusting the furnace-side feed is transferred to the reforming tubes predominantly by radiative heat transfer inside the high-temperature furnace chamber. At the reformer floor, the rows of reforming tubes are separated by the rectangular intrusions known as flue gas tunnels or coffin boxes which extend from the front to the back of the combustion chamber along the rows of reforming tubes with a height of 2.86 m from the floor. Additionally, there are thirty-five extraction ports evenly distributed in a row along each side of the flue gas tunnels that allow the furnace flue gas to enter the flue gas tunnels, and then to exit the combustion chamber through the front openings of the flue gas tunnels. In this work, we will focus on the development and analysis of a CFD model of a pilot-scale version of the industrial-scale reformer described above.

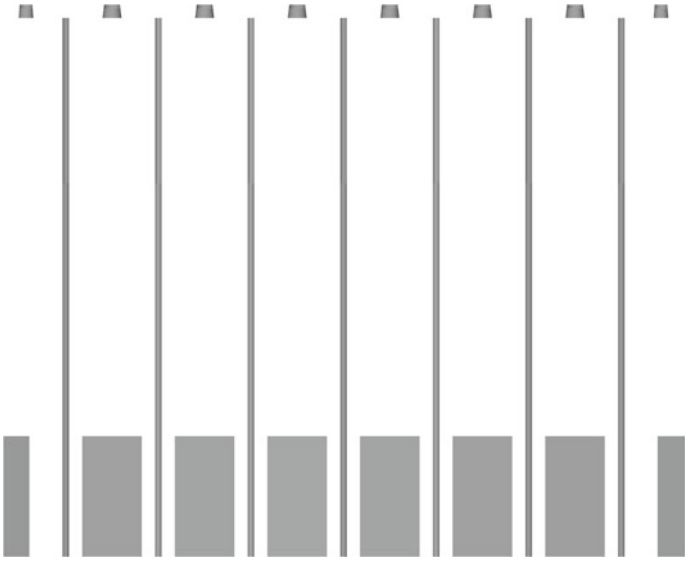

Fig. 4.3 The front view of an industrial-scale, top-fired, co-current reformer. It is important to note that the refractory wall of the combustion chamber is modeled to be transparent, so that the interior components, which include eight rectangular boxes located on the floor representing the flue gas tunnels, eight frustums of cones located on the ceiling representing the corresponding rows of burners and seven slender rectangles connecting the ceiling and floor representing the corresponding rows of reforming tubes, can be seen

4.2.1 Geometry of Pilot-Scale SMR Unit

As noted, this work focuses on the development of a pilot-scale version of the industrial-scale reformer described in the previous section. The pilot-scale reformer contains the reforming tubes, inner-lane burners, outer-lane burners and extraction ports, with the same dimensions and geometry as those of the industrial-scale reformer. It is critical for our work to obtain a pilot-scale reformer geometry such that the environment of each individual component is as close as possible (if not identical) to that in the industrial-scale unit. Thus, in this pilot-scale reformer CFD model, there are two rows of two reforming tubes, and each is surrounded by two rows of burners. Based on the estimated thermal energy released by combusting the furnace-side feed for one inner-lane burner and the expected thermal energy absorbed by each reforming tube from the available industrial data, it is determined that only one burner is needed in each row of burners in this pilot-scale reformer to provide the desired heat to the reforming tubes. As mentioned in the previous section, there are two types of burners in the industrial-scale reformer, i.e., the outer-lane burners, which are adjacent to one row of reforming tubes and to the refractory wall of the combustion chamber, and the inner-lane burners, which are adjacent to two rows of reforming tubes. Additionally, it is important to note that the outer-lane burners are

smaller in size, and are fed with 40 % less of the furnace-side feed than the inner-lane burners. Furthermore, in this pilot-scale reformer CFD model, the rows of reforming tubes are also separated by the flue gas tunnels with extraction ports at the floor of the combustion chamber as they are in the industrial-scale reformer. The geometry of the pilot-scale reformer is shown in Fig. 4.4.

4.2.2 Mesh of Pilot-Scale SMR Unit

Due to the tightly coupled chemical, physical and transport phenomena taking place inside the pilot-scale reformer and the multi-step complex reaction schemes (i.e., the SMR process Xu and Froment (1989) and combustion reactions) with corresponding nonlinear empirical kinetic formulas, it is impossible to obtain an analytical solution to characterize the fluid-flow and temperature fields. Instead, the computational domain is divided into fine, discrete elements also known as grids, within which spatial variations are relatively small. The governing equations (i.e., the continuity equation and the momentum, material and energy conservation balances) required for characterizing the heat and fluid-flow fields inside the pilot-scale reformer are numerically solved for each grid. Then, the numerical solutions of the grids are patched together to reconstruct the solution of the original domain. Hence, creating a mesh with acceptable mesh quality is a critical task that determines the success level of CFD modeling because a poor quality mesh needs the most robust CFD solver and demands greater computing resources to determine a converged solution. Additionally, a CFD model with a poor quality mesh has a slow speed of convergence and is more likely to converge to an inaccurate solution because mesh quality directly determines solver discretization error Inc. (2013).

There are two major classes of meshing strategies, i.e., the unstructured tetrahedral meshing strategy (for simplicity, it is denoted as "unstructured meshing" in the following text) within which discrete grids are arranged in an irregular pattern, and the multiblock structured hexahedral meshing strategy (for simplicity, it is denoted as "structured meshing" in the following text). Although unstructured meshing is more proficient at approximating complicated geometries than structured meshing, the latter class is chosen to develop a CFD model of this pilot-scale reformer (Figs. 4.5 and 4.6) since the pilot-scale reformer geometry shown in Fig. 4.4 is composed primarily of straight edges. Additionally, for wall-bounded systems like the pilot-scale reformer, a CFD model built from structured meshing generally generates a converged solution closer to experimental data and also has a superior speed of convergence compared to other CFD models built from unstructured meshing when the system is decomposed into the same number of discrete grids Inc. (2013). Furthermore, the *ANSYS ICEM* environment provides an O-grid Block function to enhance the structured meshing's ability to approximate curvy geometry characteristics by re-arranging existing grid lines into an *O* shape to effectively improve the overall mesh quality Inc. (2013). In the pilot-scale reformer CFD model, the sub-mesh regions of the inner-lane burners, outer-lane burners and reforming tubes shown in

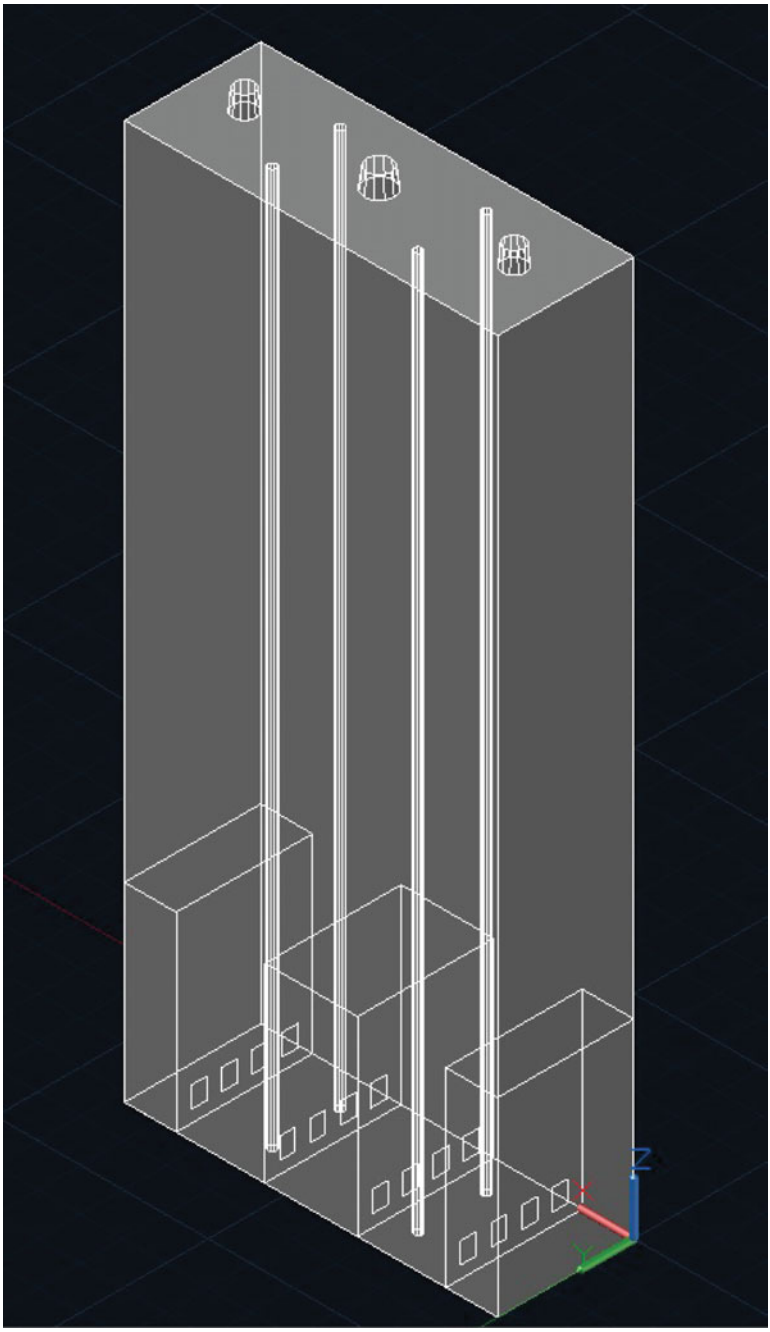

Fig. 4.4 CAD geometry of pilot-scale furnace model

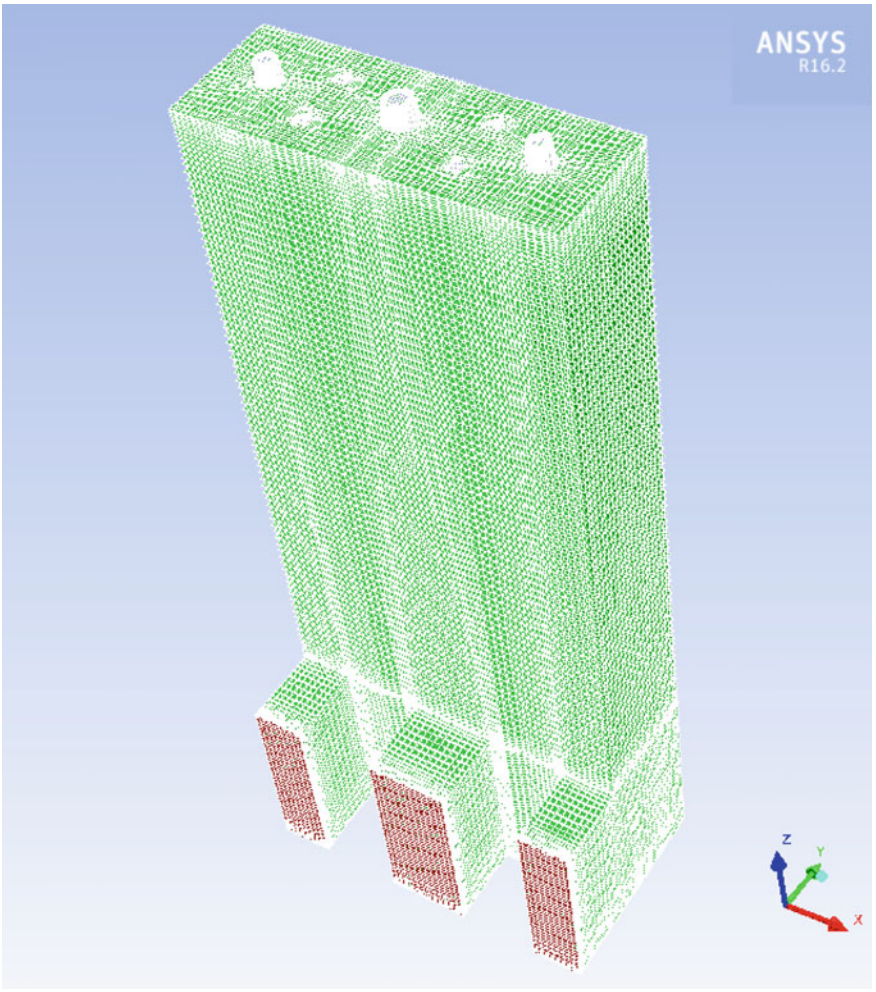

Fig. 4.5 Isometric view of the hexahedral structured mesh of the pilot-scale reformer

Figs. 4.7, 4.8 and 4.9 are created by the O-grid Block function due to their unique geometry characteristics. Specifically, the burners have a frustum-like structure, and the reforming tubes have a cylindrical structure. It is important to note that in this pilot-scale reformer the grids are not uniformly distributed but instead they are designed to be more dense in the regions that are expected to have large momentum, material and temperature gradients as shown in Figs. 4.5 and 4.6.

An example of a cluster of regions within which the grid density is higher, is within and around the reforming tubes, where the process of transferring thermal energy from the furnace-side mixture to the tube-side mixture is very complex. Specifically, the thermal energy released from combusting the furnace-side feed is

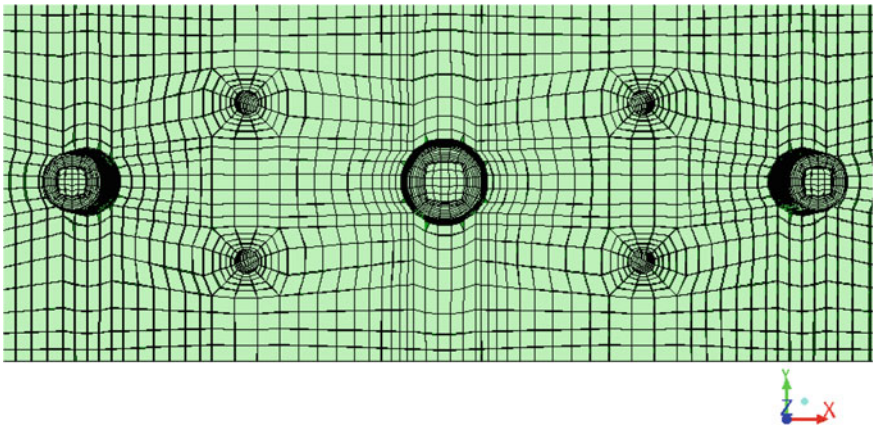

Fig. 4.6 Top view of the hexahedral structured mesh of the pilot-scale reformer. This figure demonstrates that the geometries of the outer-lane burners, inner-lane burners and reforming tubes are successfully preserved by means of the O-grid Block function of *ANSYS ICEM*

transferred to the reforming tube outer wall mostly by radiative heat transfer. Then, it is conducted to the inner reforming tube wall and to the catalyst network, and finally, it is transferred to the tube-side mixture by convective heat transfer. Therefore, to capture the process through which thermal energy is transferred from the furnace-side mixture to the tube-side mixture, the grid density in the aforementioned regions is designed to be more dense to account for the large spatial gradients. Similarly, in the flame regions where the combustion process of the furnace-side feed takes place as previously mentioned in Sect. 4.2, large spatial variations of transport variables are expected. Additionally, because the reaction rates for the combustion process are fast and are described by nonlinear empirical kinetic formulas, the formation/consumption rates of the components of the furnace-side feed are directly controlled by the transport rates of these components. Therefore, it is essential that the sub-mesh regions of the burners and their vicinity, shown in Fig. 4.6, and of the flame regions inside the combustion chamber, are composed of finer grids in order to more accurately predict the species distributions, and to simulate the flame characteristics.

The mesh of the pilot-scale reformer, shown in Fig. 4.5, accounting for all of the above considerations, contains 165,817 hexahedral grids, 516,325 quadrilateral cells and 184,636 nodes. The mesh quality of the pilot-scale reformer mesh chosen to construct the CFD model for the remaining investigation is evaluated based on three suggested criteria including the orthogonal factor, aspect ratio and ortho skew as shown in Table 4.1. It is important to note that a mesh is considered to have poor quality if the mesh quality is outside of the recommended range for any of the three criteria; the mesh quality values of the pilot-scale reformer are above all minimum recommended values and below all maximum recommended values as shown in the table, and thus, can be considered to have reasonably good quality.

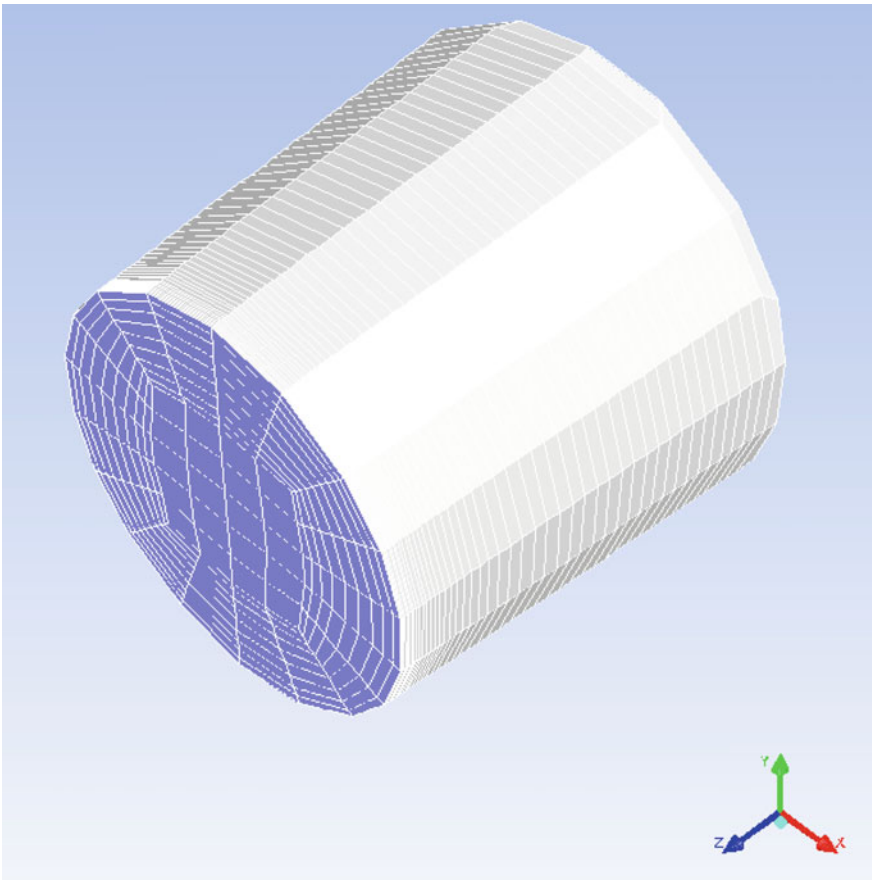

Fig. 4.7 Isometric view of the hexahedral structured mesh of the inner-lane burner

4.3 Furnace Chamber Modeling

In the pilot-scale reformer, the composition of the furnace-side feed is identical to that of the industrial-scale unit, where the furnace-side feed is composed of the fuel stream (CH_4, CO and H_2) and the oxidizer stream (Ar, N_2 and O_2). Nevertheless, due to a significant difference between the industrial-scale and pilot-scale refomers in the ratio of the number of reforming tubes to the number of burners, i.e., the ratios in the industrial-scale and pilot-scale reformers are 3.50 and 1.33, respectively, when the furnace-side feed flow rate of the inner-lane burners from the industrial-scale reformer is implemented in the pilot-scale unit, the reforming tubes will have an excessively high outer reforming tube wall temperature, which can lead to disastrous consequences and significant capital loss as discussed in Sect. 4.1. Therefore, an adjustment is made to the furnace-side feed flow rate of the inner-lane burner in

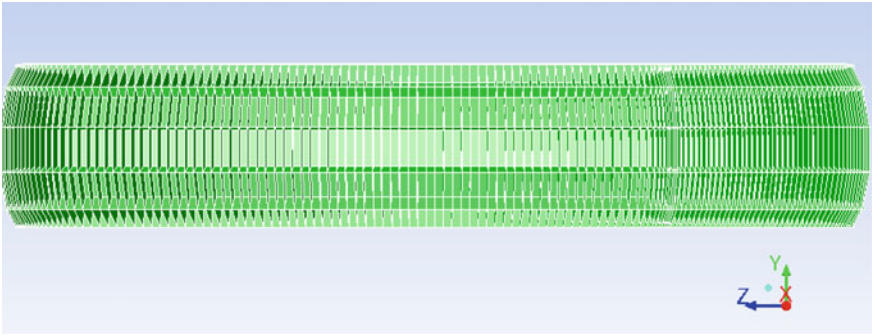

Fig. 4.8 Side view of the hexahedral structured mesh of a reforming tube. In this figure, the radial direction is scaled up by 20 times for display purposes only

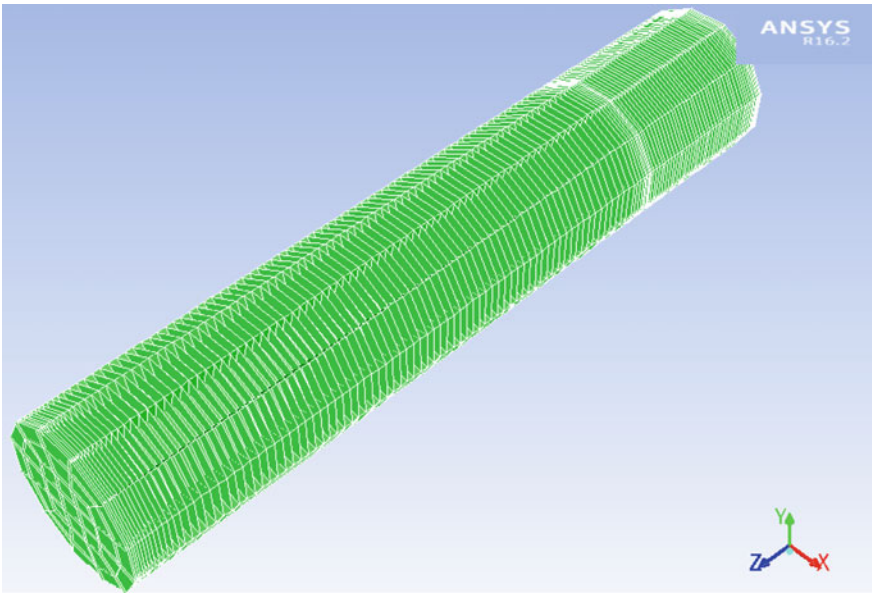

Fig. 4.9 Isometric view of the hexahedral structured mesh of a reforming tube. In this figure, the radial direction is scaled up by 20 times for display purposes only

the pilot-scale reformer to avoid the aforementioned issue. Specifically, the furnace-side feed flow rate is decreased to 37.5 % of that in the industrial-scale reformer based on the estimated thermal energy released by combusting the furnace-side feed of a single inner-lane burner in the industrial-scale reformer, the expected average heat flux across the reforming tube wall and the amount of heat absorbed by each reforming tube. Additionally, in the pilot-scale reformer CFD model, the ratio of the inner-lane burner and outer-lane burner furnace-side feed flow rates is 0.6 as in the industrial-scale reformer.

Table 4.1 Mesh quality of the pilot-scale reformer mesh

	The pilot-scale unit mesh	Recommended value
Minimum orthogonal factor	0.3934	0.1667
Maximum ortho skew	0.6066	0.8500
Maximum aspect ratio	16.14	100.0

4.3.1 Combustion Reaction Kinetic Model and Turbulence Chemistry Model

In the combustion chamber, methane and hydrogen in the furnace-side feed are completely oxidized to generate carbon dioxide, water and a large amount of thermal energy required by the SMR process. It is important to note that the chemistry of the combustion process is a complex network of sequential elementary reactions governed by the concentrations of free radicals, which are reactive molecules with unpaired electrons. For instance, the complete mechanism of the hydrogen combustion process generating water involves more than 20 elementary reactions and various intermediates Turns (1996). Therefore, the computing power required to simulate a pilot-scale reformer CFD model implemented with a complete reaction mechanism providing the detailed descriptions of all elementary steps involved in the formation and consumption rates of free radicals and intermediates of the methane and hydrogen oxidation processes would be too demanding. As a result, simplified global kinetic models for the combustion of methane Nicol (1995) and hydrogen Bane et al. (2010) are adopted to reduce the computational requirement for simulating the pilot-scale reformer CFD model. In detail, the combustion processes for natural gas and hydrogen considered in this work involve a two-step mechanism and a single-step mechanism respectively with reaction rates given by the following global kinetic models:

$$CH_4(g) + 1.5O_2(g) \xrightarrow{R_1} CO(g) + 2H_2O(g), \tag{4.1a}$$

$$R_1 = 10^{15.22}[CH_4]^{1.46}[O_2]^{0.5217} exp(-20643/T) \tag{4.1b}$$

$$CO(g) + 0.5O_2(g) \underset{R_3}{\overset{R_2}{\rightleftharpoons}} CO_2(g) \tag{4.1c}$$

$$R_2 = 10^{14.902}[CO]^{1.6904}[O_2]^{1.57} exp(-11613/T) \tag{4.1d}$$

$$R_3 = 10^{14.349}[CO_2] exp(-62281/T) \tag{4.1e}$$

$$H_2(g) + 0.5O_2(g) \xrightarrow{R_4} H_2O(g) \tag{4.2a}$$

$$R_4 = 4.61 \times 10^{15}[H_2][O_2] exp(-10080/T) \tag{4.2b}$$

where R_1, R_2, R_3 and R_4 are the intrinsic volumetric reaction rates and are measured in $\frac{\text{kmol}}{\text{m}^3\ \text{s}}$, T (K) is the temperature of the surrounding environment in which the

reactions take place and $[CO_2]$, $[CO]$, $[O_2]$, $[H_2]$ and $[CH_4]$ are the molar concentrations of the corresponding species and have the units of $\frac{\text{kmol}}{\text{m}^3}$. These empirical kinetic formulas are in the Arrhenius form, as a result, they can be directly implemented into the pilot-scale reformer CFD model to simulate the formation and consumption rates of the individual species.

In the industrial-scale reformer, the furnace-side feed composed of two separate streams, i.e., the fuel stream and the oxidizer stream, is combusted inside the combustion chamber to generate the required fired duty for the SMR process. It is important to note that the intrinsic nature of non-premixed combustion is turbulent mixing-controlled, i.e., the rate of the chemical reactions is relatively faster than that of mixing, and the observed formation and consumption rates of each individual species depends on the rate of mixing. In the remain of this section, we demonstrate the modeling strategy that allows the pilot-scale reformer CFD model with the premixed combustion model to exhibit the nature of non-premixed combustion processes.

Specifically, in the pilot-scale reformer CFD model, the fuel stream and oxidizer stream of the furnace-side feed are assumed to be well-mixed prior to being fed into the combustion chamber where the fuel stream is fully oxidized to generate the needed thermal energy for the SMR process. Thus, the pressure, temperature, flow rate and species composition of the furnace-side feed of the inner-lane burner are shown in Table 4.2, and the combustion of methane and hydrogen is modeled by the premixed combustion model. However, the intrinsic nature of non-premixed combustion processes must be shown in the simulation result generated by the pilot-scale reformer CFD model with the premixed combustion model. This issue is resolved by using the finite-rate/eddy-dissipation (FR/ED) model as the turbulence-chemistry interaction model to simulate the rates of formation and consumption of each individual species in the turbulent flow. In particular, the FR/ED model estimates the observed reaction rate of species i, for which i are components of the furnace-side mixture, based on the global kinetic models shown in Eqs. 4.1 and 4.2 and the eddy-dissipation reaction rates (as shown in Eq. 4.3) Inc. (2013):

Table 4.2 Furnace-side inlet operating conditions of the inner-lane burner

Pressure (kPa)	132.4
Temperature (K)	525
Flow rate (kg/s)	0.4056
x_{H_2O}	0.03061
x_{O_2}	0.15305
x_{Ar}	0.00765
x_{N_2}	0.57396
x_{H_2}	0.05399
x_{CO_2}	0.10797
x_{CO}	0.02113
x_{CH_4}	0.05164

$$R_{i,j} = \nu_{i,j} M_i A \rho \frac{\epsilon}{k} \min_{\mathcal{R}} \left(\frac{Y_{\mathcal{R}}}{\nu_{\mathcal{R},j} M_{\mathcal{R}}} \right) \tag{4.3a}$$

$$R_{i,j} = \nu_{i,j} M_i A B \rho \frac{\epsilon}{k} \frac{\sum_{\mathcal{P}} Y_{\mathcal{P}}}{\sum_n^N \nu_{n,j} M_n} \tag{4.3b}$$

$$R_{i,j} = \nu_{i,j} M_i R_j \tag{4.3c}$$

where $R_{i,j}$ ($\frac{\text{kg}}{\text{m}^3\ \text{s}}$) and $\nu_{i,j}$ are the consumption/formation rate and stoichiometric coefficient of species i in reaction j, M_i is the molecular weight of species i, $Y_{\mathcal{R}}$ and $M_{\mathcal{R}}$ are the mass fraction and molecular weight of a specified reactant $\mathcal{R}$, $\nu_{\mathcal{R},j}$ is the stoichiometric coefficient of a specified reactant $\mathcal{R}$ in reaction j, A and B are empirical constants equal to 4.0 and 0.5, k and ϵ are the turbulence kinetic energy and dissipation rate (which will be discussed in a subsequent section), $Y_{\mathcal{P}}$ is the mass fraction of a product species $\mathcal{P}$ in reaction j, R_j ($\frac{\text{kmol}}{\text{m}^3\ \text{s}}$) is the consumption/formation rate of reaction j from Eqs. 4.1 and 4.2, ρ is the local average density of the furnace-side mixture and n is the index of the product species involved in reaction j Inc. (2013).

When the FR/ED model as shown in Eq. 4.3 is adopted in the pilot-scale reformer CFD model, the observed reaction rate of species i is equal to the minimum reaction rate between Eqs. 4.3a, 4.3b, and 4.3c to reduce the discrepancy between the premixed combustion model and non-premixed combustion processes. Particularly, the furnace-side feed is fed into the combustion chamber at 525 K, and at this temperature the species i reaction rate (Eq. 4.3c) calculated based on the given reaction kinetic models (Eqs. 4.1 and 4.2) is smaller than that calculated based on the eddy-dissipation reaction rate because of the well-mixed furnace-side feed assumption. Thus, the observed reaction rate of species i at the burner inlet is estimated based on the former approach. However, within the flame region the species i rate of change calculated based on the given reaction kinetic models is greater than that calculated based on the eddy-dissipation reaction rate because of the flame temperature. Thus, the observed reaction rate of species i is determined based on the eddy-dissipation reaction rate. This shows that by using the minimum reaction rate from Eqs. 4.3a–4.3c, the fuel is prevented from being immediately oxidized when it first enters the combustion chamber Inc. (2013) As a result, the premixed combustion model coupled with the FR/ED model exhibits the turbulent-mixing controlled characteristics of non-premixed combustion processes. It is worth noting that because of the choice of simplified global kinetic models for the combustion of methane and hydrogen, the FR/ED model is selected as a suitable turbulence-chemistry interaction model for the pilot-scale reformer CFD model. This is because when the global reaction mechanisms for methane and hydrogen combustion consist of less than three elementary reactions, so the CFD model implemented with FR/ED model is expected to produce an accurate converged solution Inc. (2013).

4.3.2 Radiation Heat Transfer Modeling

Inside the combustion chamber, it has been established that thermal energy is transferred predominantly by radiative heat transfer, and as a result, a computationally efficient and accurate radiation model is critical to the success of the pilot-scale reformer CFD model. Due to the complexity of the combustion chamber geometry as described in Sect. 4.2.1 and the complicated thermal coupling of the furnace side and the process side, a differential approximation radiation model, also known as $P-1$, is employed to determine the rate at which energy is transferred by radiative heat transfer inside the combustion chamber Inc. (2013). This is because the $P-1$ model can easily be applied to complicated geometries with curvilinear coordinates, while the radiative energy transfer can be accurately computed with minimal CPU requirement Inc. (2013). It is important to note that in this work, the furnace-side mixture is assumed to be a non-scattering medium, and as a result, the $P-1$ model is formulated as shown in Eq. 4.4:

$$\vec{q}_{rad} = -\frac{1}{3\sigma_a}\nabla G \tag{4.4a}$$

$$-\nabla \cdot \vec{q}_{rad} = \sigma_a G - 4\sigma_a n_{comb}^2 \sigma T^4 \tag{4.4b}$$

$$\nabla \cdot \left(\frac{1}{3\sigma_a}\nabla G\right) = \sigma_a G - 4\sigma_a \sigma T^4 \tag{4.4c}$$

$$q_{rad}^{wall} = -\frac{\epsilon_{wall}}{2(2-\epsilon_{wall})}\left(4n_{comb}^2 \sigma T_{wall}^4 - G_{wall}\right) \tag{4.4d}$$

where σ_a and n_{comb} are the absorption coefficient and refractive index of the furnace-side mixture, which are estimated based on the weighted sum of gray gas (WSGG) model, σ is the Stefan-Boltzmann constant, ϵ_{wall} is the internal emissivity coefficient of the wall, T and T_{wall} are the temperature at a location inside the furnace chamber and at the wall surface where energy is transferred by radiative heat transfer, G and G_{wall} are the incident radiation at a location inside the furnace chamber and at the wall surface where energy is transferred by radiative heat transfer, respectively, and analogously $-\nabla \cdot \vec{q}_{rad}$ and q_{rad}^{wall} are the radiative heat flux at a location inside the furnace chamber and at the wall surface where energy is transferred by radiative heat transfer. It is noteworthy that the transport equation for G (as shown in Eq. 4.4c) is formed by substituting Eq. 4.4a into Eq. 4.4b, which is solved to determine the local radiation incident, and to ultimately estimate the local radiative heat transfer Inc. (2013). Additionally, it is critical to a successful modeling task to realize that the internal emissivity of the wall surface is an intrinsic property of the surface, and therefore, it only depends on the surface's characteristics, e.g., the surface texture,

Table 4.3 Combustion chamber refractory walls properties Latham (2008)

Density $\left(\frac{\text{kg}}{\text{m}^3}\right)$	3950
Heat capacity $\left(\frac{\text{J}}{\text{kg}\cdot\text{K}}\right)$	718
Thermal conductivity $\left(\frac{\text{W}}{\text{m}\cdot\text{K}}\right)$	2.6
Emissivity	0.65

instead of the surface material. In the pilot-scale reformer CFD model, the assumption of a temperature-independent emissivity coefficient for the wall surfaces is employed. Specifically, the emissivity coefficients of the reforming tubes, refractory wall and tunnel wall are chosen to be 0.85, 0.65 and 0.65, respectively Latham (2008) (Table 4.3).

4.4 Reforming Tube Modeling

4.4.1 Reforming Reaction Kinetic Model

In a reforming tube, the endothermic reforming reactions and water-gas shift reaction taking place at the catalyst active sites to convert reactants, i.e., steam and methane, into products, i.e., hydrogen, reach equilibrium before the tube-side mixture exits the reforming tube. Specifically, reactants are transported from the bulk of the tube-side mixture to the surface of the catalyst network by convective mass transfer driven primarily by the reactant concentration difference, which is generated by the external diffusion resistance of the catalyst network. Then, they diffuse down the reactant concentration gradient from the surface of the catalyst network through the catalyst medium to the catalyst active sites, where the reforming and water-gas shift reactions occur to generate the desired hydrogen along with carbon oxides. This reactant concentration gradient within the catalyst is generated by the internal diffusion resistance of the catalyst network. Next, the products diffuse from the catalyst active sites back to the surface of the catalyst network, and eventually emerge back into the tube-side mixture. A reaction kinetic model that provides a detailed description of all elementary steps involved in these catalyst-specific transport and reaction phenomena would be unnecessarily complex for the purpose of this work. Therefore, a simplified global kinetic model Xu and Froment (1989), which is widely-accepted, and is frequently used in CFD modeling and first-principles modeling of the SMR process, is also used in this paper to lessen the computational demand without sacrificing the accuracy of the simulation results as follows:

$$CH_4(g) + H_2O(g) \rightleftarrows CO(g) + 3H_2(g),$$

$$R_5 = \frac{k_1}{p_{H_2}^{2.5}} \left(p_{CH_4} p_{H_2O} - \frac{p_{H_2}^3 p_{CO}}{K_1} \right) / DEN^2 \tag{4.5a}$$

$$CO(g) + H_2O(g) \rightleftarrows CO_2(g) + H_2(g),$$

$$R_6 = \frac{k_2}{p_{H_2}} \left(p_{CO} p_{H_2O} - \frac{p_{H_2} p_{CO_2}}{K_2} \right) / DEN^2 \tag{4.5b}$$

$$CH_4(g) + 2H_2O(g) \rightleftarrows CO_2(g) + 4H_2(g),$$

$$R_7 = \frac{k_3}{p_{H_2}^{3.5}} \left(p_{CH_4} p_{H_2O}^2 - \frac{p_{H_2}^4 p_{CO_2}}{K_3} \right) / DEN^2 \tag{4.5c}$$

$$DEN = 1 + \frac{K_{H_2O} p_{H_2O}}{p_{H_2}} + K_{CO} p_{CO} + K_{H_2} p_{H_2} + K_{CH_4} p_{CH_4} \tag{4.5d}$$

where K_{H_2}, K_{CH_4} and K_{CO} are adsorption constants for H_2, CH_4 and CO, K_{H_2O} is a dissociative adsorption constant of H_2O, K_1, K_2, and K_3 are equilibrium constants of the reactions in Eqs. 4.5a, 4.5b and 4.5c, k_1, k_2 and k_3 are kinetic constant coefficients of the reactions in Eqs. 4.5a, 4.5b, and 4.5c, respectively, *DEN* is a dimensionless parameter and p_{H_2}, p_{CH_4}, p_{H_2O}, p_{CO} and p_{CO_2} are the partial pressures of H_2, CH_4, H_2O, CO and CO_2 respectively in the tube-side mixture. This kinetic model accounts for the presence of the catalyst network, and it can be reformulated to also account for the external and internal diffusion resistances of the catalyst network by multiplying the formation and consumption rates of each individual species shown in Eq. 4.5 with a universal effectiveness factor of 0.1 Wesenberg and Svendsen (2007). However, unlike the global kinetic models of the methane and hydrogen combustion processes, the empirical kinetic formulas shown in Eq. 4.5 are not in the Arrhenius form, and thus, they cannot be directly implemented within the pilot-scale reformer CFD model. Nevertheless, *ANSYS Fluent* allows these non-Arrhenius form kinetic formulas to be incorporated into the CFD model by means of user-defined functions, i.e., *DEFINE_VR_RATE* and *DEFINE_NET_REACTION_RATE*, to simulate the formation and consumption rates of the individual components.

Inside the reforming tubes, the Reynolds number estimated based on the tube-side feed information shown in Table 4.4 indicates that the tube-side mixture is a turbulent flow. Therefore, a proper turbulence-chemistry interaction model is required to accurately simulate the individual species reaction rates in the turbulent reacting flow. It is worth noting that when the global kinetic model consists of three or more dependent elementary steps as does the global kinetic model of the SMR process as shown in Eq. 4.5, the CFD model implemented with the FR/ED model is likely to converge to an inaccurate solution Inc. (2013). This issue is remedied by means of an alternate chemistry-turbulence interaction model known as the eddy dissipation concept (EDC) model. The most important advancement offered by the EDC model

Table 4.4 Tube-side gas inlet operating conditions

Pressure (kPa)	3038.5
Temperature (K)	887
Flow rate (kg/s)	0.1161
x_{CH_4}	0.2487
x_{H_2O}	0.7377
x_{CO}	0.0001
x_{H_2}	0.0018
x_{CO_2}	0.0117

is that the detailed multi-step reaction mechanism can be employed to determine the formation and consumption rates of individual species in the turbulent reacting flow with a high level of accuracy. Additionally, the EDC model with default parameters is a robust turbulence-chemistry interaction model, and can be directly applied for a vast variety of reaction-limited and diffusion-limited systems without the need to adjust its parameters Magnussen (2005). When the EDC model is used to describe the chemistry and turbulence interactions taking place inside the tube-side mixture, the grids in the reforming domain are decomposed into a special reaction space, which is the collection of turbulence structures Inc. (2013). Each turbulence structure is treated as a perfect stirred tank reactor (PSTR) for which the characteristic dimension can be calculated as shown in Eq. 4.6a. Each PSTR is assumed to operate at constant pressure over an interval equal to that of the mean residence time (estimated by Eq. 4.6b) with the initial conditions taken from the corresponding local grid, and the reactions are governed by the chosen kinetic model (Eq. 4.5) Inc. (2013), Magnussen (2005). The characteristic dimension of the turbulence structure and the mean residence time are as follows:

$$\zeta^* = C_\zeta \left(\frac{\nu\epsilon}{k^2}\right)^{\frac{1}{4}} \tag{4.6a}$$

$$\tau^* = C_\tau \left(\frac{\nu}{\epsilon}\right)^{\frac{1}{2}} \tag{4.6b}$$

where ζ^* and τ^* are the characteristic dimension and mean residence time of the turbulence structure, $C_\zeta = 2.1377$ is the volume fraction constant, $C_\tau = 0.4082$ is the time scale constant, k and ϵ are the turbulence kinetic energy and dissipation rate, respectively, and ν is the kinematic viscosity of the turbulent reacting flow. At the end of each reacting time interval, the mass fractions of species i, Y_i^*, evaluated at the scale of the turbulence structures are used to compute the net formation and consumption rates of the corresponding species on grid scale as shown in Eq. 4.7:

$$R_i = \frac{\rho\,(\zeta^*)^2}{\tau^*\left[1-(\zeta^*)^3\right]}\left(Y_i^* - Y_i\right) \tag{4.7}$$

where R_i is the observed reaction rate of species i in the tube-side mixture which has the units of $\frac{\text{kmol}}{\text{m}^3\ \text{s}}$, and Y_i is the average mass fraction of species i. Lastly, it is important to note that because of the intensive computational requirement needed to simulate a detailed multi-step kinetic model, the EDC model should only be implemented when the FR/ED model predicts a highly inaccurate converged solution.

4.4.2 Porous Zone Design

In the reforming tubes, the tube-side feed composed of steam and methane enters at the ceiling of the pilot-scale reformer and passes through a wall-bounded, tightly packed catalyst network to be converted into hydrogen and carbon oxides. It is essential to the development of a pilot-scale reformer CFD model that the contribution of the catalyst network to the SMR process is well understood. Specifically, the catalyst network facilitates the formation of hydrogen from the naturally stable and slowly-reacting tube-side reactants, i.e., steam and methane, and it also enhances the rate of convective energy transfer from the inner reforming tube wall to the tube-side mixture by increasing the contact area. Additionally, the presence of the catalyst network inside the reforming tubes interferes with the tube-side flow, generating turbulence which homogenizes the tube-side mixture and prevents the tube-side mixture from quickly exiting, which ensures the SMR reactions Xu and Froment (1989) reach the desired equilibrium state. Furthermore, a large pressure difference between the tube-side mixture at the reforming tube inlet and outlet is expected due to the presence of the catalyst network inside the reforming tubes. Therefore, it is desirable to also model the catalyst network for the pilot-scale reformer CFD model in order to gain insight into the transport and reaction phenomena inside the reforming tube. Based on the known pressure difference, tube geometry and available intrinsic characteristics of the catalyst network as shown in Table 4.5, a set of modeling parameters required by the pilot-scale reformer CFD model to design a realistic representation of the catalyst network can be estimated through the semi-empirical Ergun equation Ergun and Orning (1949), which is suitable for a wide range of Reynolds numbers and various packing patterns:

$$\frac{\Delta P}{L} = \frac{150\mu}{D_p^2}\frac{(1-\gamma)^2}{\gamma^3}v_\infty + \frac{1.75\rho}{D_p}\frac{(1-\gamma)}{\gamma^3}v_\infty^2 \tag{4.8}$$

Table 4.5 Johnson Matthey's KATALCO 23 − 4*Q* catalyst properties

Density, ρ_c	3960 kg/m^3
Heat capacity, $C_{p,c}$	880 J/(kg · K)
Thermal conductivity, k_c	33 W/(m · K)
Particle diameter, D_p	3.5 mm (average)

where ΔP is the pressure difference between the tube-side mixture at the reforming tube inlet and outlet, L is the length of the reforming tubes, μ is the viscosity of the tube-side mixture, γ is the porosity of the catalyst network, v_∞ is the bulk velocity of the tube-side mixture, ρ is the density of the catalyst network, and $\frac{150\mu}{D_p^2}\frac{(1-\gamma)^2}{\gamma^3} = 8782800\ \frac{1}{\mathrm{m}^2}$ and $\frac{1.75\rho}{D_p}\frac{(1-\gamma)}{\gamma^3} = 1782\ \frac{1}{\mathrm{m}}$ are the viscous resistance coefficient and inertial resistance coefficient of the catalyst network. In this CFD model, the assumption of uniform packing is used for the catalyst network. Additionally, the coefficients of viscous resistance and inertial resistance of the catalyst network are assumed to be constant and uniform along the axial and radial directions.

4.5 Equation of State of Pilot-Scale SMR Unit

In the pilot-scale reformer, the furnace-side mixture can be assumed to possess ideal gas characteristics, and can be described by the ideal gas law, due to its temperature and pressure. It is important to note that the ideal gas equation of state can only be used to model the behavior of a real gas at high temperature and low pressure. The maximum temperature of the furnace-side mixture is approximately ~2000 K due to the thermal energy released by the combustion processes inside the furnace chamber. Additionally, the operating pressure of the combustion chamber is nearly at atmospheric pressure, and is ~25–28 times smaller than that of the reforming tubes Latham (2008). On the contrary, the thermodynamic behavior of the tube-side mixture deviates significantly from that governed by the ideal gas law due to the severely high operating pressure inside the reforming tubes, i.e., ~3,000 kPa. Moreover, the SMR process taking place inside the reforming tubes is driven in the desired direction primarily by the thermal energy generated by the oxidation of methane and hydrogen inside the combustion chamber. Furthermore, in the pilot-scale reformer CFD model, the Mach number of the tube-side feed is estimated to be ~0.3 based on the available plant information as shown in Table 4.4 Latham (2008), and therefore, the density variations of the tube-side flow due to the high operating pressure cannot be ignored. Hence, an appropriate equation of state must be incorporated into the pilot-scale reformer CFD model to accurately describe the thermodynamics and intrinsic properties of the turbulent tube-side reacting flows.

Due to the tightly coupled chemical, physical and transport phenomena, and the complex interactions between turbulence and the multi-step nonlinear kinetic models taking place inside the pilot-scale reformer, it is extremely challenging to obtain a converged solution of the model. Additionally, simulating the pilot-scale reformer CFD model shown in Fig. 4.5 composed of 165,817 hexahedral grids, 516,325 quadrilateral cells and 184,636 nodes is expected to require large computing power and a long computation time to find a converged solution. Therefore, it is not practical to use the pilot-scale reformer CFD model as a means to obtain the most suitable equation of state that can account for the aforementioned considerations.

Table 4.6 Mesh quality of the industrial-scale reforming tube mesh

	Industrial-scale reforming tube mesh's value	Ideal value	Recommended value
Minimum orthogonal factor	1.000	1.000	0.1667
Maximum ortho skew	4.281×10^{-7}	0.000	0.8500
Maximum aspect ratio	9.253	1.000	100.0

From our previous work, we successfully developed the CFD model of an industrial-scale reforming tube that can reproduce publicly available industrial plant data given the plant tube-side feed and outer reforming tube wall temperature Lao et al. (2016). The industrial-scale reforming tube mesh contains 23,030 quadrilateral cells and 24,460 nodes, and is considerably smaller than the pilot-scale reformer mesh, i.e., the number of quadrilateral cells inside the former mesh is approximately 5 % of that inside the latter mesh. Additionally, the mesh quality of the industrial-scale reforming tube mesh is nearly ideal based on the three suggested criteria including the orthogonal factor, aspect ratio and ortho skew as shown in Table 4.6. As a result, the CFD model of the industrial-scale reforming tube allows a large reduction in the computation time required to determine a suitable equation of state for the pilot-scale reformer CFD model.

A rigorous study is conducted to select the most suitable equation of state for the CFD model of the pilot-scale reformer by means of the industrial-scale reforming tube CFD model. In this effort three potential candidates, i.e., the compressible ideal gas, the real gas Soave-Redlich-Kwong (*SRK*) and the real gas Peng-Robinson (*PR*) equations of state, are selected based on the aforementioned modeling challenges and considerations, and the corresponding simulation results are shown in Table 4.7. The differences between the simulation results generated by the CFD model implemented with the compressible ideal gas equation of state and that generated by the CFD model implemented with the more sophisticated equations of state (*SRK* and *PR*) are negligible. Specifically, in the compositions of the tube-side mixture at the reforming tube outlet, the differences occur at the fifth digit after the decimal point. Additionally, the CFD model implemented with the compressible ideal gas equation of state requires less computation time to find the converged solution than the other CFD models with the more complex equations of state, and the simulation results presented in Table 4.7 show that it yields approximately the same converged solution. As a result, the compressible ideal gas equation of state is adopted in the pilot-scale reformer CFD model to describe the thermodynamics and intrinsic properties of both the tube-side and furnace-side mixtures.

Table 4.7 Simulation results of industrial-scale reforming tube with the selected equations of state

Equation of state	Compressible ideal gas	Real gas SRK	Real gas PR
ΔP (kPa)	212.83	213.99	213.60
$\bar{P}_{outlet}$ (kPa)	3044.0	3044.0	3044.0
Average Heat Flux $\left(\frac{\text{kW}}{\text{m}^2}\right)$	68.972	69.380	69.391
$\bar{x}_{H_2}^{outlet}$	0.464485	0.464493	0.464492
$\bar{x}_{H_2O}^{outlet}$	0.346720	0.346716	0.346717
$\bar{x}_{CH_4}^{outlet}$	0.042605	0.042607	0.042608
$\bar{x}_{CO}^{outlet}$	0.087348	0.087347	0.087347
$\bar{x}_{CO_2}^{outlet}$	0.058842	0.058837	0.058837

4.6 Turbulence-Chemistry Interaction Model of Pilot-Scale SMR Unit

The discussion presented in Sect. 4.3.1 demonstrates that the FR/ED turbulence-chemistry interaction model and premixed combustion model can successfully describe the characteristics of the non-premixed combustion process, i.e., the overall reaction rates of the individual species depend only on the rate of mixing. Additionally, the global kinetic models chosen for the combustion of hydrogen and natural gas are compatible with the FR/ED model, and as a result, the FR/ED model is chosen as the turbulence-chemistry interaction model for the furnace-side mixture. However, the SMR process consists of three catalytic reversible reactions driven to yield the desired product, hydrogen, by the thermal energy generated from the combustion chamber, and therefore, the assumption of fast reaction rates is not applicable, so the FR/ED model may not be accurate. In Sect. 4.4.1, we have demonstrated that the EDC model is an alternative turbulence-chemistry interaction model with the ability to account for multi-step reaction mechanisms. Nevertheless, the computational requirement is demanding, and therefore, it should only be used as the last resort if the FR/ED model yields unacceptably incorrect simulation results. It is noteworthy that the accuracy of the FR/ED model is questioned only for the reforming tubes, and therefore, the CFD model of the industrial-scale reforming tube provides an effective means (based on the reasons stated in Sect. 4.5) for choosing the most suitable turbulence-chemistry interaction model for the reforming tubes of the pilot-scale reformer CFD model.

In this effort, two potential turbulence-chemistry interaction models, i.e., the EDC model and the FR/ED model, are implemented in the CFD model of the industrial-scale reforming tube. It is important to note that for this study using the CFD model of the industrial-scale reforming tube, all boundary conditions, i.e., the outer reforming tube wall temperature and the tube-side feed conditions, and other simulation settings, are the same as in the pilot-scale reforming tubes. The comparison between

Table 4.8 Simulation results of the industrial-scale reforming tube with the selected turbulence-chemistry interaction models

Turbulence-chemistry interaction model	EDC	FR/ED	Difference (%, with respect to EDC results)
ΔP (kPa)	208.3	212.8	2.2
$\bar{P}_{outlet}$ (kPa)	3044.0	3044.0	0.0
Average Heat Flux $\left(\frac{\text{kW}}{\text{m}^2}\right)$	66.484	68.942	3.7
$\bar{x}_{H_2}^{outlet}$	0.453	0.464	2.4
$\bar{x}_{H_2O}^{outlet}$	0.356	0.347	2.4
$\bar{x}_{CH_4}^{outlet}$	0.049	0.043	12
$\bar{x}_{CO}^{outlet}$	0.082	0.087	6.0
$\bar{x}_{CO_2}^{outlet}$	0.060	0.059	1.9

the simulation results generated by the two turbulence-chemistry models is shown in Table 4.8, where the differences (%) are computed by dividing the deviation between data extracted from the two sets of simulation results by the corresponding values generated by the CFD model implemented with the EDC model.

The differences between the simulation result generated by the CFD model implemented with the FR/ED model and the result generated by the CFD model implemented with the EDC model are negligible as shown in Table 4.8. Therefore, the errors associated with the FR/ED model are determined to be insignificant compared to the reduction in the required computation time. Specifically, the former CFD model takes ~2000 iterations to determine the converged solution while the latter CFD model takes ~6500 iterations. As a result, the FR/ED model is adopted in the pilot-scale reformer CFD model to describe the reaction rates of individual species in the turbulent reacting flows of both the tube-side and furnace-side mixtures.

4.7 Governing Equation of Pilot-Scale SMR Unit

Based on the inlet information of the furnace-side feed (Sect. 4.3.1) and tube-side feed (Sect. 4.4.1), the flow profiles inside the combustion chamber and reforming tubes are speculated to be turbulent. Additionally, the catalyst network inside the reforming tubes plays the role of a physical obstacle interfering with the tube-side mixture flow, which generates turbulent flow patterns and enhances mixing for the tube-side mixture. Therefore, it is essential that a proper turbulence model is implemented within the CFD model of the pilot-scale reformer in order to accurately capture the characteristic parameters of the turbulent reacting flow profiles, i.e., the turbulence kinetic energy and turbulence dissipation rate; as a result, a standard

$k-\epsilon$ turbulence model is adopted. This turbulence model, presented in Eqs. 4.9e and 4.9f Jones and Launder (1972), Launder and Sharma (1974), is developed from the Reynolds-averaged Navier-Stokes (RANS) equations and can yield reasonably accurate estimates of characteristic parameters of the turbulent reacting flow, while it is also computationally inexpensive compared to other turbulence models, i.e., the realizable $k-\epsilon$ and RNP $k-\epsilon$ models. In the remainder of the section, the equations used to describe the complex physical phenomena within the combustion chamber and reforming tubes are described.

4.7.1 Furnace Chamber

In the combustion chamber, predicting the temperature field remains the most difficult modeling task because of the coupling between the mass, momentum and energy transport phenomena, the exothermic oxidation of the fuel stream and the thermal coupling of the combustion chamber with the reforming tubes. Specifically, these processes must be accurately modeled in order to produce a realistic temperature field, which can only be obtained by solving for the steady-state solution of the complete reformer model. Based on the above considerations and those discussed in Sects. 4.3, 4.5 and 4.6, the governing equations including the continuity equation (Eq. 4.9a) and the momentum (Eq. 4.9b), energy (Eq. 4.9c) and species (Eq. 4.9d) balances, and the turbulence model (Eqs. 4.9e and 4.9f) required for characterizing the heat and fluid-flow fields inside the combustion chamber are formulated. It is worth recalling that due to the mixing-limited nature of the combustion process, the detailed combustion kinetic models of hydrogen and methane, i.e., models including kinetics for free radicals, are rendered unnecessary for the purpose of this work because the formation/consumption rates of all furnace-side species are directly controlled by the rate of turbulent mixing, which in turn justifies the implementation of computationally efficient global kinetic models for these processes. It is important to realize that in the turbulent reacting fluid flow, velocity components, pressure, energy, temperature, density and species concentration fluctuate about their corresponding time-averaged values, and therefore, their instantaneous values are not well-defined. As a result, the governing equations of the furnace-side are written in terms of time-averaged variables as follows:

Continuity equation:

$$\frac{\partial}{\partial t}(\rho_{comb}) + \nabla \cdot (\rho_{comb}\vec{v}_{comb}) = 0 \tag{4.9a}$$

Momentum conservation equation:

$$\frac{\partial}{\partial t}(\rho_{comb}\vec{v}_{comb}) + \nabla \cdot (\rho_{comb}\vec{v}_{comb}\vec{v}_{comb}) = -\nabla P_{comb} + \nabla \cdot \bar{\bar{\tau}}_{comb} + \rho_{comb}\vec{g} \tag{4.9b}$$

Energy conservation equation:

$$\frac{\partial}{\partial t}(\rho_{comb}E_{comb}) + \nabla \cdot (\vec{v}_{comb}(\rho_{comb}E_{comb} + P_{comb})) =$$
$$\nabla \cdot \left[k^{eff}_{comb} \nabla T_{comb} - \left(\sum_i h^i_{comb} \vec{J}^i_{comb} \right) + \left(\overline{\overline{\tau}}_{comb} \cdot \vec{v}_{comb} \right) \right]$$
$$+ S^{h,rxn}_{comb} + \nabla \cdot \vec{q}_{rad} \tag{4.9c}$$

Species material conservation equation:

$$\frac{\partial}{\partial t}(\rho_{comb}Y^i_{comb}) + \nabla \cdot (\rho_{comb}\vec{v}_{comb}Y^i_{comb}) = -\nabla \cdot (\vec{J}^i_{comb}) + R^i_{comb} \tag{4.9d}$$

Transport equations of the standard$k-\epsilon$turbulence model:

$$\frac{\partial}{\partial t}(\rho_{comb}k_{comb}) + \nabla \cdot (\rho_{comb}k_{comb}\vec{v}_{comb}) =$$
$$\nabla \cdot \left[\left(\mu_{comb} + \frac{\mu^t_{comb}}{\sigma_k} \right) \nabla k_{comb} \right] + G^k_{comb} + G^b_{comb} - \rho_{comb}\epsilon_{comb} \tag{4.9e}$$

$$\frac{\partial}{\partial t}(\rho_{comb}\epsilon_{comb}) + \nabla \cdot (\rho_{comb}\epsilon_{comb}\vec{v}_{comb}) = \tag{4.9f}$$
$$\nabla \cdot \left[\left(\mu_{comb} + \frac{\mu^t_{comb}}{\sigma_\epsilon} \right) \nabla \epsilon_{comb} \right] + C_{1\epsilon} \frac{\epsilon_{comb}}{k_{comb}} G^k_{comb} - C_{2\epsilon}\rho_{comb} \frac{\epsilon^2_{comb}}{k_{comb}}$$

$$\vec{J}^i_{comb} = \left(\rho_{comb} D^{m,i}_{comb} + \frac{\mu^t_{comb}}{Sc^t_{comb}} \right) \nabla Y^i_{comb} - D^{T,i}_{comb} \frac{\nabla T_{comb}}{T_{comb}} \tag{4.10a}$$

$$\overline{\overline{\tau}}_{comb} = \mu_{comb} \left[\left(\nabla \vec{v}_{comb} + \nabla \vec{v}^T_{comb} \right) - \frac{2}{3} \nabla \cdot \vec{v}_{comb} I \right] \tag{4.10b}$$

$$E_{comb} = h_{comb} + \frac{v^2_{comb}}{2} - \frac{P_{comb}}{\rho_{comb}} \tag{4.10c}$$

$$h_{comb} = \sum_j Y^j_{comb} h^j_{comb} \tag{4.10d}$$

$$h^j_{comb}(T_{comb}) = \int_{T_{ref}}^{T_{comb}} C^j_{p,comb} dT \text{ with } T_{ref} = 298.15\ K \tag{4.10e}$$

$$S^{h,rxn}_{comb} = -\sum_j \frac{h^f_j}{M_j} R^j_{comb} \tag{4.10f}$$

$$R^j_{comb} = \sum_k^3 \nu^{k,j}_{comb} R^{k,j}_{comb} \tag{4.10g}$$

$$k^{eff}_{comb} = k^l_{comb} + k^t_{comb} \tag{4.10h}$$

$$k^t_{comb} = \frac{C_p \mu^t_{comb}}{Pr_t} \tag{4.10i}$$

$$\mu_{comb}^{t} = \rho_{comb} C_{\mu} \frac{k_{comb}^{2}}{\epsilon_{comb}} \tag{4.10j}$$

$$G_{comb}^{k} = -\rho_{comb} \overline{v'_{comb,i} v'_{comb,j}} \frac{\partial v_{comb,j}}{\partial x_i} \tag{4.10k}$$

$$G_{comb}^{b} = \beta g_i \frac{\mu_{comb}^{t}}{Pr_t} \frac{\partial T_{comb}}{\partial x_i} \tag{4.10l}$$

where ρ_{comb}, $\vec{v}_{comb}$, μ_{comb}, μ_{comb}^{t}, k_{comb}^{eff}, k_{comb}^{l}, k_{comb}^{t}, C_p, T_{comb} and P_{comb} are the furnace-side mixture density, mass-averaged velocity, laminar mixture viscosity, turbulent mixture viscosity (calculated as shown in Eq. 4.10j), effective thermal conductivity (estimated as shown in Eq. 4.10h), laminar thermal conductivity, turbulent thermal conductivity (estimated as shown in Eq. 4.10i), heat capacity, temperature and pressure of the furnace-side mixture in the combustion chamber, respectively, $\vec{g}$ is the universal gravitational acceleration, $\overline{\overline{\tau}}_{comb}$ is the stress tensor (estimated as shown in Eq. 4.10b), and I is the unit tensor. It is noteworthy that the aforementioned governing equations are capable of characterizing the turbulent reacting furnace-side flow under the influence of complex interactions of mass, momentum and energy transport, reaction and turbulence, and also accounting for the major microscopic/macroscopic phenomena that yield the heat and fluid-flow fields inside the combustion chamber. Specifically, the turbulent mass diffusion flux of species i, $\vec{J}_{comb}^{i}$, driven by concentration and temperature gradients, is shown in Eq. 4.10a, where Y_{comb}^{i}, $D_{comb}^{m,i}$ and $D_{comb}^{T,i}$ are the furnace-side mass fraction, laminar mass diffusion coefficient and laminar thermal diffusion coefficient of species i. It is necessary to note that the ratio $\frac{\mu_{comb}^{t}}{Sc_{comb}^{t}}$, in which Sc_{comb}^{t} and μ_{comb}^{t} are the turbulent Schmidt number and turbulent viscosity of the furnace-side mixture, is used to account for the effect of turbulence on the mass diffusion flux of species i, and therefore, it can be written as $\rho_{comb} D_{comb}^{m,t}$ where $D_{comb}^{m,t}$ is the turbulent mass diffusion coefficient. Additionally, the specific internal energy (E_{comb}) of the furnace-side mixture which can be computed as the sum of the furnace-side specific sensible enthalpy (h_{comb}) which depends on the furnace-side specific sensible enthalpy of species j at temperature T_{comb} ($h_{comb}^{j}(T_{comb})$), specific kinetic energy ($v_{comb}^{2}/2$) and external work per unit weight of the furnace-side mixture ($-P_{comb}/\rho_{comb}$), is shown in Eqs. 4.10c, 4.10d and 4.10e. It is important to note that the value of $T_{ref} = 298.15\ K$ in Eq. 4.10e is chosen automatically by *ANSYS Fluent*'s parallel/pressure based solver, and $C_{p,comb}^{j}$ is the heat capacity of species j in the combustion chamber. In addition, from Eq. 4.9c, $\nabla \cdot (k_{comb}^{eff} \nabla T_{comb})$, $-\nabla \cdot \left(\sum_i h_{comb}^{i} \vec{J}_{comb}^{i} \right)$, $\nabla \cdot (\overline{\overline{\tau}}_{comb} \cdot \vec{v}_{comb})$ and $\nabla \cdot \vec{q}_{rad}$ represent four distinct mechanisms, i.e., conduction, species diffusion, viscous dissipation and radiation respectively, through which energy is transferred. Furthermore, the overall rate at which thermal energy is generated from combustion processes inside the combustion chamber, $S_{comb}^{h,rxn}$, is computed as shown in Eq. 4.10f in which R_{comb}^{j} and h_{j}^{f} represent the overall volumetric consumption/formation rate and enthalpy of formation

of species j, and $\nu_{comb}^{k,j}$ and $R_{comb}^{k,j}$ are the stoichiometric coefficient and volumetric consumption/formation rate of species j in reaction k. It is noteworthy that $R_{comb}^{k,j}$ is determined by the FR/ED turbulence-chemistry interaction model (Sect. 4.3.1). Transport equations of the standard $k-\epsilon$ turbulence model are presented in Eqs. 4.9e and 4.9f, in which k_{comb} and ϵ_{comb} are the turbulence kinetic energy and turbulence dissipation rate of the furnace-side mixture, β is the coefficient of thermal expansion of the furnace-side mixture, $\sigma_k = 1.3$ and $\sigma_\epsilon = 1.0$ are the default values of the turbulent Prandtl numbers for k_{comb} and ϵ_{comb}, $C_{1\epsilon} = 1.44$, $C_{2\epsilon} = 1.92$, $C_\mu = 0.09$ and $Pr_t = 0.85$ are default constants of the standard $k-\epsilon$ turbulence model, respectively, and G_{comb}^k and G_{comb}^b represent the generation of turbulence kinetic energy in the furnace-side mixture due to the mean velocity gradients (Eq. 4.10k) and buoyancy effect (Eq. 4.10l). We would like to point out that the standard $k-\epsilon$ turbulence model can capture the characteristic parameters of turbulent reacting flow profiles. Specifically, in Eq. 4.10k, the term $-\rho_{comb}\overline{v'_{comb,i}v'_{comb,j}}$ is the Reynolds stress representing the effect of turbulence on the velocity profile of the furnace-side mixture that arises from the Reynolds-averaged Navier-Stokes (RANS) equations, and $\overline{v'_{comb,i}}$ is the time-averaged fluctuating component of $\vec{v}_{comb}$ in the x_i direction. It is worth noting that all default constants of the standard $k-\epsilon$ turbulence model are determined empirically by experiments for fundamental turbulent flows, and have been proven to be suitable for a wide range of wall-bounded and free shear flow applications Inc. (2013).

4.7.2 Reforming Tube

In the reforming tubes, the most challenging modeling task is to accurately simulate the formation/consumption rates of the tube-side species in turbulent flow as they undergo a series of microscopic processes. Specifically, the reactants including methane and steam from the bulk tube-side mixture are convectively driven to the stationary layer around the catalyst network by reactant concentration gradients generated by the external mass transfer resistance of the catalyst network. Then, these reactants are diffusively driven by the reactant concentration gradient generated by the internal mass transfer resistance of the catalyst network and by temperature gradients generated by the SMR process to the catalyst active sites, where the aforementioned endothermic reactions take place generating hydrogen and byproducts. These products diffuse back out of the catalyst and stationary layer to reenter the bulk tube-side mixture. It is noteworthy that the effects of the internal and external mass transfer resistances and the presence of the catalyst network on the tube-side mixture are taken into account in the tube-side governing equations of this CFD model by the porosity, viscous resistance coefficient, inertial resistance coefficient of the catalyst network and the universal effectiveness factor (as mentioned in Sects. 4.4.1 and 4.4.2). Based on the above considerations and those discussed in Sects. 4.4, 4.5 and 4.6, the governing equations including the continuity equation and the momentum,

energy and tube-side species balances, and the turbulence model required to simulate the SMR process inside the reforming tubes, are constructed in a similar manner to that of the combustion chamber, which has been described in Sect. 4.7.1. Additionally, the governing equations of the tube-side mixture have also been explicitly presented in our recent publication Lao et al. (2016), and therefore, they are not repeated here for brevity.

4.8 Simulation Results

The solution of the pilot-scale reformer CFD model is computed in ~24 h by *ANSYSFluent*'s parallel solver with 48 cores on UCLA's Hoffman2 Cluster for which an arbitrarily chosen core is designated as a host process, and the remaining 47 cores are compute-node processes. Under *ANSYSFluent*'s parallel solver, the pilot-scale reformer mesh is partitioned into 47 parts for which each partition is assigned to a different compute-node process. It is noteworthy that the host process does not participate in the computing task to determine the converged solution. Instead, it is responsible for interpreting commands given in the graphical user interface (GUI), then redistributing them to all compute-node processes by a message-passing library, e.g., the Message Passing Interface (MPI).

In the remainder of this section, the steady-state simulation results of the pilot-scale reformer CFD model are presented. Based on the inlet information for the furnace-side feed (Sect. 4.3) and tube-side feed (Sect. 4.4), the steady-state furnace-side temperature field and oxidized species distributions are shown in Figs. 4.10, 4.11, 4.12 and 4.13. Additionally, the corresponding tube-side temperature, pressure, and major species distributions are presented in Figs. 4.14, 4.15, 4.16, 4.17 and 4.18. We would like to point out that due to the high heated reforming tube length to reforming tube diameter ratio, ~85:1, the radial direction is scaled up by 20 times in Figs. 4.14, 4.15, 4.16, 4.17 and 4.18 for display purposes only. It is worth noting that the radially uniform pressure profile inside the reforming tubes is the result of the uniformly packed catalyst network assumption, and the definition of the porous zone with uniform coefficients of viscous resistance and inertial resistance of the catalyst network along the axial and radial directions as presented in Sect. 4.4.2. Lastly, the average temperature profiles of the outer and inner reforming tube wall and flue gas in the vicinity of the reforming tubes is shown in Fig. 4.19, and the corresponding steady-state heat flux profile through the reforming tube wall is displayed in Fig. 4.20. Figure 4.19 suggests that the maximum temperature of the outer reforming tube wall of 1177.48 K is below the maximum allowable operating temperature, which if the maximum temperature of the outer reforming tube wall were to exceed for a sufficient length of time, the reforming tube would rupture. The presented simulation results of the pilot-scale reformer CFD model are consistent with the publicly available plant data Latham (2008), many sources in the literature Rostrup-Nielsen (1984), Dybkjaer (1995), Froment and Bischoff (1990), Pantoleontos et al. (2012) and well-established knowledge about the SMR process,

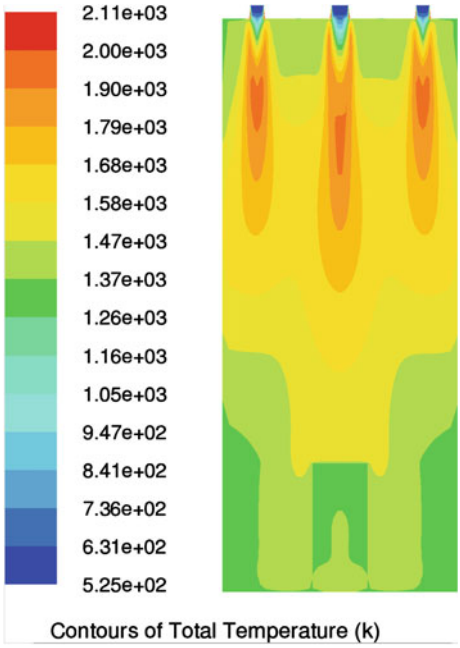

Fig. 4.10 Temperature field inside the combustion chamber from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

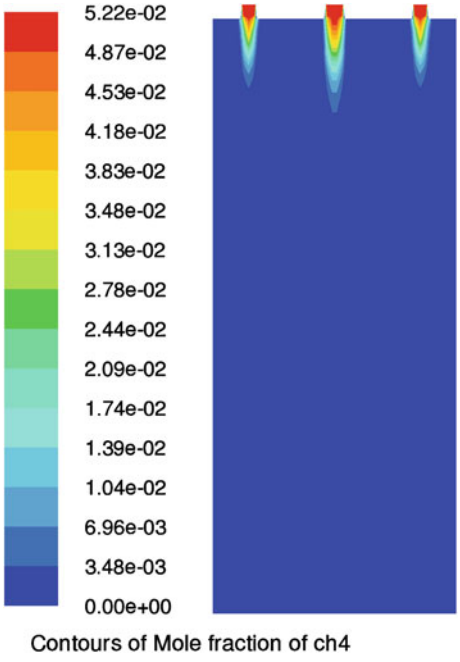

Fig. 4.11 Methane mole fraction profile inside the combustion chamber from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

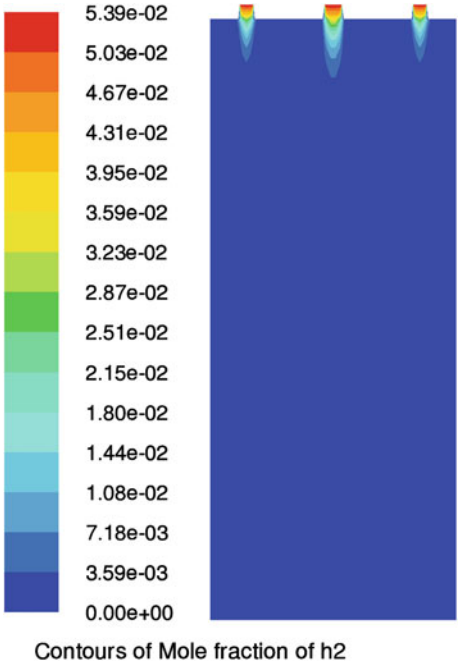

Fig. 4.12 Hydrogen mole fraction profile inside the combustion chamber from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

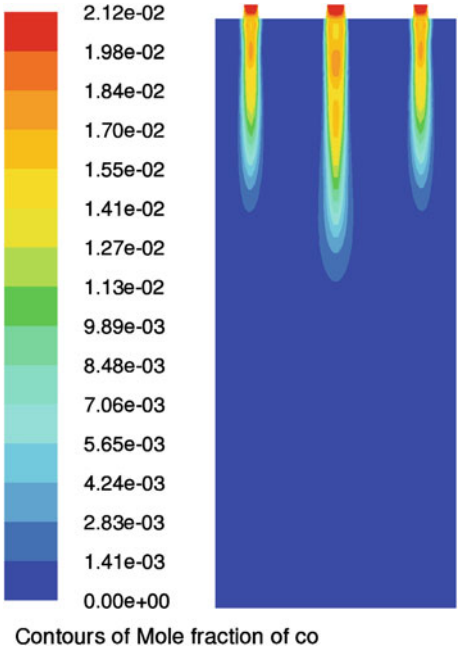

Fig. 4.13 Carbon monoxide mole fraction profile inside the combustion chamber from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

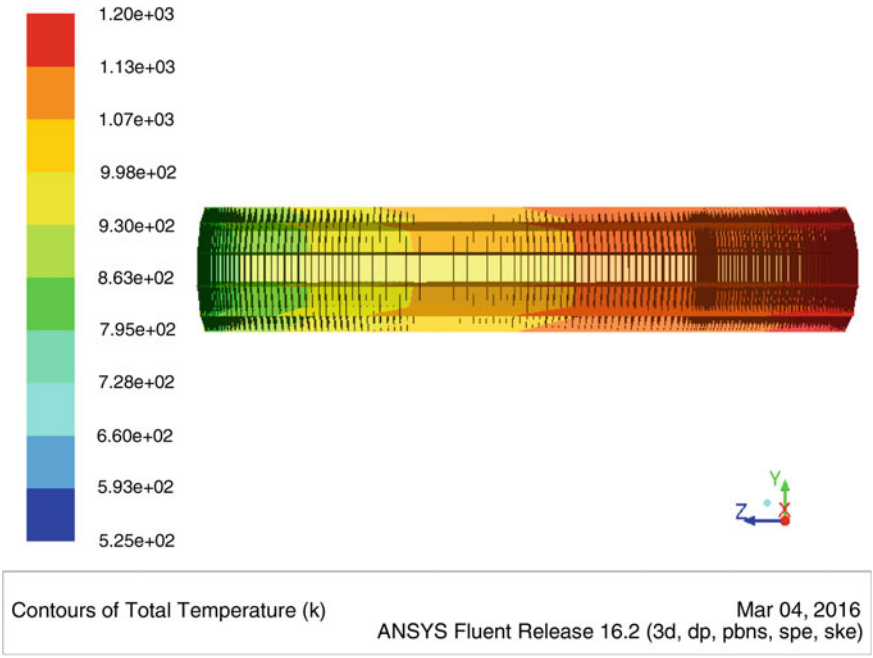

Fig. 4.14 Temperature profile inside the reforming tube from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

which supports the validity of the assumptions that we have made while constructing this CFD model. For instance, inside the combustion chamber, the furnace-side feed is fully oxidized to generate the necessary thermal energy required to drive the reversible endothermic SMR reactions in order to produce the desired hydrogen. The kinetic models for the chemical processes inside the combustion chamber required to construct this pilot-scale reformer CFD model were selected based on the assumption that methane, hydrogen and carbon monoxide are completely consumed prior to exiting the reformer. Figures 4.11, 4.12 and 4.13 demonstrate that the mole fractions of these species are effectively zero outside of the flame region, and as a result, the assumption of a lean combustion process is validated. Additionally, it is important to note that the outer-lane burners and inner-lane burner have the estimated lengths of 4 and 6 m, respectively, which are visible in Fig. 4.21. These results also match well with those observed in industry. Furthermore, the highest temperature inside the pilot-scale reformer is located at 1/3 of the combustion chamber's height from the ceiling as shown in Figs. 4.10, 4.11, 4.12, 4.13, 4.14, 4.15, 4.16, 4.17, 4.18 and 4.19, which is consistent with that observed and reported in industry.

Furthermore, the validation of the simulation results generated by the pilot-scale reformer CFD model is performed by comparing its tube-side mixture data at the reforming tube outlet to that generated by the industrial-scale reforming tube CFD

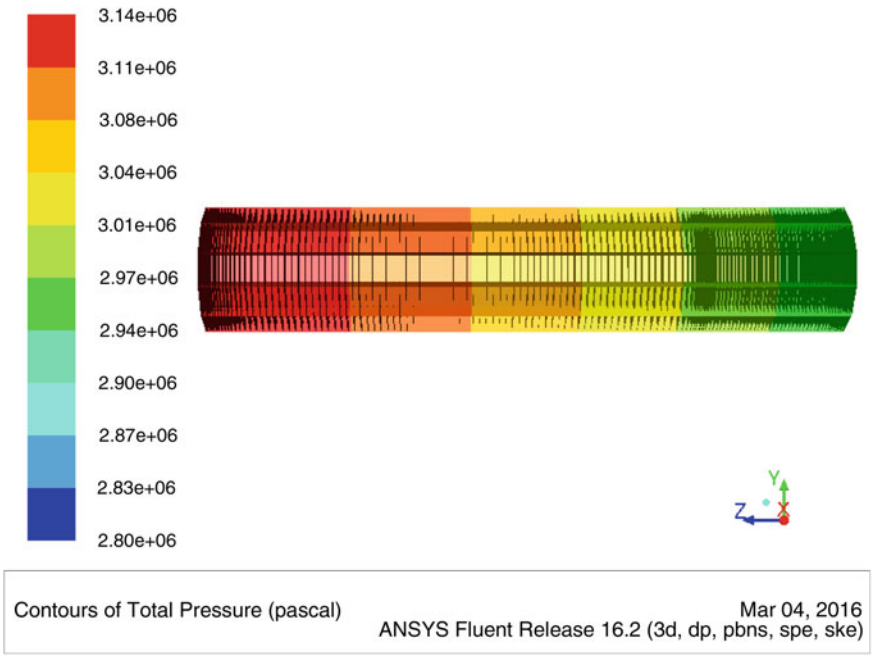

Fig. 4.15 Pressure profile inside the reforming tube from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

model as shown in Table 4.9. The industrial-scale reforming tube CFD model is constructed with the same modeling parameters as described in Sects. 4.2, 4.4, 4.5 and 4.6, and implemented with the same tube-side inlet conditions and outer reforming tube wall profile as those in the pilot-scale reformer CFD model Lao et al. (2016). We found that the differences between the simulation results generated by the industrial-scale reforming tube CFD model and the pilot-scale reformer CFD model are negligible. Finally, the simulation results generated by this pilot-scale reformer are further justified by comparing the area-weighted average heat flux across the reforming tube wall with that reported in the literature as shown in Table 4.10. In addition to the average heat flux results reported in this table, the average heat flux for the reformer described by Latham (2008) is estimated to be ~65.6 kW/m^2 based on the outer and inner reforming tube wall temperature profiles, the reforming tube thermal conductivity of 106,500 $\frac{\text{J}}{\text{m}\cdot\text{h}\cdot\text{K}}$ and the reforming tube wall thickness of 0.015 m, which is an average heat flux that is consistent with the results of the pilot-scale CFD model and thus supports the accuracy of the pilot-scale reformer CFD model. The heat flux profile of the pilot-scale CFD model and the average heat flux are shown in Fig. 4.20.

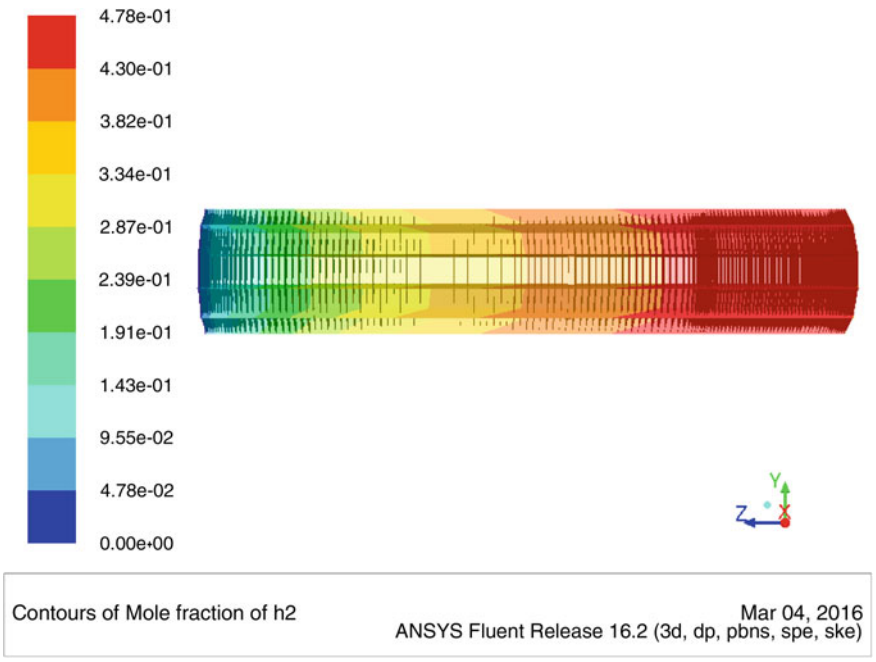

Fig. 4.16 Hydrogen mole fraction profile inside the reforming tube from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

4.9 Sample Case Study

We are primarily interested in developing a pilot-scale reformer CFD model that can be employed for reformer parametric study which has been traditionally conducted by on-site experiments. The necessity and motivation for reformer parametric study has been discussed in great depth in Sect. 4.1, and specifically, it allows chemical process plants to determine the optimized and more profitable operating conditions, which result in an increase of profit margin for the plant. Nevertheless, the traditional on-site parametric study must be conducted over a long time period with small changes in the parameter of interest throughout time in order to avoid unit failure which can result in significant production and capital losses (Sect. 4.1). On the other hand, it is noteworthy that a well-developed CFD model of a reformer can generate simulation data that reasonably resembles data that would be collected from on-site experiments, and this idea in turn motivates us to construct a CFD model of a pilot-scale reformer.

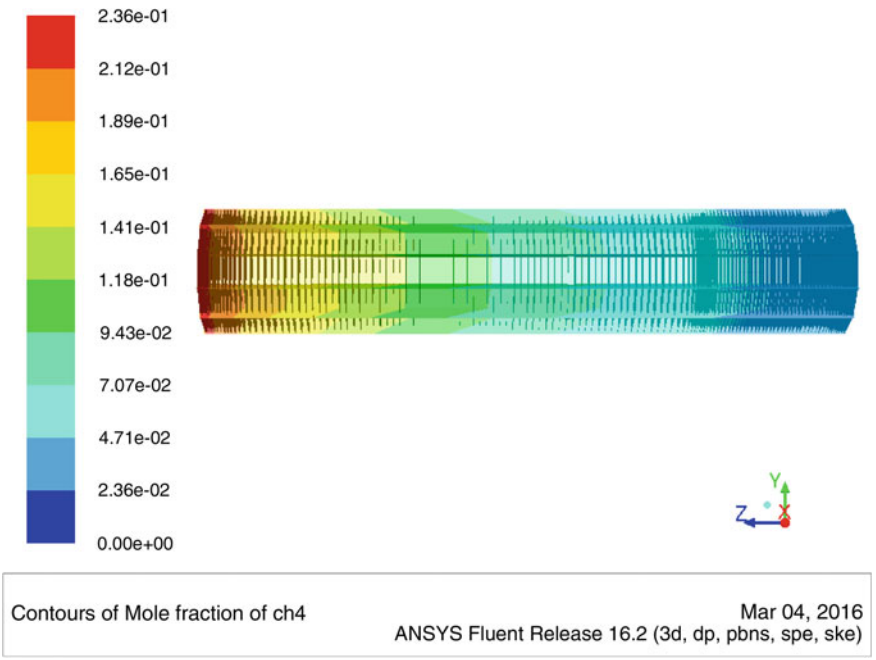

Fig. 4.17 Methane mole fraction inside the reforming tube from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

This task is successfully accomplished by the detailed analysis of physical phenomena, i.e., the transport of momentum, material, and energy in addition to turbulence, and chemical reactions taking place inside the reformer and furnace, and allows us to select the most suitable models, e.g., the $P-1$ radiation model (Sect. 4.3.2), standard $k-\epsilon$ turbulence model, compressible ideal gas equation of state (Sect. 4.5) and FR/ED turbulence-chemistry interaction model (Sect. 4.6), to simulate the macroscopic and microscopic events in the reformer. Simulation data generated by the resulting pilot-scale reformer CFD model is consistent with that of publicly available plant data reported in the literature and also with the simulation data generated by a well-developed industrial-scale single reforming tube CFD model as discussed in Sect. 4.8. As a result, this pilot-scale reformer CFD model is adopted for a reformer parametric study of the burner furnace-side feed.

In this effort, a 20 % increase in the furnace-side feed mass flow rate of each burner is implemented, which results in changes in the furnace-side feed mass flow rates of the inner-lane burner and outer-lane burners from 0.4056 $\frac{\text{kg}}{\text{s}}$ to 0.4867 $\frac{kg}{\text{s}}$, and from

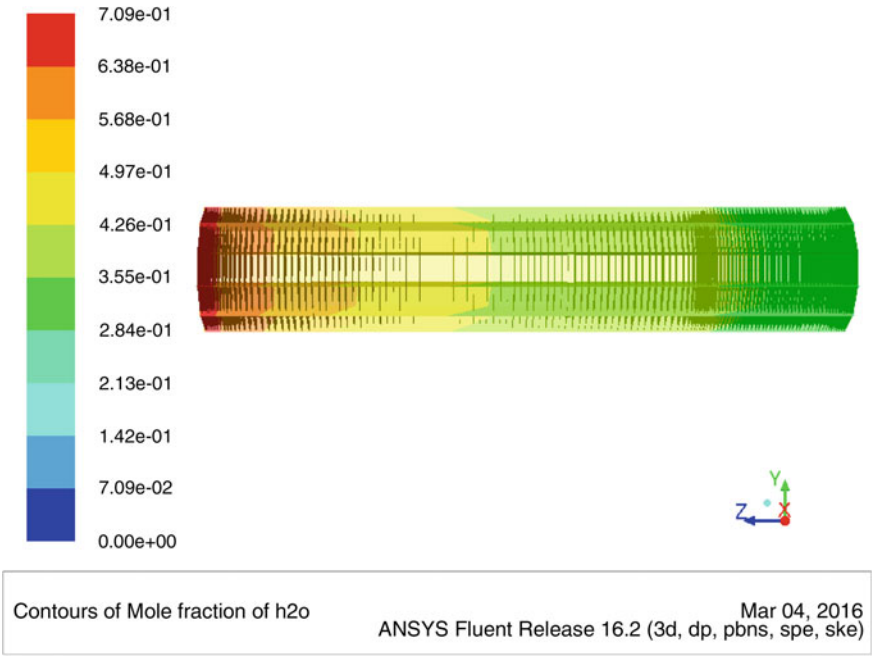

Fig. 4.18 Superheated steam mole fraction profile inside the reforming tube from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

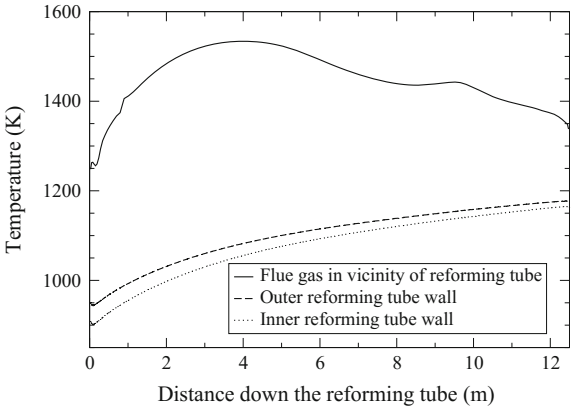

Fig. 4.19 Temperature profiles of flue gas in the vicinity of the reforming tube (*solid line*), of the outer reforming tube wall (*dashed line*) and of the inner reforming tube wall (*dotted line*) from the pilot-scale reformer CFD simulation, where the parameters of tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

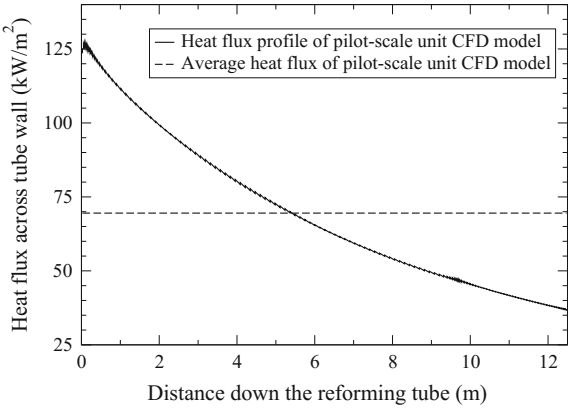

Fig. 4.20 Heat flux profile (*solid line*), and average heat flux (*dashed line*) across the reforming tube wall from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

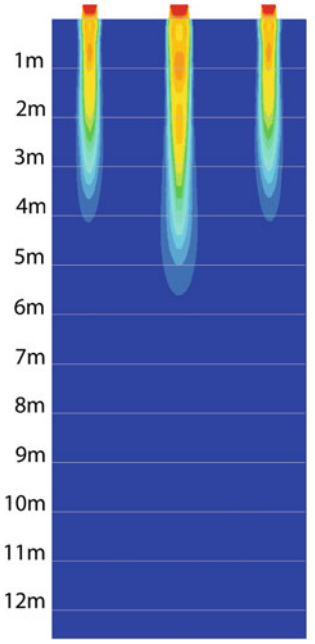

Fig. 4.21 Flame length inside the combustion chamber from the pilot-scale reformer CFD simulation, where the parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

0.2434 $\frac{\text{kg}}{\text{s}}$ to 0.2920 $\frac{\text{kg}}{\text{s}}$, respectively. During the parametric study of the furnace-side feed mass flow rate, the temperature, pressure and composition of the furnace-side feed, and all parameters of the tube-side feed and of the combustion chamber refractory walls, are kept constant, so that the effect of increasing the furnace-side feed mass flow rate can be identified and quantified. The simulation results generated

Table 4.9 Validation of pilot-scale reformer CFD model by the industrial-scale reforming tube CFD model

	Industrial-scale reforming tube CFD model (∗)	Pilot-scale unit CFD model	Difference (%, with respect to (∗))
ΔP (kPa)	212.83	201.73	5.215
$\bar{P}_{outlet}$ (kPa)	3044.0	2942.94	3.320
Average heat flux $\left(\frac{\text{kW}}{\text{m}^2}\right)$	68.972	69.523	0.799
$\bar{x}_{H_2}^{outlet}$	0.4645	0.4625	0.431
$\bar{x}_{H_2O}^{outlet}$	0.3467	0.3483	0.461
$\bar{x}_{CH_4}^{outlet}$	0.0426	0.0436	2.347
$\bar{x}_{CO}^{outlet}$	0.0873	0.0867	0.687
$\bar{x}_{CO_2}^{outlet}$	0.0588	0.0588	0.000

Table 4.10 Validation of pilot-scale reformer CFD model by available plant data from literature

	Average heat flux $\left(\frac{\text{kW}}{\text{m}^2}\right)$
Pilot-scale reformer CFD model	69.5
Industrial-scale reforming tube CFD model Lao et al. (2016)	69.0
Rostrup-Nielsen (1984)	45–90
Dybkjaer (1995)	78.5
Froment and Bischoff (1990)	75.6
Pantoleontos et al. (2012)	< 80.0

by the pilot-scale reformer CFD model indicate that the maximum outer reforming tube wall temperature exceeded the maximum allowable operating temperature of the reforming tube as shown in Fig. 4.22. This method for conducting parametric study has a number of advantages over the traditional on-site method. Firstly, it did not require that only incremental increases in the parameters (e.g., furnace-side feed mass flow rate) be implemented, allowing the study to be completed efficiently. In addition, the study can be performed without risk to the equipment, and thus reasonably large variations in the parameters as in the above parametric study can be evaluated without concern.

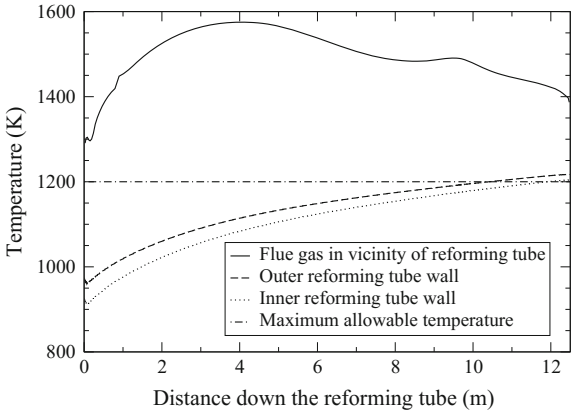

Fig. 4.22 Temperature profiles of the flue gas in the vicinity of the reforming tube (*solid line*), of the outer reforming tube wall (*dashed line*) and of the inner reforming tube wall (*dotted line*) from the pilot-scale reformer CFD simulation, where the furnace-side feed mass flow rate is increased by 20 % from its nominal value and all other parameters of the tube-side feed, furnace-side feed and combustion chamber refractory walls are consistent with typical plant data Latham (2008)

4.10 Conclusion

The present work initially described the detailed procedure for CFD modeling of a pilot-scale reformer, and subsequently demonstrated that the pilot-scale reformer CFD model could be employed for reformer parametric study. Specifically, the $k - \epsilon$ turbulence model, FR/ED turbulence-chemistry interaction model and global kinetic models of combustion Bane et al. (2010), Nicol (1995) were selected for use in the furnace side of the CFD model by studying the physical and chemical phenomena inside the combustion chamber, and were demonstrated to successfully capture the reaction rates of the furnace-side species and the thermal energy released by combustion in the turbulent reacting flow. The applicability of the FR/ED turbulence-chemistry interaction model was evaluated for the tube side of the CFD model due to the form of the global kinetic model of the SMR process Xu and Froment (1989). Simulations with a previously developed and validated industrial-scale reforming tube CFD model suggested that the errors due to implementing FR/ED as the turbulence-chemistry interaction model were negligible. The simulation results generated by the pilot-scale reformer CFD model were in agreement with available plant data from literature, and also with the simulation results generated by the industrial-scale reforming tube CFD model. Lastly, we employed the pilot-scale reformer CFD model to conduct a sample parametric study in which the mass flow rate of the furnace-side feed was increased by 20 % from its original value. We demonstrated that conducting the parametric study with the CFD model has a number of advantages compared to conducting it with the traditional on-site method, in particular

that it allows for efficient testing of even reasonably large changes in the process parameters, without risk to the equipment.

Acknowledgments Financial support from the Department of Energy is gratefully acknowledged.

References

Amirshaghaghi, H., Zamaniyan, A., Ebrahimi, H., & Zarkesh, M. (2010). Numerical simulation of methane partial oxidation in the burner and combustion chamber of autothermal reformer. *Applied Mathematical Modelling*, *34*, 2312–2322.

Bane, S. P. M., Ziegler, J. L., & Shepherd, J. E. (2010). Development of One-Step Chemistry Models for Flame and Ignition Simulations. Technical Report. GALCIT Report GALCITFM:2010.002.

Dybkjaer, I. (1995). Tubular reforming and autothermal reforming of natural gas—an overview of available processes. *Fuel Processing Technology*, *42*, 85–107.

de Lasa, H. I., Dogŭ, G., & Ravella, A. (Eds.) (1992). *Chemical reactor technology for environmentally safe reactors and products*. volume 225 of NATO ASI Series. Springer Science and Business Media, Dordrecht, The Netherlands.

Ergun, S., & Orning, A. A. (1949). Fluid flow through randomly packed columns and fluidized beds. *Industrial and Engineering Chemistry*, *41*, 1179–1184.

Froment, G. F., & Bischoff, K. B. (1990). *Chemical reactor analysis and design*. New York: Wiley.

ANSYS Inc. (2013). ANSYS Fluent Theory Guide 15.0.

Jones, W. P., & Launder, B. E. (1972). The prediction of laminarization with a two-equation model of turbulence. *International Journal of Heat and Mass Transfer*, *15*, 301–314.

Kroschwitz, J. I., & Howe-Grant, M. (Eds.). (1999). *Kirk-othmer encyclopedia of chemical technology*. New York, NY: Wiley.

Lao, L., Aguirre, A., Tran, A., Wu, Z., Durand, H., & Christofides, P. D. (2016). CFD modeling and control of a steam methane reforming reactor. *Chemical Engineering Science*, *148*, 78–92.

Latham, D. (2008). Masters thesis: Mathematical modeling of an industrial steam methane reformer. Queen's University.

Launder, B. E., & Sharma, B. I. (1974). Application of the energy dissipation model of turbulence to the calculation of flow near a spinning disc. *Letters in Heat and Mass Transfer*, *1*, 131–137.

Magnussen, B. F. (2005). The eddy dissipation concept: A bridge between science and technology. In *ECCOMAS Thematic Conference on Computational Combustion*, Lisbon, Portugal.

McGreavy, C., & Newmann, M. W. (1969). Development of a mathematical model of a steam methane reformer. In *Institution of Electrical Engineering, Conference on the Industrial Applications of Dynamic Modelling*, Durham, NC.

Nicol, D. G. (1995). Ph.D. Thesis: A Chemical Kinetic and Numerical Study of NOx and Pollutant Formation in Low-emission Combustion. University of Washington.

Pantoleontos, G., Kikkinides, E. S., & Georgiadis, M. C. (2012). A heterogeneous dynamic model for the simulation and optimisation of the steam methane reforming reactor. *International Journal of Hydrogen Energy*, *37*, 16346–16358.

Rostrup-Nielsen, J. R. (1984). *Catalysis: Science and technology* (pp. 1–117). Berlin, Germany: Springer. chapter Catalytic Steam Reforming.

Sadooghi, P., & Rauch, R. (2013). Pseudo heterogeneous modeling of catalytic methane steam reforming process in a fixed bed reactor. *Journal of Natural Gas Science and Engineering*, *11*, 46–51.

Turns, S. R. (1996). *An introduction to combustion: Concepts and applications*. Boston, MA: McGraw-Hill.

Udengaard, N. R. (2004). Hydrogen production by steam reforming of hydrocarbons. *Preprint Papers-American Chemical Society, Division of Fuel Chemistry*, *49*, 906–907.

Wesenberg, M. H., & Svendsen, H. F. (2007). Mass and heat transfer limitations in a heterogeneous model of a gas-heated steam reformer. *Industrial and Engineering Chemistry Research*, *46*, 667–676.

Xu, J., & Froment, G. F. (1989). Methane steam reforming, methanation and water-gas shift: I. intrinsic kinetics. *AIChE Journal*, *35*, 88–96.

Zamaniyan, A., Behroozsarand, A., & Ebrahimi, H. (2010). Modeling and simulation of large scale hydrogen production. *Journal of Natural Gas Science and Engineering*, *2*, 293–301.

Chapter 5
Data-Based Model Reduction for Refinery-Wide Optimization

Taoufiq Gueddar and Vivek Dua

5.1 Introduction

Production optimization and planning is a key part of the modern process industry operations. It plays a major role in sustaining the profitability of the operations and providing decision making guidelines to attain business objectives through achieving best margins and minimum operating costs. Numerous industries have benefited from planning models to achieve optimal operations subject to technologies involved, resources requirements, production constraints and available markets for the products sales.

The oil refining is one of the industries, which represents the core of the global oil supply chain process (Gary and Handwerk 1994). It processes crude oils to produce valuable refined petroleum products. Shipment of these products to the local and international market places is subject to meeting commercial specifications which depend on the season and the geographical region where the refinery is operating. The fuel products demands are driven by economics and industrial sectors growth as well as other factors. Another constraint is the introduction of environmental restrictions that are leading to tighter fuel specifications that require investing in new refinery technologies and/or a better crude feedstock to improve the quality of the products downstream. Challenges from the operational point of view include the selection of feed stocks and operating conditions. These should be considered for each refinery unit in order to achieve higher margins whilst meeting the final product specifications.

The issues mentioned above require refinery-wide optimization which is computationally complex. Crude oil distillation unit is one of the highest energy consumers in refineries due to their complex configuration (Gadalla et al. 2003). Such

T. Gueddar · V. Dua (✉)
Department of Chemical Engineering, Centre for Process Systems Engineering, University College London, London, UK
e-mail: v.dua@ucl.ac.uk

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_5

complex systems with high energy-yield integration requires intensive modeling and is computationally challenging (Caballero and Grossmann 2004; Zhou 2015).

Similarly, refinery reactor modeling involves rigorous kinetic modeling of complex catalytic and thermal reactions; fluid catalytic cracking modeling aspects are discussed in (Van Landeghem et al. 1996; Theologos and Markatos 1993), catalytic reforming modeling aspects (Ancheyta-Juarez and Villafuerte-Macias 2000) and thermal cracking modeling aspects in (Sundaram and Froment 1977; Del Bianco et al. 1993). Catalytic Hydrocracking modeling aspects are discussed in (Ancheyta et al. 2005; Umana 2014). Such rigorous models are large scale and computationally intensive as confirmed by Powell and Yu (2004).

Furthermore, process simulators represent oil refineries more realistically, however are not suitable for optimization due to the size and the nonlinearity of the problem. Therefore model reduction techniques are required to address these issues.

5.1.1 Refinery Operation Overview

The crude oil after being shipped to the refinery is firstly processed through a desalter preheated with the downstream Crude and Vacuum units products before being heated in a fired heater. It is then fed to the atmospheric crude distillation unit (CDU) to be fractioned into: LPG (liquefied petroleum products), naphtha, kerosene, atmospheric gas oil, and atmospheric residue (Fig. 5.1). The residue is sent to the vacuum distillation tower and separated into vacuum gas oil and a vacuum residue. The vacuum residue can be used as a feedstock to a delayed coker unit to produce fuel gas, coker naphtha, and coker gas oil and coke (Gary and Handwerk 1994).

The visbreaker is another thermal cracking process that can use the Vacuum Residue (VR) as a feedstock. In the refineries where no deep conversion units are available, the VR is sold as a heavy fuel oil or as bitumen. The atmospheric gas oil and light vacuum gas oils as well as the coker gas oil are used as feedstocks for the fluid catalytic cracking (FCC) or hydrocracking/hydroprocessing. These units crack the heavy molecules into lower molecular weight compounds boiling in the naphtha and distillates cut ranges.

The products from the hydrocracker are saturated. The unsaturated catalytic cracker products are saturated and improved through processing into hydrotreating and/or reforming units. The straight run naphtha cut from the crude distillation unit is sent to a Naphtha splitter that separate the naphtha full range into a light and heavy naphtha streams. The light naphtha cut from the naphtha splitter, coker and cracking units are used as a feedstock to the isomerization unit to upgrade the research octane[1] numbers and motor octane numbers by increasing the branching of

[1]Octane number: The octane number is a quantitative measure of the maximum compression ratio at which a particular gasoline/naphtha product can be utilized in internal combustion engines without some of the fuel-air mixture undergoing self-ignition.

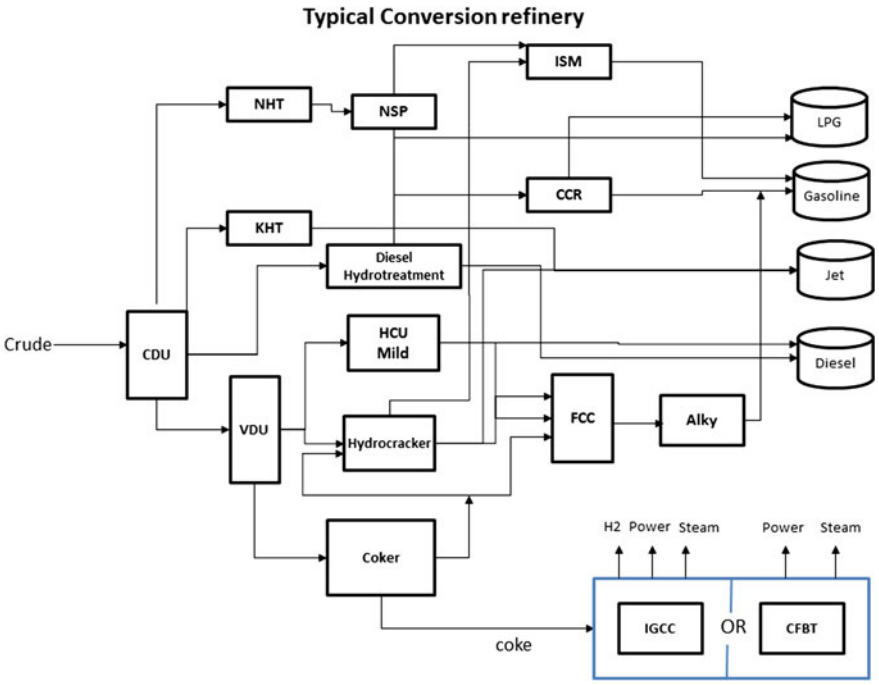

Fig. 5.1 Refinery flowsheet

the paraffinic chains. The heavy naphtha cuts from the crude distillation unit, coker, and cracking units are fed to the reformer in order to improve their octane numbers. The reformate product from the catalytic reformer is blended along with other products, into regular and premium gasoline for sale. The gas produced in the crude unit, coker, and cracking units is separated in the gas plant into fuel gas, liquefied petroleum gas (LPG), unsaturated hydrocarbons (propylene, butylenes, and pentenes), normal butane, and isobutane. The fuel gas is used in the fired heaters. The normal butane can be blended into gasoline or LPG.

The unsaturated hydrocarbons and isobutane are sent to the alkylation unit for processing. The middle distillates from the crude unit, coker, and cracking units are blended into diesel and jet fuels.

5.1.2 Refinery Optimization

The oil refinery remains a key component of the global oil production and distribution process. It consists of an industrial process plant where crude oils are processed and refined into more valuable petroleum products (Gary and Handwerk 1994).

These products should meet certain commercial specifications before being shipped to the domestic and international markets. The oil refining industry is currently facing new challenges. In fact, there is a change in fuel demand towards lighter products, as well as a change in the quality of crude oils. Furthermore, environmental restrictions are leading to tighter fuel specifications that require refiners to either invest in new refinery technologies or use better crude feedstock to improve the quality of the products downstream. Challenges from the operational point of view include the selection of crude oils to be fed to the refinery in order to achieve higher margins and to meet the final product specifications, impacting the cost per barrel.

The crude, refined products prices fluctuations as well as utilities costs make refinery-wide optimization a key practice in each operating refinery over the world (Symonds 1955). The refinery-wide optimization models reflect the refinery strategic plans, the nature of the configuration, markets they serve and time periods they cover. Therefore, refinery models differ in complexity and sizes.

The complexity of oil refineries operations makes the operations decision-making process an extremely difficult task, especially when uncertainty in feed specifications, product demand and economic parameters is involved (Dua et al. 2004; Yang and Barton 2015) Refinery wide optimization is a key requirement to overcome these challenges to achieve a high standard of performance and to stay competitive within the market (Slaback and Riggs 2007). Through refinery-wide optimization, optimal refinery crude feed mixture as well as gasoline and distillates blending can be achieved (Mendez et al. 2006). It can also be used to minimize low-value fractions and quality giveaway whilst ensuring a better energy management, which is becoming of increasing importance in process industries (Georgiadis et al. 2008).

5.1.3 Current State of the Art

Historically, in the oil refining industry, all refinery optimization tasks were initially done by performing utility and material balances calculation manually until an optimal solution is reached. The manual simplex procedure was tedious and time consuming (Parkash 2003). Then after the implementation of the simplex algorithm in computer programs, oil refineries gained interest in optimizing their operations by solving linear programming models (Favennec 2001). The typical refinery LP model by then had approximately 300–500 equations and 800–1500 variables (Parkash 2003).

The LP models have been and still are the standard method for production planning and the underlying method for most commercial packages such as PIMS (Aspen Tech), GRTMPS (Haverley Systems) and RPMS (Honeywell Hi-Spec Solutions).

LP modeling in the oil refining industry is still the main method for the refinery-wide optimization since it was introduced. LP models are routinely used to

optimize refinery operations and work processes such as crude acquisition, final product blending and product sales as well as assess new investment opportunities. For such important financial decisions it is crucial to have models that are as accurate and robust as possible. However it is well known that LP models are attractive from the computational point of view but their main drawback is the lack of consistency and accuracy.

Although the linearity assumption simplifies the complexity of the refinery behaviors leading to short computational times and robust models, it fails to obtain accurate solution to the refinery-wide optimization problem. Attempts to solve the NLP refinery-wide optimization problem were hindered by convergence issues and excessive computational times in an environment that requires business decisions to be made within short times. Therefore, there is currently a demand for achieving accurate refinery-wide optimization solutions within reasonable timelines without having to be penalized on the accuracy side.

5.1.4 Data-Based Model Reduction

The mathematical models for chemical operation units are usually highly nonlinear particularly the ones in the refining industry (e.g. crude separation, fixed bed based reactors, fluidized bed reactors, etc.). Therefore, these high fidelity models have been difficult to use for optimization purposes although they are successfully used in simulation and are well recognized in reproducing the nonlinear behavior of these units. Within the oil refining industry, there are in fact different rigorous simulation packages available. These softwares are able to simulate the refinery units with accuracy. However, these packages are mainly used for process design, unit monitoring, and energy management. Attempts to achieve refinery-wide optimization with rigorous simulators have not been very successful. Therefore, many research efforts have tried to propose models that approximate the original models accurately but have smaller dimensional size, as it is easier to implement into an optimization problem. Indeed, many researchers responded to the increasingly need for accurate and simpler models by developing efficient strategies in order to replace rigorous physical property calculations by simplified ones (Slaback and Riggs 2007; Alattas et al. 2011; Guerra and Le Roux 2011; Yang and Barton 2015).

Systems data sets hold valuable information. Therefore, models based on actual input/output data allow us to capture and understand the modeled system's behavior from these data sets. They can also be considered as substitutes for more computationally intensive models in fields where computational time is critical or where the relationships between variables are difficult to model with usual mathematical formulations. In this present work, a data-based Artificial Neural Network (ANN) model reduction technique is used to model the nonlinearity of the Refinery models, which will then be used in refinery-wide optimization.

5.2 MIPANN Based Model Approximation for Refinery Optimization

The oil refining industry mainly uses Linear Programming (LP) for refinery optimization and planning purposes, on a daily basis. LP is attractive from the computational time point of view; however it has limitations such as the nonlinearity of the refinery processes is not taken into account. The main aim of this work is to develop approximate models to replace the rigorous ones providing a good accuracy without compromising the computational time, for refinery optimization. The data for deriving approximate models has been generated from rigorous process models in Petrosim®, a commercial software, which is extensively used in the refining industry. In this work we present novel model reduction techniques based upon optimal configuration of artificial neural networks to derive approximate models and demonstrate how these models can be used for refinery-wide energy optimization.

5.2.1 Artificial Neural Networks

Artificial neural networks (ANN) are very efficient in approximating complex nonlinear systems (Hornik et al. 1989; Venkatasubramanian and Chan 1989) and have been widely used for solving many engineering problems (Himmelblau 2008; Hussain and Kershenbaum 2000). The ANN consists of a network structure of neurons, where neurons are nonlinear, unbounded, parameterized functions. Variables are given in the form of inputs to the neurons, which use its internal activation function to calculate outputs (Fig. 5.2).

Figure 5.2 shows typical artificial neural network architecture. It is composed of interconnected nodes, which can be categorized into Input layer, Hidden layers and an Output layer. The input layer consists essentially of a group of nodes, which receives input signal from the external system while the output layer consists of groups of nodes that returns the signal to the external environment. The hidden layers receives signal from input layers and gives out signal to the output layers and does all computational analysis.

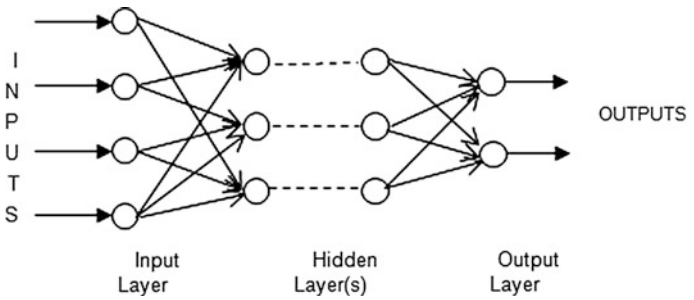

Fig. 5.2 Artificial neural network

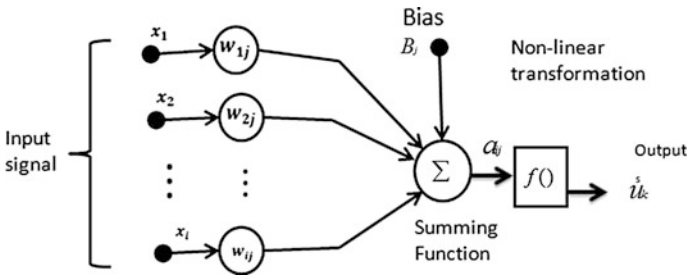

Fig. 5.3 Single processing node

Figure 5.3 illustrates a single processing node. The node receives an input signal x_i either from an external source or from another node. The input signal is usually adjusted according to the weight factor w_{ij} which value determines the output from any given node in the network. The biases B_j also called internal threshold determines the internal activation of the node it is associated with.

Every node calculates the total activation and it is equal to the sum of all $w_{ji}^1 x_i$ (weight factor multiply by input data) plus the internal threshold of that node. It can be expressed mathematically as:

$$Total\ activation = a_j^1 = \sum_{i=1}^{N_n} w_{ji}^1 x_i + b_j^1 \tag{5.1}$$

The calculated total activation is then sent to the transfer function $f()$. The transfer function can be any mathematical function but in this work the hyperbolic tangent mathematical function shown graphically in Fig. 5.4 will be used.

$$h_j^l = \tanh(a^l) \quad l = 1, \ldots, N_n \tag{5.2a}$$

The transfer function for the node calculation is given by:

$$h_j^l = \tanh\left(\sum_{i=1}^{N_n} w_{ji}^1 x_i + b_j^1\right) \quad l = 1, \ldots, N_n \tag{5.2b}$$

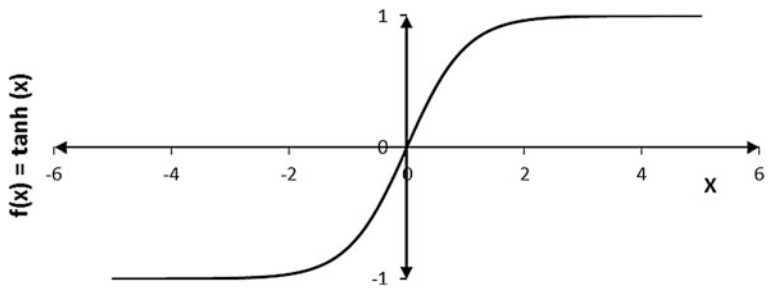

Fig. 5.4 Hyperbolic tangent function

The next stage is learning and training the network. The most commonly used learning method is called the error correction learning. Here, the neural network weights are adjusted in proportion to output error vector $\boldsymbol{\varepsilon}$. The error associated with a single processing node is expressed mathematically as:

$$\varepsilon_k = (\hat{u}_k - u_k)^2 \tag{5.3}$$

where $\hat{u}_k$ is the desired output and u_k corresponds to the calculated output associated with the kth node in the output layer.

The artificial neural network is trained by solving the following non-linear programming problem.

5.2.1.1 Nonlinear Formulation

Each input has an assigned weight. The output of the neuron will be calculated by using a nonlinear combination of the inputs, and the weights. The computation of the synaptic weights can be achieved through the process of learning. The objective of the learning is to calibrate the network with a set of data for which we know the input and the corresponding output (Dua 2010). The learning process is achieved by solving the following Nonlinear Programming (NLP) problem:

ANN prediction error objective function:

$$\min_{a,b,w,h,W,B,u} E_1 = \sum_{k=1}^{N_o} (\hat{u}_k - u_k)^2 \tag{5.4}$$

Subject to:

Activation variables of the first hidden layer:

$$a_j^1 = \sum_{i=1}^{N_x} w_{ji}^1 x_i + b_j^1, \quad j = 1, \ldots, N_n \tag{5.5}$$

Nonlinear transformation of activation variables:

$$h_j^l = \tanh\left(a_j^l\right) \quad l = 1, \ldots N_h \tag{5.6}$$

Activation variables of remaining hidden layers:

$$a_j^l = \sum_{i=1}^{N_n} w_{ji}^l h_j^{l-1} + b_j^l, \quad j = 1, \ldots, N_n, \quad l = 2, \ldots, N_h \tag{5.7}$$

ANN Output:

$$u_k = \sum_{i=1}^{N_n} W_{ki} h_j^{N_h} + B_k, \quad k = 1, \ldots, N_o \tag{5.8}$$

where x_i denotes the input values to the network, $i = 1, \ldots, N_x$ is the number of inputs, N_n linear combinations of these inputs gives the activation variables, a_j^i, where N_n is the number of nodes in the hidden layer, the superscript 1 denotes the index of the first hidden layer, w_{ij} are the weights and b_j the biases. These activation variables are then transformed non-linearly to provide h_j^l, the output of a hidden layer; note that nonlinear transformations other than tanh are also used in the literature. h_j^l, the output of a hidden layer then becomes the input to the next hidden layer. The outputs from the last hidden layer h_j^{Nn} are then combined to provide the outputs, u_k. N_0 is the number of nodes in the output layer and w_{ki} and B_k are the weights and biases respectively. Let $\hat{u}_k$ denote the desired output, the training of the network can then be formulated as minimization of the error function, E_1. For simplicity in presentation, u_k, $\hat{u}_k$, a_j^l, h_j^l and x_i represent vectors for all the points in the training data set.

5.2.1.2 MINLP Formulation

Artificial neural network especially with feed forward neural network topology has been extensively used in modeling engineering applications mainly because of its ability to model highly non-linear system. However, the major setback of using feed forward neural network topology is the difficulty in determining optimum network configuration. Determining the optimum network configuration is important because a neural network structure that is too small (simple) may have difficulty learning complex behaviors and patterns in the training dataset. And a network that is too large (complicated) will lead to over-fitting of the training dataset hence may perform poorly over unseen data (Dua 2010).

There are several nodes and interconnections in large neural network that do not contribute to network performance. These nodes and interconnections are usually referred to as redundant nodes and interconnections.

The rationale behind Mixed Integer Programming for Artificial Neural Network (MIPANN) is to model the training of the neural network as a mixed-integer nonlinear programming problem. Binary variables [0, 1] are used to model the non-existence/existence of nodes and interconnections. The ANN training is then implemented by minimizing the number of nodes and interconnections consecutively while keeping the training error below a predefined tolerance. With the determination of optimum network configuration, a new ANN is trained using the simplified network topology by minimizing the sum-squared error between predicted ANN output and desired ANN output. Figure 5.5 shows the process for obtaining an optimal network configuration from an original configuration. For further details refer to (Gueddar and Dua 2012).

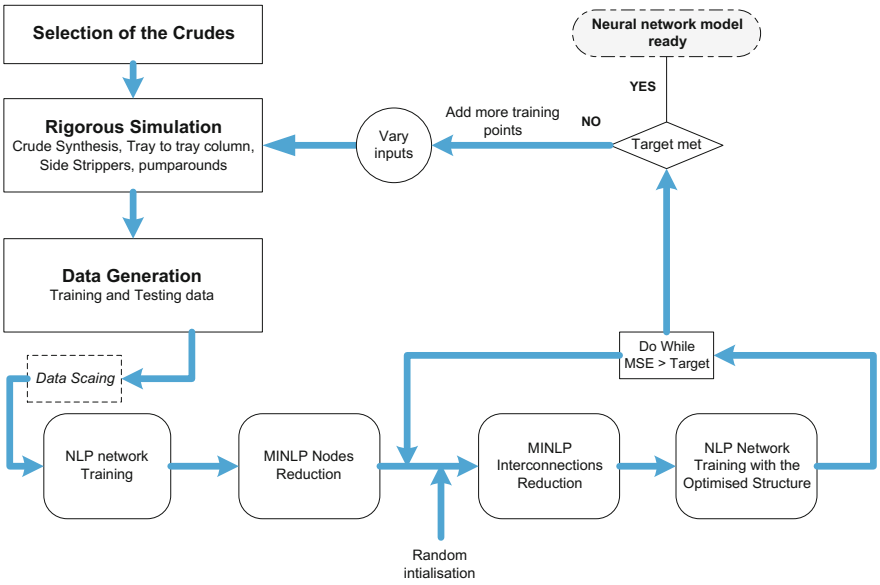

Fig. 5.5 MIPANN process

5.3 Disaggregation–Aggregation Based Model Approximation for Refinery Optimization

A Disaggregation–Aggregation method is proposed for the training of ANNs for medium and large datasets. The method proposes the splitting (disaggregation) of a dataset into a number of smaller subsets, but instead of training the ANN on each subset sequentially; the subsets are used for the training of different sub-ANNs individually. The sub-ANNs are then used to train a new ANN over the whole dataset (aggregation) where the Sum Squared Error (SSE) is given by the squared error of the predicted output of the ANN from the predicted error of the sub-ANN. The procedure of disaggregation and aggregation consists of splitting the given training data into n subsets, then training of sub-ANNs over the n subsets to obtain n sub-ANNs. This is then followed by training the full ANN model by minimizing the (SSE) between the full ANN outputs and the sub-ANNs outputs. Each sub-ANN is only considered for input data within the range of inputs for which it was trained.

Successful application of this algorithm could potentially allow for the parallel computation of (Fig. 5.6) ANNs and therefore decrease computational times for the training for ANNs. The Disaggregation–Aggregation method has been applied to obtain the reduced models for CCR and naphtha splitter units described in Sect. 5.4. This method is compared to the traditional method, which uses all the data points

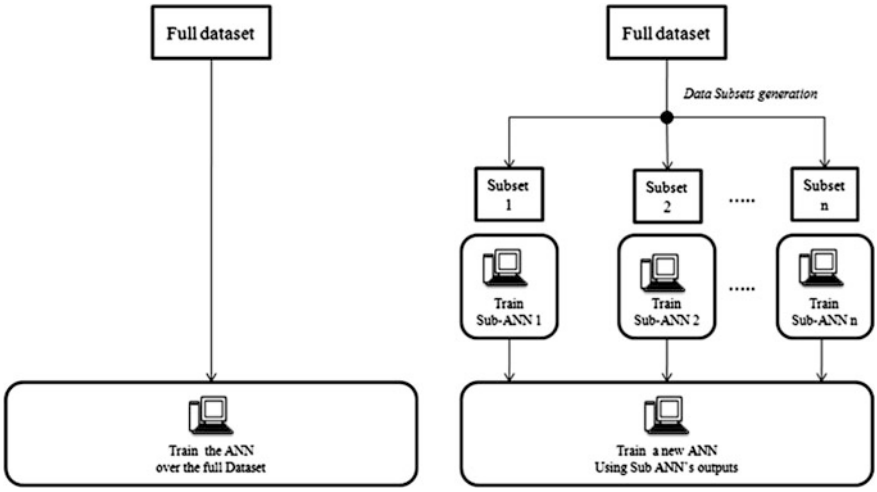

Fig. 5.6 Implementation of Disaggregation/Aggregation method in a parallel computing platform

simultaneously, in terms of the network prediction accuracy and computational time requirements. For further details refer to (Gueddar and Dua 2011) (Fig. 5.7).

5.4 Oil Refining System Under Study

After the shipment of crude oil to the refinery, it is first processed through a desalter preheated with the downstream Crude and Vacuum unit's products before being heated in a fired heater. It is then fed to the atmospheric crude distillation unit (CDU) to be fractioned into: LPG (liquefied petroleum products), Naphtha, kerosene, atmospheric gas oil, and atmospheric residue. The straight run naphtha cut from the crude distillation unit is sent to a naphtha splitter that separates the naphtha full range into light and heavy naphtha streams. The light naphtha cut from the naphtha splitter, coker and cracking units are used as a feedstock to the isomerization unit to upgrade the research octane numbers and motor octane numbers by increasing the branching of the paraffinic chains. The heavy naphtha cuts from the CDU, coker, and cracking units are fed to the reformer in order to improve their octane numbers. The reformate product from the catalytic reformer is blended along with other products, into regular and premium gasoline for sale. In this chapter the main focus is on the Naphtha processing and gasoline blending sections of the oil refinery.

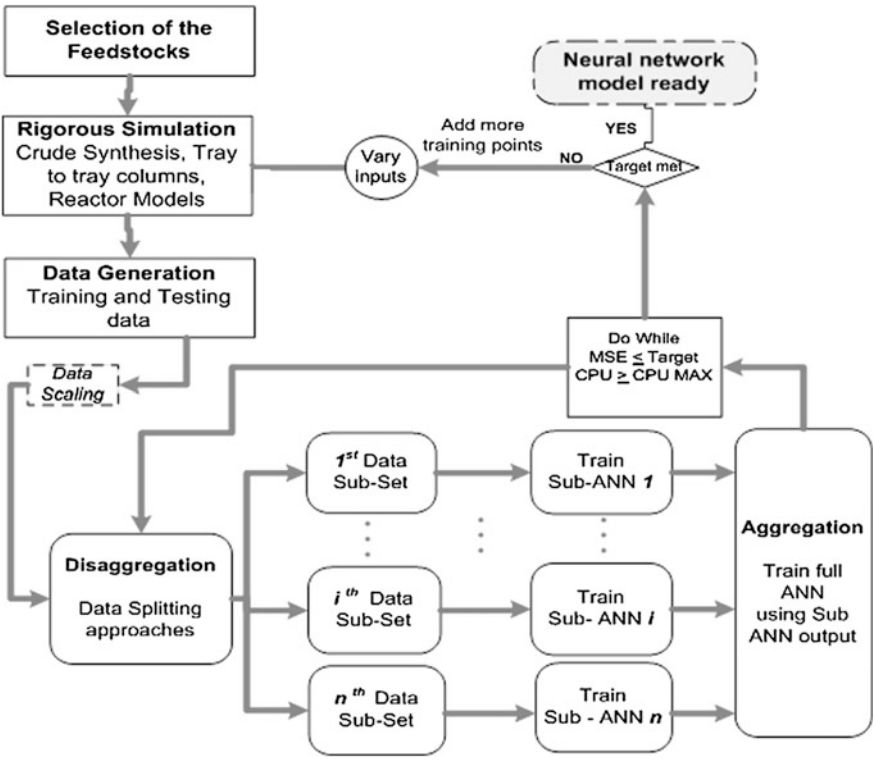

Fig. 5.7 Disaggregation Aggregation process

5.4.1 Refinery Units Modeling and Data Generation

In this chapter, reduced ANN models for the CCR and naphtha splitter are obtained by using a rigorous process simulator to generate input-output training and testing data. The obtained reduced models are used later for refinery and energy optimization application.

5.4.1.1 Refinery Units Background (Naphtha Splitter and CCR)

After the fractionation of the crude oil the overhead of the distillation tower is stabilized to recover gas and liquefied petroleum gas (LPG). The bottom of the stabilizer which is full range naphtha is then routed to the hydrotreatment section to

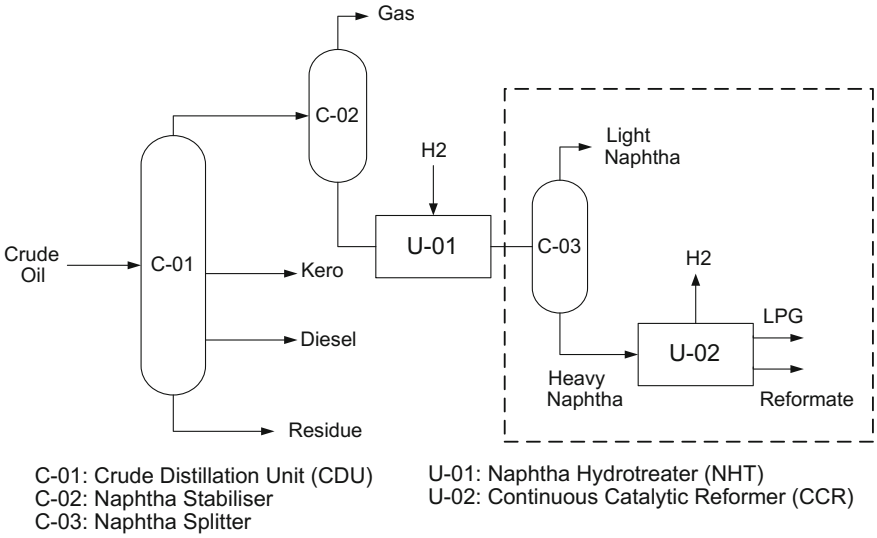

Fig. 5.8 Naphtha splitter and Continuous Catalytic Reforming flowsheet

remove most of the sulphur content. The hydrotreated naphtha is then separated in a Naphtha Splitter distillation tower into a light and heavy naphtha. The cut point of this separation will be set by production targets and units operating downstream. The flowhseet of this section of the refinery is shown in Fig. 5.8. In fact, if the refinery has an isomerization unit, then the presence of C7 components in the light naphtha feed may impact the performance of the unit as well as the physical properties of the isomerate. The Reid Vapor Pressure (RVP), specific gravities, paraffins, naphthenes and aromatics are other factors that are directly impacted by the cut point. In regards to the CCR operation, depending on the Light/heavy naphtha cut point, C6 precursors may or may not be present in the heavy naphtha feed. If it is the case, then benzene content of the reformate will increase which can be an issue in terms of meeting gasoline benzene specification (1 vol.%). Clearly there is a trade-off between increasing the yield of heavy naphtha against allowing C6 precursors presence in it. Another key operating variable is the severity of the CCR. The higher is the severity the higher is the Research Octane Number (RON), the Motor Octane Number (Symonds 1955), the aromatics and the benzene content of the reformate. To meet operations targets and maximize economical margins the use of optimization models is a must. Optimization cases from real refinery scenarios are presented in Sect. 5.5.

5.5 Combined Disaggregation–Aggregation and MIPANN Approach for Model Approximation

This chapter proposes a combination of MIPANN and disaggregation–aggregation reduction methods proposed earlier in Sects. 5.3 and 5.4 to handle medium and large size data sets. It is anticipated that it will result in faster training computational times for complex systems (e.g. Oil refinery units) and result in simplified models.

5.5.1 Combined Disaggregation–Aggregation MIPANN Formulation

As outlined in the previous chapters, the training of artificial neural network (ANN) is computationally expensive especially for large data sets. In this chapter a disaggregation–aggregation mixed integer programming method is proposed for training artificial neural network model for medium to large dataset. Overview of the proposed approach is shown in Fig. 5.9. It involves the division (disaggregation) of the initial large dataset, which otherwise will be computationally expensive, into smaller sub-datasets. A sub-ANN is trained individually using the sub-data by minimizing the sum-squared error (SSE) between the predicted ANN output from the individual sub-data and the known output data of the given sub-data. A new aggregated ANN is then trained using the combined (aggregation) dataset by minimizing the sum-squared error between predicted aggregated ANN output and sub-ANN outputs. In the next phase, redundant nodes and interconnections are eliminated from the aggregated ANN model by solving two mixed integer nonlinear programming models; thereby producing an ANN with reduced network structure. The existence and nonexistence of nodes and interconnections are modelled through binary variables [0, 1] in the aggregated mixed integer programming model, the sum of the binary variables associated with the nodes and interconnections are minimized consecutively. The rest of the chapter is organized as follows. In the next section note that the mathematical formulations from the previous chapters are repeated here for the sake of completeness and continuity.

5.5.2 Mathematical Formulation of the Combined Approach

5.5.2.1 Mathematical Formulation of Disaggregation

In this formulation, the large dataset is divided into smaller sub-data $s = s_1, s_2, s_N$ and for each sub-data the following optimization problem is solved:

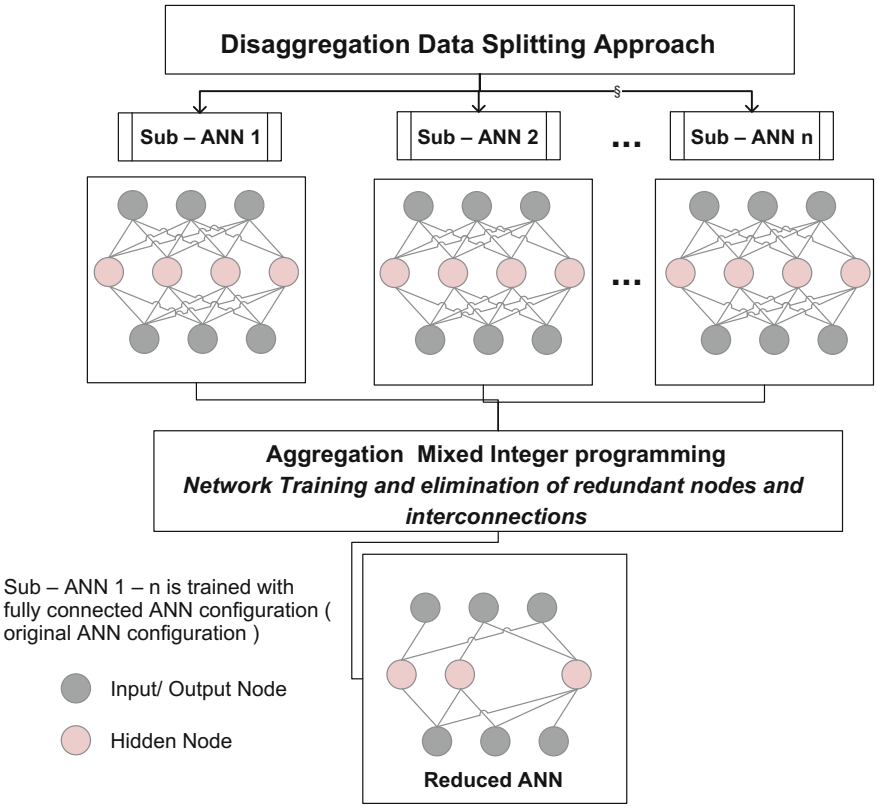

Fig. 5.9 Disaggregation–aggregation mixed integer-programming implementation

Objective function:

$$\min_{a,b,w,h,B,W,u,y,\delta} E_s = \sum_{k=1}^{No} \left(\hat{u}_{k,s} - u_{k,s}\right)^2 \tag{5.9}$$

Subject to:

$$a_{j,s}^1 = \sum_{i=1}^{N_x} w_{ji,s}^1 x_{i,s} + b_{j,s}^1 \quad j = 1, \ldots, N_n, \quad s = 1, \ldots, N_s \tag{5.10}$$

$$h_{j,s}^l = \tanh\left(a_{j,s}^l\right) \quad i = 1, \ldots, N_h \tag{5.11}$$

$$a_{j,s}^{l} = \sum_{i=1}^{N_x} w_{ji,s}^{l-1} h_{j,s}^{l-1} + b_{j,s}^{l} \quad j = 1, \ldots, N_n, \quad l = 2, \ldots, N_h \tag{5.12}$$

$$u_{k,s} = \sum_{i=1}^{N_n} W_{ki,s} h_{j,s}^{N_h} + B_{k,s} \quad k = 1, \ldots., N_o, \quad m = 1, \ldots., N_{ds} \tag{5.13}$$

5.5.2.2 Mathematical Formulation of Aggregation

Here, the sub-ANN outputs is used to train a combined (aggregation) ANN model by minimizing the sum squared error between the predicted aggregated ANN output and sub-ANN output. The following optimization problem is solved:

Objective function:

$$\begin{aligned} \min_{a,b,w,h,B,W,u,y,\delta} E = & \sum_{k=1}^{Nds1} \left(\tilde{u}_{k,s1} - u_k\right)^2 \\ & + \sum_{k=Nds1}^{Nds2} \left(\tilde{u}_{k,s2} - u_k\right)^2 + \cdots + \sum_{k=Nds,n-1}^{Ndsn} \left(\tilde{u}_{k,sn} - u_k\right)^2 \end{aligned} \tag{5.14}$$

Subject to:

$$a_j^1 = \sum_{i=1}^{N_n} w_{ji}^1 x_i + b_j^1 \quad j = 1, \ldots, N_n \tag{5.15}$$

$$h_j^l = \tanh\left(a_j^l\right) l = 1, \ldots, N_n \tag{5.16}$$

$$a_j^l = \sum_{i=1}^{N_n} w_{ji} h_j^{i-1} + b_j^1 \quad j = 1, \ldots, N_n, \quad l = 2, \ldots, N_n \tag{5.17}$$

$$u_k = \sum_{i=1}^{N_n} W_{ki} h_j^{N_h} + B_k \quad k = 1, \ldots, N_o \tag{5.18}$$

5.5.2.3 Mathematical Formulation of Aggregation Mixed Integer Programming—Elimination of Nodes

In this formulation, the sub-ANN predicted output is used to train a new combined (aggregated) ANN by eliminating redundant nodes. The existence and nonexistence of node is modeled as a binary variable y such that, if $y = 0$ then the associated node is redundant and otherwise if $y = 1$.

Consider the formulation of the proposed approach:
Objective function:

$$\min_{a,b,w,h,B,W,u,y} E_1 = \sum_{l=1}^{Nh} \sum_{j=1}^{Nn} y_j^l \tag{5.19}$$

Subject to:

$$E = \sum_{k=1}^{Nds1} (\tilde{u}_{k,s1} - u_k)^2 + \sum_{k=Nds1}^{Nds2} (\tilde{u}_{k,s1} - u_k)^2 + \cdots + \sum_{k=Nds,n-1}^{Ndsn} (\tilde{u}_{k,s1} - u_k)^2 \tag{5.20}$$

$$a_j^1 = \sum_{i=1}^{N_n} w_{ji}^1 x_i + b_j^1 \quad j = 1, \ldots, N_n \tag{5.21}$$

$$h_j^l = \tanh\left(a_j^l\right) l = 1, \ldots, N_n \tag{5.22}$$

$$a_j^l = \sum_{i=1}^{N_n} w_{ji} h_j^{i-1} + b_j^1 \quad j = 1, \ldots, N_n, \quad l = 2, \ldots, N_n \tag{5.23}$$

$$u_k = \sum_{i=1}^{N_n} W_{ki} h_j^{N_h} + B_k \quad k = 1, \ldots, N_o \tag{5.24}$$

$$-M \cdot y_j^l \le a_j^l \le M \cdot y_j^l \quad j = 1, \ldots, N_n, \quad l = 1, \ldots, N_h \tag{5.25}$$

$$E \le \varepsilon \tag{5.26}$$

In this formulation, the training error E is kept below a predefined tolerance ε. In Eq. (5.19) E_1 is the objective function. Equations (5.14)–(5.18) are identical to Eqs. (5.20)–(5.24) and the variables have the same meaning. Equation (5.25) is a nonlinear constraint equation that models the existence and nonexistence of nodes. In Eq. (5.25) M is a positive large number and y_j^l is a binary variable associated with the jth node in the lth hidden layer.

5.5.2.4 Mathematical Formulation of Aggregation Mixed Integer Programming-Elimination of Interconnections

Similarly to the elimination of nodes considered in the previous section, the sub-ANN predicted output is also used to train a new aggregated ANN by eliminating redundant interconnections. The existence and nonexistence of interconnections is modeled as a binary variable δ such that, if $\delta = 0$ then the associated interconnections is redundant and otherwise if $\delta = 1$. Furthermore, there are no interconnections between redundant nodes eliminated in the previous step.

Consider the formulation of the proposed approach (note that the formulation below considers the simplified network from which redundant nodes have been removed):

Objective function:

$$E_2 = \sum_{l=1}^{Nh} \sum_{j=1}^{Nn} \sum_{i=1}^{N} \delta_{ji}^{l} \tag{5.27}$$

Subject to:

$$E = \sum_{k=1}^{Nds1} \left(\tilde{u}_{k,s1} - u_k\right)^2 + \sum_{k=Nds1}^{Nds2} \left(\tilde{u}_{k,s1} - u_k\right)^2 + \cdots + \sum_{k=Nds,n-1}^{Ndsn} \left(\tilde{u}_{k,s1} - u_k\right)^2 \tag{5.28}$$

$$a_j^1 \check{y}_j^1 = \sum_{i=1}^{N_n} w_{ji}^1 x_i + b_j^1 \check{y}_j^1 \quad j = 1, \ldots, N_n \tag{5.29}$$

$$h_j^l \check{y}_j^l = \tanh\left(a_j^l \check{y}_j^l\right) l = 1, \ldots, N_h \tag{5.30}$$

$$a_j^l \check{y}_j^l = \sum_{i=1}^{N_n} w_{ji} h_j^{l-1} + b_j^l \check{y}_j^l j = 1, \ldots, N_n, \quad l = 2, \ldots, N_h \tag{5.31}$$

$$u_k = \sum_{i=1}^{N_n} W_{ki} h_j^{N_h} + B_k \quad k = 1, \ldots, N_o \tag{5.32}$$

$$E \leq \varepsilon \tag{5.33}$$

$$-M \cdot \delta_{ji}^1 \leq w_{ji}^1 \leq M \cdot \delta_{ji}^1 \quad j = 1, \ldots, N_x \tag{5.34}$$

$$-M \cdot \delta_{ji}^l \leq w_{ji}^l \leq M \cdot \delta_{ji}^l \quad j = 1, \ldots, N_n, \quad l = 1, \ldots, N_h \tag{5.35}$$

In this formulation, the training error E is kept below a predefined tolerance ε. In Eq. (5.27) E_2 is the objective function. Equations (5.29)–(5.31) is identical to Eqs. (5.21)–(5.23) except with the introduction of the binary variables y_j^l used as a parameter $\check{y}_j^l$. The parameter $\check{y}_j^l$ is put in place to limit the computation of model in reduced structure since there is no interconnections between redundant nodes eliminated from the previous step (i.e. $\check{y}$ is solution of the problem (5.19)–(5.26)). Equations (5.34) and (5.35) are nonlinear constraint equations that model the existence/nonexistence of interconnections in the first hidden layer and l hidden layer. In Eqs. (5.34) and (5.35), M is a positive large number and δ_{ji}^l is a binary variable associated with the interconnection between jth node and ith node in the lth hidden layer.

5.5.2.5 Mathematical Formulation of Aggregation—Reduced ANN

After the successful elimination of the redundant of nodes and interconnections, a new simplified aggregated ANN is trained using the reduced network structure. Consider the formulation of the proposed approach (note that the formulation below considers the simplified network from which redundant nodes and interconnections have been removed):

Objective function:

$$E = \sum_{k=1}^{Nds1} (\tilde{u}_{k,s1} - u_k)^2 + \sum_{k=Nds1}^{Nds2} (\tilde{u}_{k,s1} - u_k)^2 + \cdots + \sum_{k=Nds,n-1}^{Ndsn} (\tilde{u}_{k,s1} - u_k)^2 \quad (5.36)$$

Subject to:

$$a_j^1 \check{y}_j^1 = \sum_{i=1}^{N_n} w_{ji}^1 \delta_{ji}^1 x_i + b_j^1 \check{y}_j^1 j = 1, \ldots, N_n \quad (5.37)$$

$$h_j^l \check{y}_j^l = \tanh\left(a_j^l \check{y}_j^l\right) \quad l = 1, \ldots, N_n \quad (5.38)$$

$$a_j^l \check{y}_j^l = \sum_{i=1}^{N_n} w_{ji} \delta_{ji}^{l-1} h_j^{l-1} + b_j^l \check{y}_j^l \quad j = 1, \ldots, N_n, \quad l = 2, \ldots, N_n \quad (5.39)$$

$$u_k = \sum_{i=1}^{N_n} W_{ki} \delta_{ji}^{N_h} h_j^{N_h} + B_k \quad k = 1, \ldots, N_o \quad (5.40)$$

Similar to the previous formulation, the binary variables y_j^l used as a parameter $\check{y}_j^l$ and binary variables δ_{ji}^l used as parameter $\check{\delta}_{ji}^l$ are introduced in this formulation to run the aggregation while using the finalized structure after removing the redundant nodes and interconnections, i.e. δ is solution of the problem (5.27)–(5.35).

5.5.3 *Data Generation Process*

The training and testing data sets used in this chapter are the same ones generated for the Sect. 5.4 to model the CCR and naphtha splitter units. The process is also shown in Fig. 5.10 and described below:

Data Generation and Scaling: Before the training of the ANN, the data from the simulator (the input and desired output values) need to be scaled in order to achieve a better estimation of the ANNs internal parameters. Scaling is achieved by dividing each input and output data value from Petrosim® by the maximum value

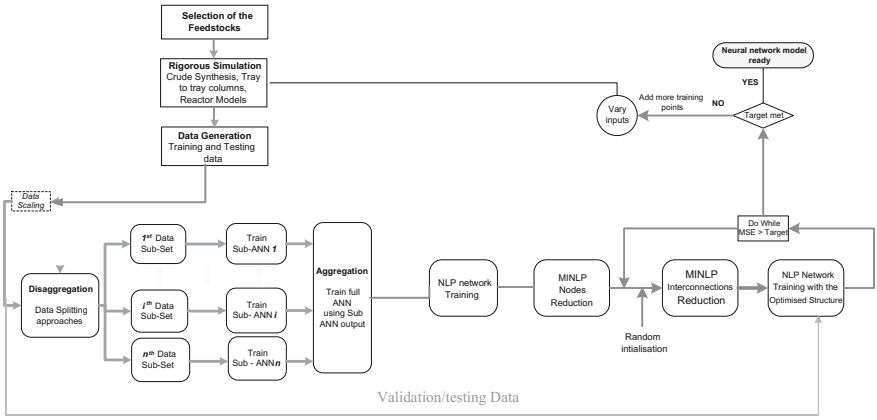

Fig. 5.10 Model reduction process

within the dataset of that variable. All the data values were scaled in the $[0, 1]$ interval.

Disaggregation: The first step consists of splitting the whole dataset using the equal adjacent intervals approach. Furthermore each subset will be allocated to sub-ANN, which will be trained by minimizing the objective function E_1 which is the Sum of Squared Errors (SSE) between the predicted outputs from the ANN and the desired outputs obtained from the simulator. The different sub-ANNs can be trained in parallel.

Aggregation: The full ANN will be trained using the outputs from all the sub-ANNs simultaneously. Note that the sub-ANNs were already trained in the Disaggregation step over the full dataset.

Aggregated ANN Model Nodes Reduction (MIPANN): The completion of previous step will allow the reduction to be initialized by identifying a tolerable error value above the E_1 value achieved by NLP training. This Program will minimize a new objective variable E_2 (total number of nodes), whilst keeping the training error under the chosen tolerable error for E_1. In this way, the number of nodes within the network is reduced while keeping the accuracy of the network within the allowable range.

Aggregated ANN Model Interconnection Reduction (MIPANN): This program will first enforce the weights of unused nodes to zero. The next step is to minimize a new objective function E_3 which is the sum of the interconnection binary variables from the inputs to the hidden layer nodes and from the hidden layer to the outputs. The network error is also kept under the chosen error. These operations are carried out in a loop where the weights are initialized randomly during each iteration, the loops stops when the prediction MSE (Mean Square Error) is under the MSE target. We can also run the loop for several iterations and select the best configuration among the different runs.

Training the Aggregated ANN Model with the optimized structure: This program is the same as in the aggregation step. But in this step we use the optimized structure by fixing the binary variables on their optimum values computed in the previous two steps.

Validation test: Interconnections reduction steps are carried out in a loop where the weights are initialized randomly during each iteration. Also for each run SSE are calculated using testing data. The loop stops when the prediction MSE is under the MSE target. We can also run the loop for several iterations and select the best configuration among the different runs.

In the conventional ANN learning method the number of neurons and interconnections are fixed. Model reduction (Nodes and interconnections reduction) is achieved by reducing over parameterization of the ANN by removing redundant nodes and interconnections while the training error is kept under a defined limit. This is achieved by using mixed integer nonlinear programming techniques. In many cases the MIPANN prediction has actually improved compared to the conventional ANN. The simplified configuration will also result in reduced computational effort for computing the outputs for given inputs, e.g. where ANN is used as property and yields prediction correlation for refinery units. Refinery optimization problems are complex and large where computational performance can be an issue. Since we are using the ANN approximate models instead of the rigorous ones in the optimization problems, any further structure reduction with reduced computational effort and without compromising the performance of the network, is highly desirable.

5.5.4 ANN Modeling Results

The training and validation, node reduction and interconnection reduction programs were modelled in GAMSIDE and ran using SBB and SNOPT solvers for the MINLP and NLP formulations respectively. All the runs were done using GAMS version 23.3.2 on Genuine Intel ~2.925 GHz and 4.00 GB RAM.

5.5.4.1 Aggregation Phases Stopping Criteria

After the disaggregation phase, the aggregation phase takes place where a number of data points is selected randomly from each disaggregation subset (e.g. 20 points selected randomly from the first disaggregation dataset, 20 points selected from the second one and so on). The sensitivity analysis is carried out in Sect. 5.4 showed that beyond a minimum number of data points, the network performances is not significantly impacted. Therefore, to save computational time further, the data points selected in the aggregation (AGG) has been minimized.

To ensure the network training error does not deteriorate in the nodes reduction aggregation (AGG-1) and interconnections reduction (AGG-2) aggregation an inequality constraint was added to their mathematical formulation.

$$E_{Agg-1,2} \leq SSE_{Agg} \tag{5.41}$$

where $E_{AGG-1,2}$ are the training errors from AGG-2 and AGG-3 and SSE_{AGG} is training error from AGG.

When the network structure has been optimized, the aggregation (AGG-3) is re-run now with the new structure (reduced nodes and interconnections). The first step is to run Agg-3 with the inclusion of the following constraint:

$$E_{Agg-3} \leq SSE_{Agg} \tag{5.42}$$

The prediction and training errors from AGG-2 and AGG-3 are then compared and the best network is then selected. When both training and prediction error of AGG-2 are lower than AGG-3, then AGG-2 is the network to be selected. Also when both training and prediction error of AGG-3 are lower than AGG-2, then AGG-3 is the network to be selected this time. Two cases remain when one of the networks has the best training error but not the best prediction and vice versa. To tackle these two cases, we propose to introduce a new metric α defined as:

$$\alpha_i = \sum_{k=1}^{Ng} (\tilde{u}_{k,s1} - u_k)^2 \tag{5.43}$$

where $Ng = N_d \cup N_t$ and $N_d \cap N_t = \emptyset$, where N_d is the aggregation training dataset and N_t is the testing dataset. Therefore α can be written as:

$$\alpha = \sum_{k=1}^{Nd} (\tilde{u}_k - u_k)^2 + \sum_{k=1}^{Nt} (\tilde{u}_k - u_k)^2 \tag{5.44}$$

Hence:

$$\alpha_i = SSE_i + PP_i \tag{5.45}$$

where SSE_i is the calculated error over the training dataset for network i while PP_i is the error for the same network calculated over the testing dataset for network i (both testing and training datasets are separate with no data overlaps). $\boldsymbol{\alpha}$ is the error calculated over the entire dataset combining training and testing datasets. When comparing SSE and PP is not conclusive α is calculated to provide a network error over the entire dataset (including testing dataset) and compare both AGG-2 and AGG3 on a same basis.

The algorithm below summarizes the different four cases discussed above:

IF $\mathrm{SSE}_2 < SSE_3$ **AND** $\mathrm{PP}_2 < PP_3$
Select AGG − 2
ELSE IF $\mathrm{SSE}_3 < SSE_2$ **AND** $\mathrm{PP}_3 < PP_2$
Select AGG − 3
ELSE
$\alpha_2 = \mathrm{SSE}_2 + \mathrm{PP}_2$
$\alpha_3 = \mathrm{SSE}_3 + \mathrm{PP}_3$

IF $\alpha_2 < \alpha_3$
Select AGG − 2
ELSE IF $\alpha_3 < \alpha_2$
Select AGG − 3
END

Sometimes, the error upper bound additional constraint can prove AGG-3 to be infeasible. In that case, AGG-3 is solved with a relaxed error constraint. In that case, predictions from AGG-2 and relaxed AGG-3 are compared and best prediction network is then selected. Figure 5.11 above summarizes that.

5.5.4.2 Continuous Catalytic Reformer Unit ANN Modeling Results

Detailed model of a continuous catalytic reformer unit will comprise of many variables and nonlinearities to capture the inherent nonlinear behavior of kinetic

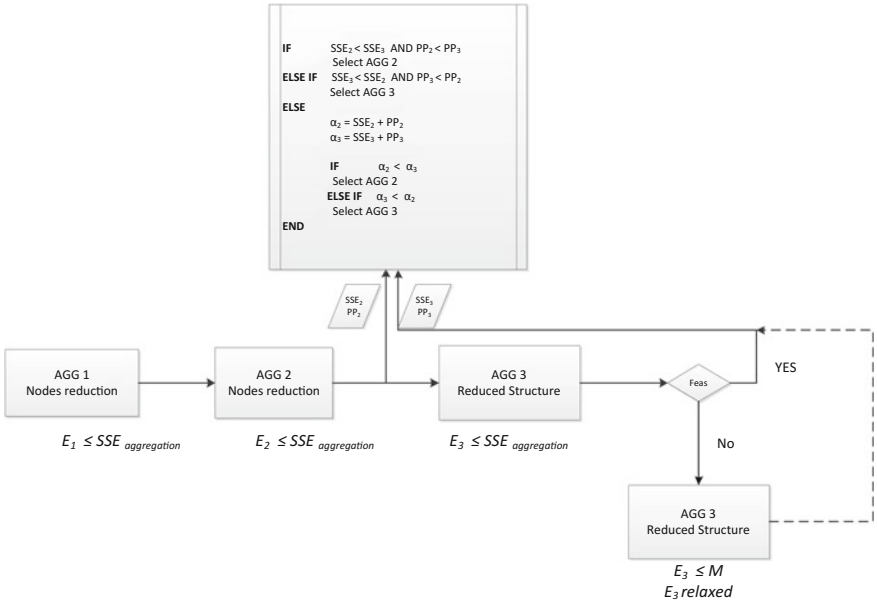

Fig. 5.11 Aggregation phases stopping criteria

reactions. Generally, the continuous catalytic reactor is modeled using kinetic or unit-level model. In kinetic models detailed kinetic analysis of reaction mechanism and catalytic behavior of the process is used to model the CCR. Kinetic models are developed by conducting series of experiments on chemical compounds in order to determine their reaction kinetics. On the other hand, unit-level models are developed by integrating kinetic models in the framework of pilot-scale or commercial reactors (Chang et al. 2013). For a detailed review of catalytic reforming model refer to (Chang et al. 2013; Arani et al. 2009; Taskar 1996).

A reduced nonlinear model will be developed using input-output data generated by a rigorous simulator to replace the rigorous nonlinear models found in the CCR. A new ANN training technique proposed in the previous chapter will be used to train the CCR ANN model. The original data obtained from a rigorous simulator consists of 595 data points consisting of 16 input and 16 output variables. Tables 5.1 and 5.2 summarize inputs and output variables and corresponding values of one data-point.

Both ANN input and outputs node are determined by CCR inputs and output variables. But choice of number of nodes in hidden layer is usually fixed a priori. In practice, different network topologies are used to train ANN by trial-and-error until ANN topology that gives the best prediction of unseen data is obtained. In this work, instead of the trial-and-error approach, a mixed integer programming technique is used to eliminate redundant nodes and interconnections consecutively.

Table 5.1 CCR input variables for different feed stream at one data-point

Heavy naphtha stream	Input values
C6 naphthenes	0.0827
Benzene	0.0305
C7 Aromatics	0.5768
C7 paraffins	14.9197
C7 naphthenes	11.3812
C8 paraffins	22.4372
C8 naphthenes	9.3068
C8 aromatics	2.7814
C9 naphthenes	15.4065
C9 paraffins	27.8727
C9 aromatics	0.3372
Motor octane number (Symonds)	36.7500
Research octane number (RON)	39.0657
Reid vapour pressure (RVP)	0.9568
Specific gravity	0.7353
Reactor	**Input value**
Severity (target RON)	90

Table 5.2 CCR output variable for different product stream atone data-point

Hydrogen stream	Output values
Yield	0.4104
Liquefied petroleum gas (LPG)	**Output values**
Yield	0.0300
C1 %	5.2353
C2 %	0.3749
iC3 %	24.3015
iC4 %	27.9517
nC4 %	39.5731
Reformate stream	**Output values**
Yield	0.8206
Benzene	0.1980
Aromatics (ARO)	58.1326
Naphthenes (NAP)	1.4376
Paraffin (PAR)	38.8230
Motor octane number (Symonds)	81.9758
Research octane number (RON)	90.6474
Reid vapor pressure (RVP)	2.9120
Specific petroleum gravity (SPG)	0.8005

Continuous Catalytic Reformer Unit ANN Modeling Results

The aim of this section is to assess the impact of applying disaggregation–aggregation mixed integer programming (DA-MIPANN) model reduction technique on a previously trained ANN structure through disaggregation–aggregation only described in Sect. 5.5.

A new neural network model of CCR is trained using DA-MIPANN method. The training and computational results are summarized in Tables 5.3, 5.4, 5.5, 5.6, 5.7 and 5.8

Agg-ANN: aggregation ANN. AGG-1: Elimination of nodes in aggregated ANN trained by minimizing number of nodes, AGG-2: aggregated ANN trained by minimizing number of interconnections, Agg-ANN3: new aggregated simplified ANN trained using optimum network configuration.

The full dataset was split into three subsets and a sub-ANN trained for each subset. An aggregated ANN is trained by randomly selecting 60 data points from each subset. Afterwards, redundant nodes and interconnections in the aggregated ANN is eliminated consecutively by minimizing sum of number of nodes and interconnections, resulting into two new ANN model AGG-1 and AGG-2 respectively while maintaining the training error below a predefined tolerance. With the determination of optimal network configuration, a simplified ANN (AGG-3) is trained using the optimal network configuration.

Table 5.3 shows the training error of ANN trained with only disaggregation-–aggregation and disaggregation–aggregation mixed integer-programming method.

Table 5.3 CCR ANN training results for different training techniques

ANN	Data points	DAMIPANN		Data points	DA	
		SSE	MSE		SSE	MSE
Sub-ANN 1	198	14.008	0.066	198	13.998	0.066
Sub-ANN 2	198	5.495	0.042	198	5.595	0.042
Sub-ANN 3	198	0.813	0.016	198	0.813	0.016
Agg-ANN	60	0.795	0.029	595	13.661	0.038
AGG-1	60	0.795	0.029			
AGG-2	60	0.795	0.029			
AGG-3	60	0.795	0.029			

Table 5.4 CCR Initial ANN structure

ANN training technique	Inputs	Outputs	Initial node	Initial interconnection
Dis-aggregation MIPANN	16	16	10	320
Dis-aggregation ANN	16	16	10	320

Table 5.5 Summary ANN results with optimal network configuration

ANN training technique	Reduced nodes	Reduced IC	% Reduction on nodes (%)	% Reduction on IC (%)	SSE	MSE
Dis-aggregation MIPANN	6	69	40	78	0.795	0.029
Dis-aggregation ANN (base)	10	320	0	0	13.661	0.038

Table 5.6 Comparison of computational time for different ANN training methods

ANN	Data points	DAMIPANN	Data points	DA
		CPU		CPU
Sub-ANN 1	198	4.406	198	2.125
Sub-ANN 2	198	7.750	198	9.359
Sub-ANN 3	198	11.813	198	11.34
Agg-ANN	60	2.594	595	41.62
AGG-1		16.797		
AGG-2		48.844		
AGG-3		4.266		
Total CPU		84.314		53.016

Table 5.7 Comparison of unseen data prediction error for different ANN training methods

ANN	Data points	DAMIPANN (from AGG-2)	DAMIPANN (from AGG-3)	Disaggregation–aggregation
Prediction SSE	94	27.301	12.179	4.653
Prediction MSE	94	0.135	0.090	0.056

Table 5.8 Comparison of AGG-2 and AGG-3

ANN	DAMIPANN (from AGG-2)	DAMIPANN (from AGG-3)	Disaggregation–aggregation
Total data points (Aggregation training and testing data points)	154	154	689
Prediction (PP)	27.301	12.179	4.653
Training (SSE)	0.795	0.795	13.661
α	28.096	12.974	18.314
MSE(α)	0.107	0.07	0.04

In both cases the sum squared error (SSE) and mean squared error (MSE) is tabulated. The result clearly shows that the simplified ANN (AGG-3) trained using optimal network results in an improvement in MSE from 0.038 down to 0.029 when compared to using disaggregation–aggregation based method (Agg-ANN).

Note that when solving AGG-1 and AGG-2, extra constraints are added to remove redundant nodes and interconnections. The error is then likely to reduce. Therefore, by adding extra constraint:

$$E_{Agg-1,2,3} \leq SSE_{Agg} = 0.795$$

We aim to optimise network's structure while keeping training error below the initial aggregation error. It the current case, in both AGG-1 and AGG-2 this constraint has been satisfied which explains why AGG, AGG-1,2,3 have all similar SSE/MSE.

Tables 5.4 and 5.5 show the initial ANN configuration and optimum ANN configuration respectively obtained using the proposed ANN training method. The proposed ANN training approach reduces the number of nodes and interconnection by 40 % and 78 % respectively when compared to training the ANN with only disaggregation–aggregation training approach; thereby producing network with simpler configuration.

Agg-ANN: aggregation ANN. AGG-1: Elimination of nodes in aggregated ANN trained by minimizing number of nodes, AGG-2: aggregated ANN trained by minimizing number of interconnections, Agg-ANN3: new aggregated simplified ANN trained using optimum network configuration.

Table 5.6 shows that disaggregation–aggregation mixed integer programming (DAMIPANN) training approach increases the overall computational time when compared to only using disaggregation–aggregation training method by a factor of 1.6. This can be attributed to the additional steps involved in the proposed technique. However, the computational time of the simplified ANN model (AGG-3 in Table 5.6) obtained using DAMIPANN is significantly less than the original aggregated ANN model (Agg-ANN in Table 5.6). The implication is that the proposed approach has further reduced the ANN computational complexity, but at the expense of a slight drop in the ANN performance in predicting unseen data.

According to the AG-2/3 stopping criteria proposed in 5.4.1, AGG-3 and AGG-2 have similar SSE but AGG-3 have better prediction therefore it should be selected. Calculation of α in this case is not required, but is provided as an example. Comparing training and testing errors for AGG-2 and AGG-3 as shown in Table 5.8 shows that the global network error, calculated as a sum of *SSE* and *PP* for AGG-3 is better than AGG-2. Therefore for any further use of the model, AGG-3 should be selected.

Continuous Catalytic Reformer Unit Network

In Figs. 5.12 and 5.13 a graphical representation of the CCR ANN before and after the application of the DAMIPANN model reduction technique is presented.

5.5.4.3 Naphtha Splitter Unit Artificial Neural Network Model

Naphtha is a generic term used to describe refined, partly refined or unrefined petroleum products and liquid product which has a cut point between 80 and 180 °C (Pandey et al. 2004). Naphtha produced from the crude distillation unit is further processed in a naphtha splitter unit into light naphtha stream and heavy naphtha stream; afterwards the heavy naphtha stream is then charged into a catalytic reformer unit. Because, it is not economical to process the light straight-run gasoline cuts (C_5 – 88 °C) in the catalytic reformer.

A reduced nonlinear model will be developed to replace the rigorous non-linear models found in naphtha splitter. Data generated from the rigorous simulator consists of 519 data points comprising of 15 inputs variable and 26 output variables. Tables 5.8 and 5.9 summarize inputs and output variables and corresponding values of one data-point.

Both ANN input and outputs node are determined by naphtha splitter inputs and output variables. But choice of number of nodes in hidden layer is usually fixed in priori. In practice, different network topology is used to train a network by

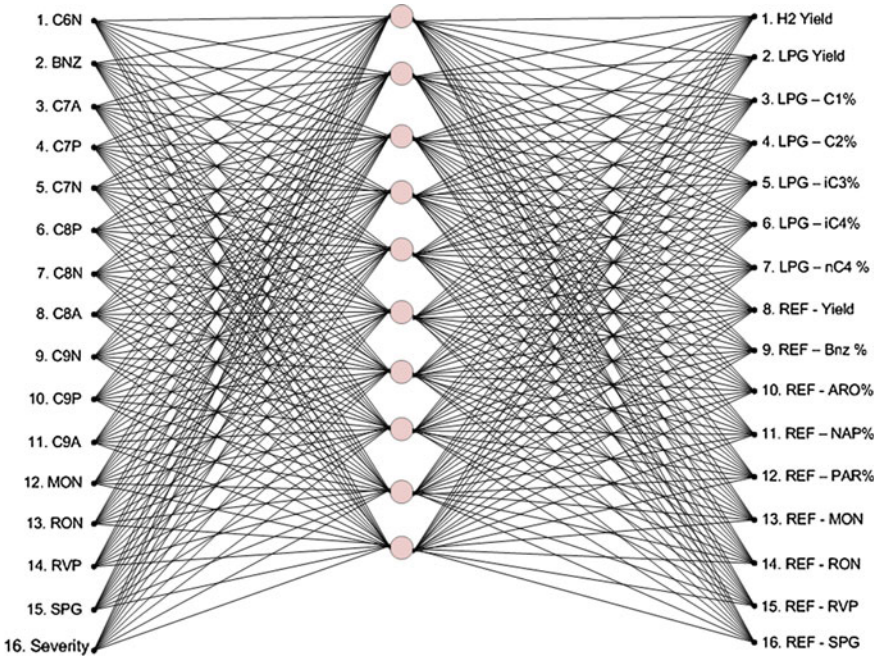

Fig. 5.12 CCR ANN before DAMIPANN

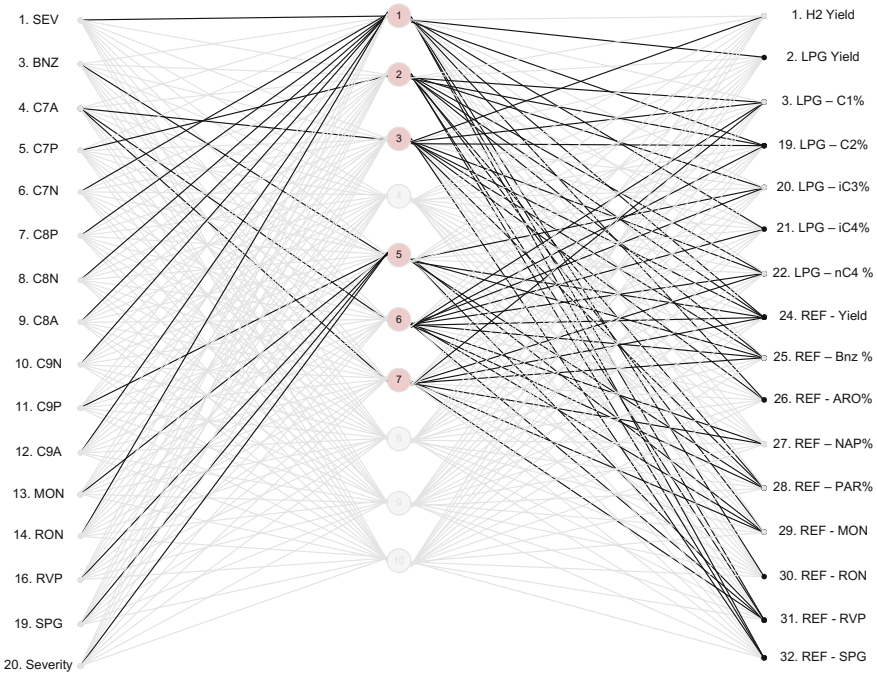

Fig. 5.13 CCR ANN after DAMIPANN

Table 5.9 Naphtha splitter input variables for different feed stream in one data-point

Naphtha feed	Input values
C_6 paraffin	9.6140
C_6 naphthenes	2.9252
C_7 paraffins	13.4719
C_7 naphthenes	8.8920
C_8 paraffins	16.1089
C_8 naphthenes	6.6819
C_8 aromatics	1.9969
C_9 aromatics	0.2421
C_9 paraffin	20.0114
C_9 naphthenes	11.0612
Motor octane number (MON)	47.6576
Research octane number (RON)	48.6971
Reid vapour pressure (RVP)	11.2711
Specific gravity	0.7111
Splitter	
Naphtha splitter cut point	68

trial-and-error until ANN topology that gives the best prediction of unseen data is obtained. In this work instead of trial-and-error approach, a mixed integer programming technique is used to eliminate redundant nodes and interconnections consecutively.

ANN Modeling Results for Naphtha Splitter Units

The aim of this section is to assess the impact of applying DAMIPANN model reduction technique on a reduced ANN structure (less nodes and interconnections). A new neural network of Naphtha splitter is modeled and trained using disaggregation–aggregation mixed integer programming (DAMIPANN) method. The training and computational results are summarized in Tables 5.10, 5.11, 5.12, 5.13, 5.14, 5.15 and 5.16.

Agg-ANN: aggregation ANN. AGG-1: Elimination of nodes in aggregated ANN trained by minimizing number of nodes, AGG-2: aggregated ANN trained by minimizing number of interconnections, AGG-3: new aggregated simplified ANN trained using optimum network.

The full dataset was split into three subsets and a sub-ANN trained for each subset. An aggregated ANN is trained by randomly selecting 60 data points from each subset. Afterwards, redundant nodes and interconnections in the aggregated ANN is eliminated consecutively by minimizing sum of number of nodes and interconnections, resulting into two new ANN model AGG-1 and AGG-2 respectively while maintaining the training error the aggregation training error (error for

Table 5.10 Naphtha splitter output variables in different product streams

Light naphtha stream	Output values
Yield	0.2124
Benzene	0.1776
C_6 paraffin	33.8108
C_6 naphthenes	8.5376
C_7 paraffin	0.1309
C_7 naphthenes	0.0342
Motor octane number (MON)	77.2992
Research octane number (RON)	78.0817
Specific petroleum gravity (SPG)	0.6331
Raid vapour pressure	37.2907
Heavy naphtha stream	**Output values**
Yield	0.7876
Benzene	0.0577
C_6 paraffin	3.0883
C_6 naphthenes	1.4116
C_7 paraffin	17.0698
C_7 naphthenes	11.2808
C_8 paraffins	20.4534
C_8 naphthenes	8.4839
C_8 aromatics	2.5355
C_9 aromatics	0.3073
C_9 paraffin	25.4083
C_9 naphthenes	14.0443
Motor octane number (MON)	38.5502
Research octane number (RON)	40.7723
Specific petroleum gravity (SPG)	0.7322
Raid vapour pressure	1.2857

Table 5.11 Naphtha splitter ANN training results

ANN	Data points	DAMIPANN		Data points	DA	
		SSE	MSE		SSE	MSE
Sub-ANN 1	173	5.084	0.034	173	5.380	0.034
Sub-ANN 2	173	4.997	0.033	173	4.864	0.033
Sub-ANN 3	173	5.884	0.036	173	5.884	0.036
Agg-ANN	60	1.495	0.031	519	11.186	0.028
AGG-1	60	1.495	0.031			
AGG-2	60	1.495	0.031			
AGG-3 $E_{Agg-3} \leq SSE_{Agg}$	60	INFEASIBLE				
AGG-3	60	10.867	0.083			

Table 5.12 Naphtha splitter initial ANN structure

ANN training technique	Inputs	Outputs	Initial node	Initial Interconnection (IC)
Dis-aggregation MIPANN	15	26	15	615
Dis-aggregation ANN	15	26	15	615

Table 5.13 Naphtha splitter summary ANN results with optimal network configuration

ANN training technique	Reduced nodes	Reduced IC	% Reduction on nodes (%)	% Reduction on IC (%)	SSE	MSE
Dis-aggregation MIPANN	8	71	47	88	1.495	0.031
Dis-aggregation ANN (base)	15	615	0	0	11.186	0.028

Table 5.14 Comparison of computational time for different naphtha ANN training method

ANN	Data points	DAMIPANN (s)	Data points	Disaggregation–aggregation (s)
Sub-1	173	32.125	173	31.766
Sub-2	173	26.203	173	28.844
Sub-3	173	30.828	173	30.625
Agg-ANN	60	5.328	519	167.59
AggANN	60	50.719		
Agg- ANN 2	60	287.125		
AGG-3	60	0.844		
Total		376.14		199.356

Table 5.15 Comparison of unseen data prediction error for different ANN training methods

ANN	Data points	DAMIPANN (Agg-ANN 2)	DAMIPANN (Agg-ANN 3)	Disaggregation–aggregation
Prediction SSE	134	121.158	33.172	48.620
Prediction MSE	134	0.186	0.098	0.118

Agg-ANN that is 1.495). With the determination of optimal network configuration, a simplified ANN (AGG-3) is trained using the optimal network configuration.

Table 5.11 shows the training error of ANN trained with only disaggregation–aggregation and disaggregation–aggregation mixed integer-programming method. In both cases the sum squared error (SSE) and mean squared error

Table 5.16 Comparison between AGG-2 and AGG-3 for Naphtha Splitter

ANN	DAMIPANN (from AGG-2)	DAMIPANN (from AGG-3)	Disaggregation–aggregation
Total data points (Aggregation training and testing datapoints)	194	194	653
Prediction (PP)	121.158	33.172	48.620
Training (SSE)	1.495	10.867	11.186
α	122.653	44.039	59.806
MSE(α)	0.156	0.093	0.059

(MSE) is tabulated. The result shows that the simplified ANN (AGG-2) trained using optimal network results in a MSE (0.031) that is comparable to the MSE when using disaggregation–aggregation based method (0.028). Running AGG-3 with the error upper bound constrained has led to infeasibility. Therefore, AGG-3 was ran with a relaxed error constraint, which resulted, into a MSE of 0.083. Analyzing prediction performances over unseen data will therefore be required before conclusion can be drawn in regards to selection of AGG-2 or AGG-3.

Tables 5.12 and 5.13 shows naphtha splitter initial ANN structure and naphtha splitter summary ANN results with optimal network configuration obtained using the proposed ANN training method. The proposed ANN training approach reduces the number of nodes and interconnections by 47 % and 88 % respectively; thereby producing network with simpler configuration.

Agg-ANN: aggregation ANN. AGG-1: Elimination of nodes in aggregated ANN trained by minimizing number of nodes, AGG-2: aggregated ANN trained by minimizing number of interconnections, Agg-ANN3: new aggregated simplified ANN trained using optimum network configuration.

Table 5.14 shows that disaggregation–aggregation mixed integer programming based approach for ANN increases the overall computational time when compared to only using disaggregation–aggregation training method by a factor of 1.9. This can be attributed to the additional steps involved in the proposed technique. However, the computational time of the simplified ANN model (AGG-3 in Table 5.14) is significantly less than the original aggregated ANN model (Agg-ANN in Table 5.14). This implies that DAMIPANN produces reduced ANN models that are less computational expensive.

Data generated from rigorous simulator was split into two sets: ANN training dataset and ANN testing dataset (data not used for ANN training). A statistical measure, sum squared error (SSE) and mean squared error (MSE) was used to determine the performance of our ANN model over unseen data. Both sum squared error and mean squared error is a measure that indicates how closely the predicted ANN output is to testing dataset. Table 5.15 shows that the proposed ANN training method resulted in a further improvement of prediction SSE and MSE.

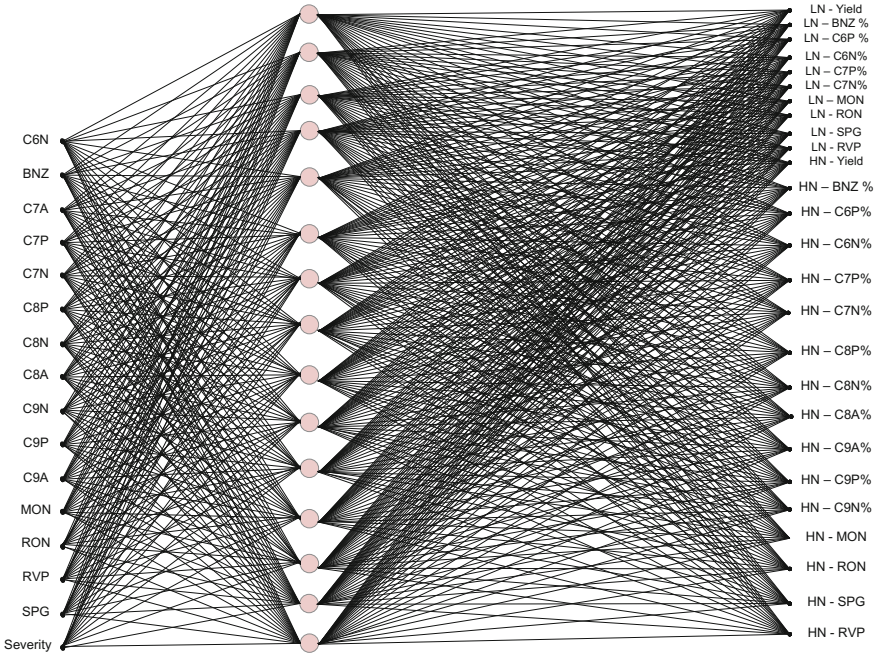

Fig. 5.14 Naphtha Splitter ANN before DAMIPANN

Comparing training and testing errors for AGG-2 and AGG-3 as shown in Table 5.16 shows that although the global network error, calculated as a sum of *SSE* and *PP* for AGG-3 is better than AGG-2. Therefore for any further use of the model, AGG-3 should be selected.

Naphtha Splitter Unit Network

Below in Fig. 5.14 and 5.15 a graphical representation of the Naphtha Splitter ANN before and after the application of the DAMIPANN model reduction technique is provided.

5.5.5 Summary

The combined disaggregation–aggregation mixed integer programming ANN reduction technique was presented in this chapter, consisting of using the previously used data (in Sects. 5.3 and 5.4) from a rigorous simulator to generate a set of

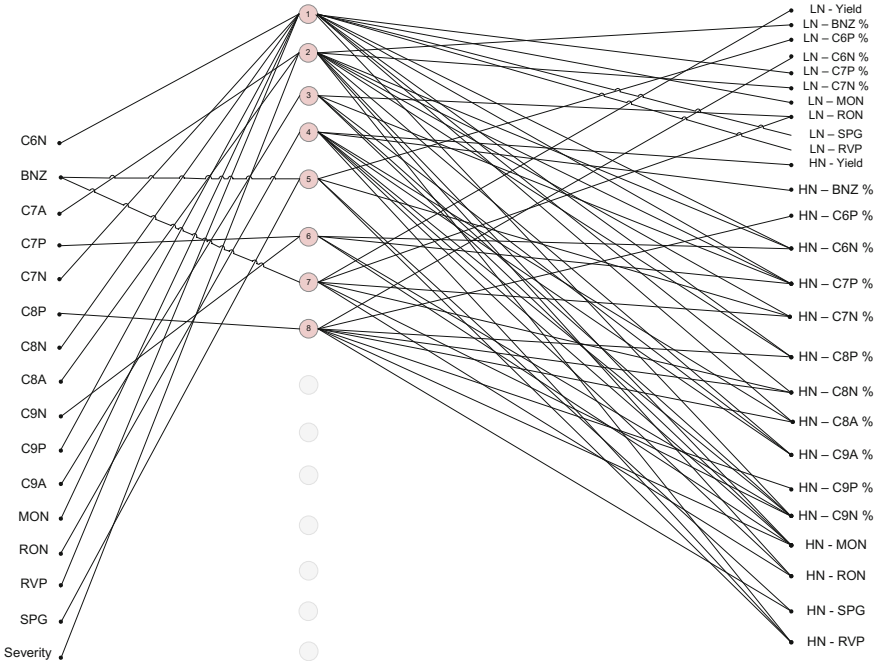

Fig. 5.15 Naphtha Splitter ANN after DAMIPANN

data that is used to train and to test large or medium neural networks. This is achieved through firstly training sub-ANNs using subset data instead of the full range data, then combining outputs from the Sub-ANNs into a single aggregated ANN model (Disaggregation Aggregation technique) followed by reduction of redundant nodes and interconnections in the aggregation model. This is achieved through solving Mixed Integer nonlinear Programming model (MIPANN technique).

This has resulted in a trained network over medium to large datasets with reduced structure presenting computational benefits when using the model in an enterprise-wide optimization models. Predictions performances over unseen data with DAMIPANN are comparable if not better than the standalone Disaggregation–Aggregation reduction. Future work will aim to address the impact of initial ANN structure on the reduced network performance when applying DAMIPANN model reduction technique.

Acknowledgments Financial support from EPSRC (EP/G059195/1) is gratefully acknowledged. The authors would like to thank KBC Process Technology Ltd. for allowing them to use Petrosim® for generating simulation data for training ANN models.

Table 5.17 Inputs and outputs of the CCR model

ANN	Stream	Physical properties[a]	Variables
Inputs	Heavy Naphtha	Benzene, C6P, C7P, C7N, Toluene, C8P, C8N, C8A, C9P, C9N, C9A, MON, RON, RVP, SPG	16
	Reactor	Severity	
Outputs	H2	Yield	16
	LPG	Yield, C1%, C2%, C3%, iC4%, nC4%	
	Reformate	Yield, Benzene, ARO, NAP, PAR, MON, RON, RVP, SPG	

[a]The description of the physical properties can be found below

Appendix 1: ANN Modeling—Input and Output Variables Description

The first step taken to model the naphtha splitter and the CCR units is to select the key variables and operating conditions of these units, which will be used in the modeling.

The Naphtha splitter and the CCR are in series so we have to make sure the inputs of the CCR are well predicted in the outputs of the naphtha splitter. Tables 5.17 and 5.18 summarize the inputs and outputs variables to be considered for the modeling section.

Table 5.18 Inputs and outputs of the Naphtha Splitter model

ANN	Stream	Physical properties[a]	Variables
Inputs	Naphtha Feed	Benzene, C6P, C7P, C7N, Toluene, C8P, C8N, C9A, MON, RON, RVP, SPG	13
	Splitter	Naphtha splitter cut point	
Outputs	Light Naphtha	Yield, Benzene, C6P, C6N, C7P, C7N, MON, RON, RVP, SPG	26
	Heavy Naphtha	Yield, Benzene, C6P, C6N, C7P, C7N, Toluene, C8P, C8N, C8A, C9P, C9N, C9A, MON, RON, RVP, SPG	

[a]The description of the physical properties can be found below
C6P C6 paraffins, vol.%
C6N C6 Naphthenes, vol.%
Bnz Benzene, vol.%
C7P C7 paraffins, vol.%
C7N C7 Naphthenes, vol.%
C8P C8 paraffins, vol.%
C8N C8 Naphthenes, vol.%
C8A C8 Aromatics, vol.%
C9P: C9 paraffins, vol.%
C9N C9 Naphthenes, vol.%
C9A C9 Aromatics, vol.%
MON Motor Octane Number
RON Research Octane Number
RVP Reid Vapor Pressure
SPG Specific Gravity

References

Alattas, A. M., Grossmann, I. E., & Palou-Rivera, I. (2011). Integration of nonlinear crude distillation unit models in refinery planning optimization. *Industrial and Engineering Chemistry Research, 50*(11), 6860–6870.

Ancheyta, J., Sanchez, S., & Rodriguez, M. A. (2005). Kinetic modeling of hydrocracking of heavy oil fractions: A review. *Catalysis Today, 109*(1), 76–92.

Ancheyta-Juarez, J., & Villafuerte-Macias, E. (2000). Kinetic modeling of naphtha catalytic reforming reactions. *Energy & Fuels, 14*(5), 1032–1037.

Arani, H., Shirvani, M., Safdarian, K., & Dorostkar, E. (2009). Lumping procedure for a kinetic model of catalytic naphtha reforming. *Brazilian Journal of Chemical Engineering, 26*(4), 723–732.

Caballero, J. A., & Grossmann, I. E. (2004). Design of distillation sequences: From conventional to fully thermally coupled distillation systems. *Computers & Chemical Engineering, 28*(11), 2307–2329.

Chang, A. F., Pashikanti, K., & Liu, Y. A. (2013). *Refinery engineering: Integrated process modeling and optimization*. Wiley.

Del Bianco, A., Panariti, N., Anelli, M., Beltrame, P., & Carniti, P. (1993). Thermal cracking of petroleum residues: 1. Kinetic analysis of the reaction. *Fuel, 72*(1), 75–80.

Dua, V. (2010). A mixed-integer programming approach for optimal configuration of artificial neural networks. *Chemical Engineering Research and Design, 88*(1), 55–60.

Dua, V., Papalexandri, K. P., & Pistikopoulos, E. (2004). Global optimization issues in multiparametric continuous and mixed-integer optimization problems. *Journal of Global Optimization, 30*(1), 59–89.

Favennec, J.-P. (2001). *Petroleum refining: Refinery operation and management* (Vol. 5). Editions Technip.

Gadalla, M., Jobson, M., & Smith, R. (2003). Optimization of existing heat-integrated refinery distillation systems. *Chemical Engineering Research and Design, 81*(1), 147–152.

Gary, J. H., & Handwerk, G. E. (1994). *Petroleum refining, technology and economics*, *Petroleum refining crude oil* (4th ed.).

Georgiadis, M. C., Pistikopoulos, E. N., & Vivek, D. (2008). *Energy systems engineering* (Vol. 5). Wiley-VCH Verlag GmbH.

Gueddar, T., & Dua, V. (2011). Disaggregation–aggregation based model reduction for refinery-wide optimization. *Computers & Chemical Engineering, 35*(9), 1838–1856.

Gueddar, T., & Dua, V. (2012). Novel model reduction techniques for refinery-wide energy optimisation. *Applied Energy, 89*(1), 117–126.

Guerra, O. J., & Le Roux, G. A. (2011). Improvements in petroleum refinery planning: 1. Formulation of process models. *Industrial and Engineering Chemistry Research, 50*(23), 13403–13418.

Himmelblau, D. M. (2008). Accounts of experiences in the application of artificial neural networks in chemical engineering. *Industrial and Engineering Chemistry Research, 47*(16), 5782–5796.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks, 2*(5), 359–366.

Hussain, M., & Kershenbaum, L. (2000). Implementation of an inverse-model-based control strategy using neural networks on a partially simulated exothermic reactor. *Chemical Engineering Research and Design, 78*(2), 299–311.

Mendez, C. A., Grossmann, I. E., Harjunkoski, I., & Kaboré, P. (2006). A simultaneous optimization approach for off-line blending and scheduling of oil-refinery operations. *Computers & Chemical Engineering, 30*(4), 614–634.

Pandey, S. C., Ralli, D. K., Saxena, A. K., & Alamkhan, W. K. (2004). Physicochemical characterization and applications of naphtha. *Journal of Scientific & Industrial Research, 63*(3), 276–282.

Parkash, S. (2003). *Refining processes handbook*. Gulf Professional Publishing.

Powell, R. T., & Yu, C.-Y. (2004). Refinery reaction modelling trends. *Hydrocarbon Engineering*, *21*.

Slaback, D. D., & Riggs, J. B. (2007). The inside-out approach to refinery-wide optimization. *Industrial and Engineering Chemistry Research, 46*(13), 4645–4653.

Sundaram, K., & Froment, G. (1977). Modeling of thermal cracking kinetics—I: Thermal cracking of ethane, propane and their mixtures. *Chemical Engineering Science, 32*(6), 601–608.

Symonds, G. H. (1955). *Linear programming: The solution of refinery problems*. Esso Standard Oil Company.

Taskar, U. M. (1996). *Modeling and optimization of a catalytic naphtha reformer.* Doctoral dissertation. Texas Tech University.

Theologos, K., & Markatos, N. (1993). Advanced modeling of fluid catalytic cracking riser-type reactors. *AIChE Journal, 39*(6), 1007–1017.

Umana, B., et al. (2014). Integrating hydroprocessors in refinery hydrogen network optimisation. *Applied Energy, 133*, 169–182.

Van Landeghem, F., Nevicato, D., Pitault, I., Forissier, M., Turlier, P., Derouin, C., et al. (1996). Fluid catalytic cracking: Modelling of an industrial riser. *Applied Catalysis A: General, 138*(2), 381–405.

Venkatasubramanian, V., & Chan, K. (1989). A neural network methodology for process fault diagnosis. *AIChE Journal, 35*(12), 1993–2002.

Yang, Y., & Barton, P. I. (2015). Refinery optimization integrated with a nonlinear crude distillation unit model.

Yang, Y., & Barton, P. I. (2015). Integrated crude selection and refinery optimization under uncertainty. *AIChE Journal.*

Zhou, L., et al. (2015). Energy configuration and operation optimization of refinery fuel gas networks. *Applied Energy, 139*, 365–375.

Chapter 6
Coal Staged Conversion Polygeneration Technology Combining with Pyrolysis and Combustion Processes

Qinhui Wang

Abstract Coal is the main primary energy in the world. However, most coal is combusted to generate electric and heat, which does not make effective of valuable hydrogen-rich volatile matter. A coal staged conversion process,which coupled circulating fluidized bed (CFB) combustion and coal pyrolysis,is a more efficient and clean technology for coal conversion and the principle is: part of the high temperature circulating material from CFB combustor as solid heat carrier is sent to the pyrolyzer to provide heat for coal pyrolysis and the coal incurs pyrolysis, producing tar and gas. The pyrolysis char is returned to CFB for further combustion to generate electricity and provide heat. Thus, poly-generation of gas, tar, heat and power can be realized in this system. According to the bed type of the pyrolyzer, the coal staged conversion technology is mainly divided into three types: the technology based on fluidized bed pyrolysis, that based on moving bed pyrolysis and that based on downer bed pyrolysis, and a lot of research has been done on this. Take the technology by Zhejiang University for example, industrial application and system analysis are discussed in detail. 12 MW and 40 t/h coal staged conversion process is established and operated. The test results showed that this system can realize poly-generation of heat, electricity, gas and tar. When the pyrolyzer does not run, the circulating fluidized bed can still be normally operated. This shows that the coal staged conversion technology can be applied to industrial scale poly-generation plant. Thermodynamic and economic of poly-generation system coupling 2×300 MW circulating fluidized bed and 2 fluidized bed pyrolyzer is analyzed. It shows that the coal staged conversion process system has higher energy and exergy efficiency and more profitable than single traditional 2×300 MW CFB plant. This justified the coal staged conversion process in a large scale is an efficient and economic technology.

Keywords The coal staged conversion process · Pyrolysis · Circulating fluidized bed combustion · System analysis

Q. Wang (✉)
Institute for Thermal Power Engineering, Zhejiang University,
No. 38 Zheda Road, Hangzhou 310027, China
e-mail: qhwang@zju.edu.cn

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_6

6.1 Introduction

The coal reserve is rich in the world and would be the major energy resource in the future in the certain countries like China, of the world. The coal consumption structure in China is shown in Fig. 6.1. Power generation, coking, coal gasification and other industrial applications accounted for a large percentage of coal consumption. In 2013, Chinese coal consumption for power was about 3.04 billion tons, accounting for about 64 % of coal consumption. However, high-value volatile matter is fired in the coal combustion and it does not make effective use of hydrogen-rich components of the coal. In addition, the pollution is diluted in the huge volume of flue gas, which added the difficulty of removing the pollution. So, currently, coal utilization is mainly focused on a single utilization with low efficiency, rather than the comprehensive utilization. Therefore, it is attractive to pursue a new coal conversion technology which can make effective of hydrogen-rich volatile matter and reduce pollution emission.

The compositions of coal are complex, and there are great differences in the chemical reactivity among different constituents of the coal. The hydrogen-rich volatile component is the most active components in the coal and usually released at lower temperatures. In addition, since the properties of volatile matter are similar to petroleum and natural gas, high-value utilization of the volatile matter, like prodicing gas and oil, can be realized more easily. While the fixed carbon is stable and the activation energy of the reaction is higher and it is not easy to react at low temperatures, which added the difficulty of high-value utilization of fixed carbon. Based on this, the idea of coal staged conversion process which extracts hydrogen-rich volatiles of coal prior combustion was born, which can be implemented by coupling coal pyrolysis and char combustion.

Based on the coal staged conversion process, hydrogen-rich volatiles of coal can be extracted into gas and tar in the pyrolysis process and the char is then fired in the combustion process. This technology has a distinct advantage compared with other coal conversion technologies, like coal combustion and coal gasification. As mentioned above, the coal combustion does not make effective use of hydrogen-rich components of the coal and give rises to serious pollution. Although

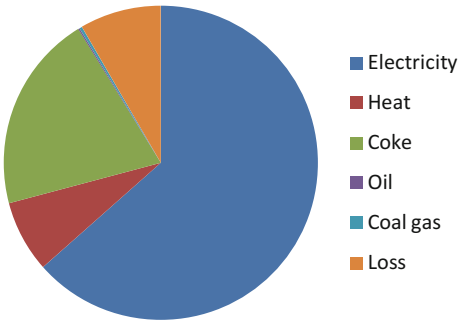

Fig. 6.1 The coal consumption structure in China in 2013

coal gasification, a process to convert coal into fuel gas or syngas, has the potential to allow efficient and clean utilization of coal, the reaction rate of the gasification would become slow with the increase of reaction extent, so high temperature, high pressure, pure oxygen and long residue time are needed to realize full or high conversion during coal gasification, which would cause that the equipment is huge and investment and production cost is high. What's more, coal gasification has high requirement for coal quality and water resource, so many coals and many districts where water is scarce are not suit for complete gasification. If the low active component of the coal is fired rather than gasification, it will greatly reduce costs, for the burning rate of the coal is far higher than that gasification rate. Based on this, the coal staged conversion process can reduce investment and reduce cost by simplifying coal gasification technology and lowering reaction condition compared with coal gasification technology.

The coal staged conversion process technology can realize comprehensive utilization of coal, shown in Fig. 6.2. The volatile matter would be extracted into gas and tar through pyrolysis. And the volatile gas could be used in the production of natural gas, synthetic liquid fuel or hydrogen production. Tar could be extracted for useful chemicals and improving the quality for fuel oil. Char with rich carbon could be used for combustion and the ash can be used for building materials. In addition, it removes some pollution during the pyrolysis and high pollution concentration in the pyrolysis gas makes it easier for pollution removal, which can solve the pollution problems in an economic way. Through the organic coupling of above process in the system, it would simplify process and reduce the cost of investment and operation. The output of the heat, electricity gas and tar can be regulated according to market demand and coal characteristics, thus realizing the highest

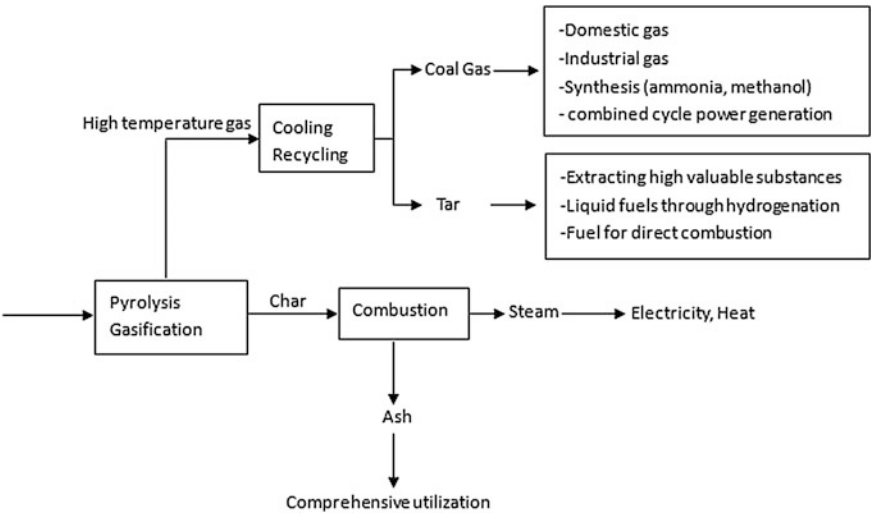

Fig. 6.2 Comprehensive utilization of coal based on coal staged conversion process

profit. Through the coupling pyrolysis-combustion classification transformation, the coal staged conversion process can be applicable to lignite or bituminous coal with different coal quality and achieve co-production of variety of products and reduce the cost, so as to realize comprehensive utilization of coal with staged conversion, improve the coal conversion and utilization efficiency, reduce pollution emissions, realize overall system efficiency optimization of coal utilization.

The coal staged conversion process technology can relieve the shortage of petroleum and natural gas, thus having great potential application in the world. Take China for example, thermal power with a capacity of 765 million kW has following potentials for coal staged conversion process reform. Volatile parts in about 1.8 billion tons of coal consumption (more than 90 % was bituminous coal and lignite) each year could be equivalent to 271.3 billion cubic meters of synthetic natural gas (23 West-East Gas Transmission amounts) or 220 million tons of oil (Chinese oil imports in 2011 was 250 million tons). The coal staged conversion process can not only be applied to the new power plant, but also can be used to reform the existing coal-fired power plants. The popularization and application of the coal staged conversion process technology is significant for the world clean and efficient coal power generation, oil and gas replacing, energy conservation, emission reduction and recycling economy, etc. Therefore, the coal staged conversion process has a wide application prospect.

6.2 The Overview of Main Technologies of Coal Staged Conversion Process Based on Coal Pyrolysis

Since the coupling of many coal conversion technologies in a system can be realized, the coal staged conversion technology coupling with coal pyrolysis and char combustion has been rapidly developed. Among all of the coal staged conversion process technologies, the technology based on circulating fluidized bed combustion is fully recognized in recent years due to the large-scale application of CFB combustion and the advantages of CFB combustion. Circulating fluidized bed (CFB) combustion is a clean coal technology (Wang et al. 1999). With the development of CFB technology in those years, CFB boiler has been widely applied in the industry. Now 600 MW utility boiler based on CFB has been used for commercial use. There is a great deal of high temperature ash in the CFB combustion and part ash from the CFB boiler provides the heat for the pyrolysis. This can solve the problems of low thermal efficiency for traditional heat carrier. The development of CFB combustion and the pyrolysis technology based on solid heat carrier lays solid foundation for the coal staged conversion process coupling pyrolysis and the CFB combustion. Based on this, the coal staged conversion process technology has been rapidly developed in recent years.

The coal staged conversion process based on pyrolysis and CFB combustion has the advantages as follows:

(1) This technology has high thermal efficiency. The heat for the pyrolysis is from the high temperature circulating ash from CFB boiler, and meanwhile the char is directly sent to the boiler. The sensible heat of the ash and char is fully utilized, so the heat loss of this system is low.
(2) This technology has wide fuel adaptation. Inferior fuels (such as gangue and oil shale) can be used in CFB boiler, so those fuel also can be used in the coal staged conversion process.
(3) Part sulfur and nitrogen in the coal can be transformed to the coal gas in the pyrolysis. The volume of coal gas is far less than that of the flue gas, so the removal of sulfur and nitrogen from the coal gas is much easier than that from the flue gas. Therefore, this technology can reduce the burden of desulfurization and denitration.
(4) The output of the heat, electricity gas and tar can be regulated according to market demand, thus realizing the highest profit.

What's more, this technology not only can be used for new plant, but also can be used to transform the existing CFB plant. The pyrolyzer and CFB boiler are independent, so the CFB boiler can run normally when there is a failure for the pyrolysis process.

The pyrolysis process is a key part in the coal staged conversion process, the main technologies of coal staged conversion process in China mainly consists of the technology based on fluidized bed pyrolysis, that based on moving bed pyrolysis and that based on downer pyrolysis (Guo 2015). A lot of studies have been carried out about the coal staged conversion process and a brief introduction will be given below.

6.2.1 The Coal Staged Conversion Process Based on Fluidized Pyrolysis

The fluidized bed is used for pyrolysis in the coal staged conversion process based on fluidized bed pyrolysis. It has many advantages: (1) The material in the fluidized bed can be fully mixed and the temperature is uniform, which provide good reaction condition for the pyrolysis. (2) The gasification intensity is high and the volume of the pyrolyzer is small, which makes the pyrolyzer easier to scale up. (3) Coal adaptability is wide. The pyrolyzer and boiler both are fluidized beds and fluidized bed has the characteristics of wide fuel adaptability. However, it also has some disadvantages, such as a great deal of fluidized gas is needed to fluidize the bed materials which can cause high heat loss. What is more, much fly ash and fine particle are taken away from the pyrolyzer, which make the purification for the gas and tar more complicated.

Zhejiang University, Tsinghua University and Institute of Engineering Thermal Physics have done a lot of research about this technology. Take the technology by Zhejiang University for example, the coal staged conversion process based on fluidized pyrolysis is introduced in detail.

Proposed by Zhejiang University, the coal staged conversion process combining with fluidized bed pyrolysis and CFB combustion is a domestic new technology with independent intellectual property rights. This technology combines the circulating fluidized bed boiler and the pyrolysis furnace closely, realizes co-production of heat, electricity, gas and tar in one system (Fang et al. 2011; Wang et al. 2002; Guo 2015; Fang et al. 1998), shown in Fig. 6.3. The main process is: operating temperature of circulating fluidized bed boiler is between 850 and 900 °C. A large number of high temperature materials are transported to the pyrolyzer via recycle device to provide heat for the pyrolysis. The coal is sent to the pyrolysis furnace through screw and then mixed with high temperature material. In the mixing process, the coal is heated and undergoes pyrolysis, getting gas, tar and char. (operating temperature for pyrolysis is between 550 and 800 °C). Coal gas and fine ash particles are separated in the cyclone. After the separation, the crude gas enters into the gas purification and cooling system for purification. The gas is split two parts, one is sent back the pyrolyzer to fluidize the bed material in the pyrolysis reactor. The other gas can be clean gas for civil or transformed and synthesized to produce relative chemical commodities after further purification, like desulfurization. Collected tar can be extracted for high value-added products or

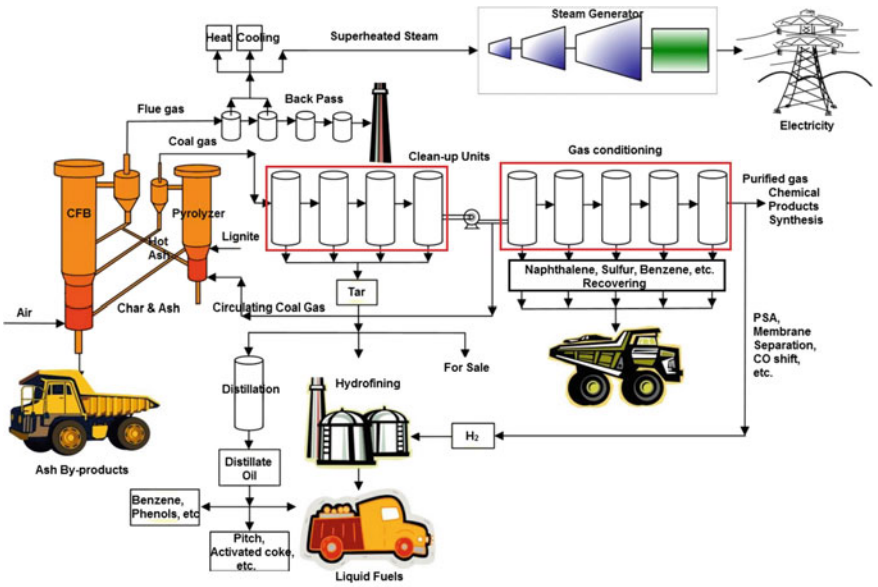

Fig. 6.3 Sketch of coal staged conversion process proposed by Zhejiang University

modified into high grade synthetic oil. Semi-coke is then sent to circulating fluidized bed boiler for combustion together with circulating material for power generation.

Zhejiang University has done a lot of research about the coal staged conversion process and has completed three stages about this technology. The first stage is 1 MW coal staged conversion process, which is developed and established in 1991. A lot of tests have been done on this equipment and theoretical and experimental study was carried out, which accumulated a great deal of theoretical knowledge and abundance operation experiences for the pilot scale and industrial scale. 12 MW and 40 t/h coal staged conversion process have been established and successfully operated in Huainan and Xiaolongtan, respectively.

The coal staged conversion process by Zhejiang University has the following advantages:

(1) The technology is simple and advanced: The circulating fluidized bed boiler and the pyrolysis and gasification furnace are combined closely. It can realize heat, electricity, tar and gas co-producing through the simple and advanced technology. High quality gas and tar can be produced in the process of steam power generation. The gas with high quality is good material of synthetic ammonia, methanol, synthesis of natural gas and other chemical products, it also can be used as fuel gas in the gas steam combined cycle power generation. The production of tar can be extracted for high value chemicals at the same time hydrogenate to produce liquid fuel. Thus, it effectively utilizes the various components of coal, and realizes comprehensive utilization of coal staged conversion.
(2) Wide fuel adaptability: The lignite, bituminous coal with volatile matter more than 20 % (ar) are suitable for this process. Also the coal particle size meets the requirements of existing circulating fluidized bed boiler. It avoids the disadvantage that existing coal gasification and distillation process which are restricted to coal types and coal particle size.
(3) Low requirement of technical parameters and low equipment investment: the coal pyrolysis and gasification under low temperature and oxygen free conditions at atmospheric pressure has low requirement for material of reactor and related equipments (conventional gasifier operating temperature is 1300–1700 °C, pressure 2–4 MPa). Equipment manufacturing cost is low. Also it does not consume oxygen and steam in the pyrolysis and gasification process, so that it does not need the oxygen preparation device and the steam boiler, and greatly reducing the cost of equipment construction in gasification system.
(4) Low operating cost: Coal pyrolysis unit doesn't need oxygen, steam as gasification agent. System energy loss is low. Compared with the conventional gasification technology, the process thermal efficiency is greatly increased. Thus, the operating cost is greatly reduced.

(5) Direct combustion of high temperature semi-coke: semi coke after pyrolysis and gasification is sent into boiler directly to combust for power generation. It avoids heat loss, making the energy fully utilized. While semi coke free water is fired in the boiler, which greatly reduces the amount of boiler flue gas, thus reduces the power consumption of fan. Energy consumption for the device is reduced. Boiler system efficiency is also improved. The process of cooling semi coke is avoided in the process of its production. Also the problem that fine semi coke particles which are difficult to transport and use is avoided.
(6) Easy to achieve large-scale: The fluidized bed pyrolysis furnace has the characteristics that hot ash and coal can be fully mixed, heat and mass transfer process is good, temperature field is uniform. They are all beneficial to the pyrolysis and gasification of coal in the furnace. Moreover, the fluidized bed pyrolysis furnace is easy to be large-scale. And it's easy to match with the circulating fluidized bed boiler in the arrangement, realizing of organic integration with circulating fluidized bed boiler. So the problem that the fixed bed or moving bed reactor is not easy to enlarge and arrange can be avoided.
(7) High gas production rate, good quality, achieving high value use of gas: Circulating ash is used as heat carrier to provide heat for the pyrolysis of circulating fluidized bed. Heat, electricity, tar and gas multi co-production can be realized. The content of effective component in the gas produced by pyrolysis is high. And the output of the gas is all used for subsequent. So as to ensure the gas content of the subsequent gas synthesis process, avoid the problem that burning of pyrolysis gas to provide heat source causes small amount of gas supply.
(8) Good control of pollutant emissions: Most of the sulfur content in coal is released as H_2S form in the pyrolysis process of the pyrolysis and gasification furnace, then entering into the gas purification system for desulfurization. While only a small amount of sulfur enters into the combustion furnace of circulating fluidized bed released in the form of SO_2. At the same time, compared with the flue gas desulfurization after direct combustion of coal, the removal of H_2S from coal gas has great advantages: (1) The amount of coal gas is greatly reduced. Therefore, the volume of desulfurization equipment, investment and operating costs are smaller. (2) At present, the by-product of gas desulfurization is the sulfur, which has great use of value. Most (80 % or more) of the nitrogen in coal is released in the form of nitrogen and ammonia in the pyrolysis process. And the combustion process in circulating fluidized bed is medium temperature combustion. Almost no thermal NOx is produced. Thus, the NOx emission concentration in the flue gas generated by the circulating fluidized bed combustion furnace is further reduced. Also extracting a small amount of ammonia from the small volume of gas is relatively easy and low-cost.

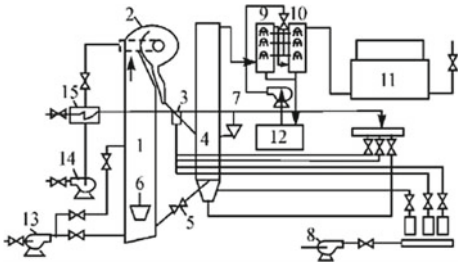

Fig. 6.4 Sketch of coal staged conversion process proposed by Tsinghua University. *1* Combustion furnace, *2* cyclone, *3* sending material device, *4* pyrolyzer-gasifier, *5* recycle material device, *6* the coal feeder for combustion furnace, *7* the coal feeder for gasifier, *8* fan, *9* the first stage cooler, *10* the second stage cooler, *11* gas tank, *12* cooling water, *13* blower, *14* induced draft fan, *15* superheater

The technology by Tsinghua University (Zhang et al. 2010) and Institute of Engineering Thermal Physics (Lv et al. 2008) are similar to that by Zhejiang University. The system sketch is shown in Figs. 6.4 and 6.5.

6.2.2 *The Coal Staged Conversion Process Based on Moving Bed Pyrolysis*

The basic principle of the coal staged conversion process based on moving bed pyrolysis is similar to that based on the fluidized bed pyrolysis. The difference is that the moving bed for the technology is used for coal pyrolysis based on the moving bed pyrolysis. This device has those characteristics: (1) Coal adaptability is wide. Bituminous coal and lignite with high content of volatile matter all can be used in this system and there is no special requirement about bondability and cokeability. (2) The slack coal that cannot be used for most other devices can be used for material for this device. However, the design and operation of the mixer for the coal and ash is the key problems for this technology.

The Institute of Coal Chemistry and Beijing Power Economic Research Institute have done a lot of research about this technology. Take Institute of Coal Chemistry for example, the coal staged conversion process based on moving bed pyrolysis is introduced in detail.

The coal staged conversion process proposed by Institute of Coal Chemistry is coupling moving bed pyrolysis and CFB combustion (Fig. 6.6) (Qu 2010; Liang et al. 2008; Qu et al. 2011). In this system, high temperate circulating ash is separated and collected in the cyclone of CFB boiler. Some ash is returned to CFB boiler and the rest ash is sent to the moving bed to provide heat for pyrolysis. The

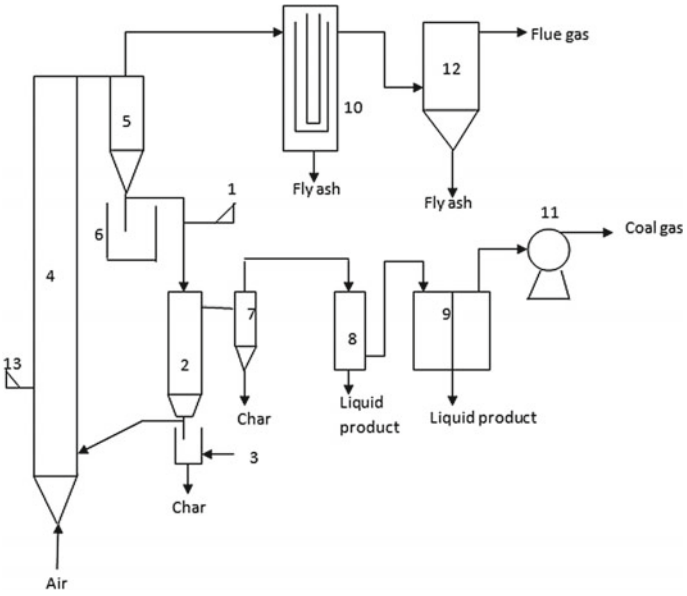

Fig. 6.5 Flow diagram of dual fluidized bed pyrolysis technology by Institute of Engineering Thermal Physics. *1* Screw feeder, *2* pyrolyzer, *3* startup recycle material, *4* boiler, *5* cyclone of the boiler, *6* U style recycle device, *7* cyclone of the pyrolyzer, *8* one stage cooler, *9* two stages cooler, *10* the cooler for the boiler, *11* induced draft fan, *12* bag-type dust remover, *13* screw feeder for startup

allocation of the high temperature ash is realized in the U-type valve. The coal is fed to pyrolyzer and mixed well with the high temperature ash. The coal undergoes pyrolysis, getting gas, tar and char. The fine ash is separated from the gas in the filter. The gas is then cooled and purified. The tar is collected in the oil tank. The purified gas can be used as domestic gas.

A lot of studies have been done on the key problems of this technology, such as hot ash allocation, mixing of hot ash and coal, and pyrolysis gas cleanup. Based on this, a pilot plant of a 75 t/h CFB boiler combined with a moving bed pyrolyzer is designed and established in Fugu, Shanxi, and the capacity of the pyrolyzer is 5 t/h coal. The operation and test results show that the process is successfully scaled up. The test has been completed in 2008. During the test, the poly-generation system can run stably and no operating problems were observed. The reasonable allocation of hot between CFB boiler and pyrolyzer could be realized by using the U-type valve and the conveying capacity could be adjusted by changing air flow rate. Hot ash and coal were mixed well in the gravitational mixer, which ensured their heat transfer. The agglomeration of char particles was not encountered in the

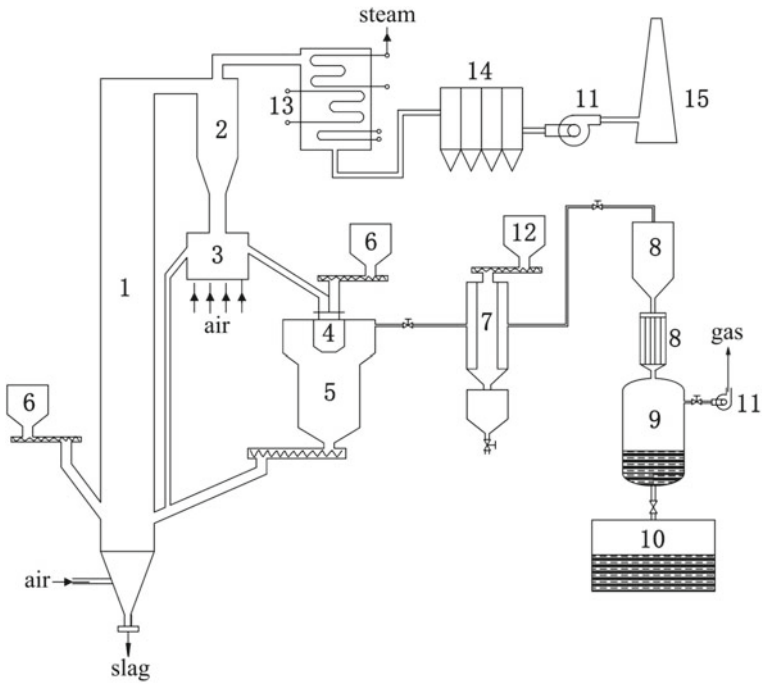

Fig. 6.6 Sketch of coal staged conversion process proposed by Institute of Coal Chemistry. *1* CFB boiler, *2* cyclone, *3* U-type valve, *4* mixer, *5* pyrolyzer, *6* screw feeder, *7* filter, *8* cooler, *9* tar separator, *10* oil tank, *11* fan, *12* feeder, *13* heat exchanger, *14* electrostatic precipitator, *15* stack

test. The solid inventory in pyrolyzer could be maintained by using the automatic controller. The purification of pyrolysis gas was carried out by coupling a granular bed filter with a two-stage indirect cooler to avoid the production of coking wastewater. Fine particles in tar could be restrained less than 4.0 wt% by using granular filter. The test results show that this system can run continuously and stable. The heat value can reach 26 MJ/Nm^3 and the tar yield reaches about 6 % wt at 600 °C.

The problems of this pilot experiment are as follows:

(1) The quality of the oil from pilot experiment is poor. Light fraction content of the oil is low, whereas asphalt content is high. This may be because the flow rate of the gas in the cooler is high and some light fraction is carried out of the cooler before condensation.
(2) The oil might be cooled and condensed in the filter for fly ash, which will affect the stable operation of pilot device.

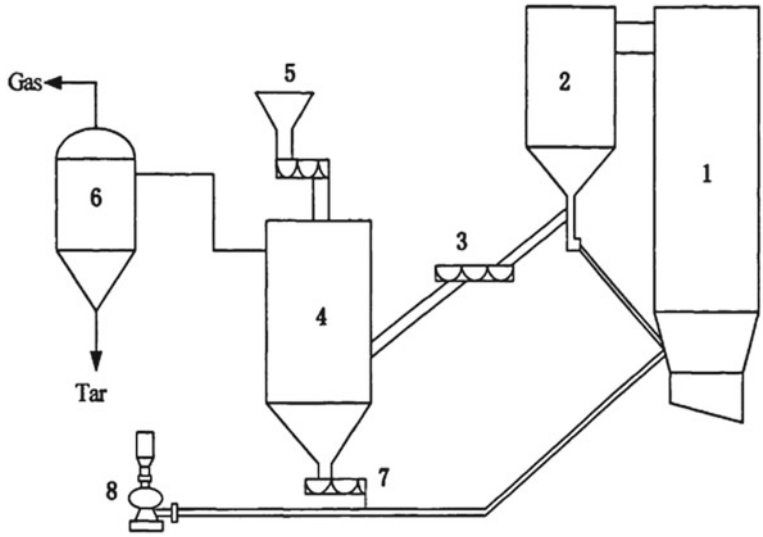

Fig. 6.7 Sketch of poly-generation system proposed by Beijing Power Economic Research Institute. *1* Boiler furnace, *2* cyclone, *3* ash screw feeder, *4* pyrolysis reactor, *5* coal hopper, *6* gas liquid separator, *7* semi-coke screw feeder, *8* root blower

The principle of the technology by Beijing Power Economic Research Institute is similar to that by Institute of Coal Chemistry, shown in Fig. 6.7.

6.2.3 The Coal Staged Conversion Process Based on Downer Pyrolysis

The main difference of coal staged conversion process with the other two technologies abovementioned is that the coal is pyrolyzed in the downer reactor. Institute of Process Engineering has done a lot of research about this technology.

Coal topping technology developed by Institute of Process Engineering is based on downer pyrolysis and CFB combustion (Zhang et al. 2014; Lin et al. 2007), shown in Fig. 6.8. Topping is a jargon in petrochemical processing industry, which meant distilling light oil from crude oil. The coal topping technology is to realize flash pyrolysis at relatively low temperature and extract combustible gas, liquid fuel and fine chemicals from coal without the addition of catalysts or H_2. It can realize poly-generation of oil, gas, heat and electricity in this system.

The principle of the coal topping technology is as follows: Coal is first sent to the solid-solid mixer and mixed with high temperature ash which is separated in the

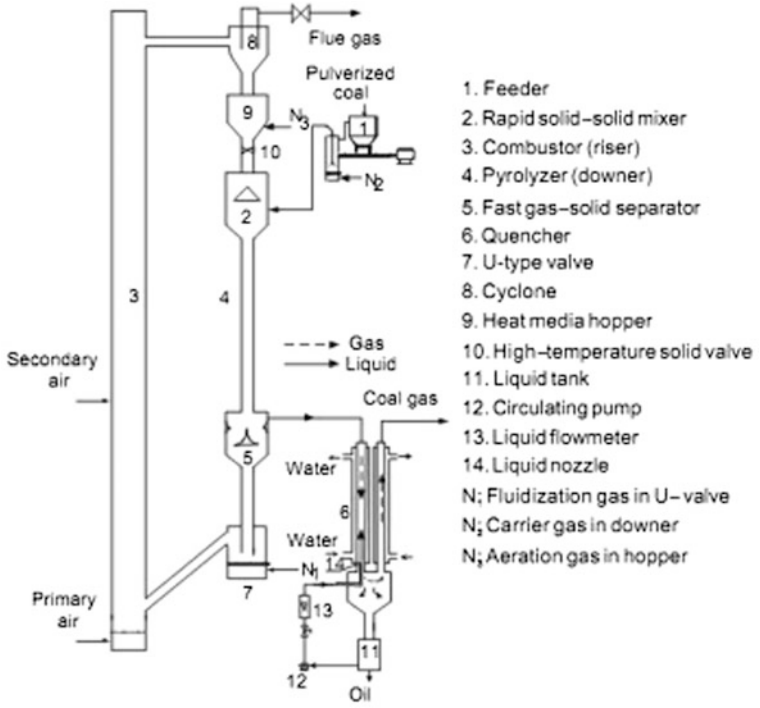

Fig. 6.8 System chart of coal staged conversion proposed by Institute of Process Engineering

cyclone of CFB boiler. The mixture of the coal and ash then goes to the pyrolyzer for pyrolysis, getting gas, tar and char. The mixture of char, ash, gas and tar goes to fast gas-solid separator due to the gravity and the purge of N_2. The raw coal gas is then separated with the char and ash in the gas-solid separator. The gas is fast cooled in the quencher and the tar is separated and collected in this process. The char and the ash are then sent to CFB boiler for combustion. The heated from combustion can be used to heat the steam and the ash. The ash then leaves the boiler from the top furnace with the flue gas and separated in the cyclone.

The coal staged conversion process based on downer pyrolysis with a capacity of 8 kg/h has been built. The effects of reaction temperature and coal particle meter on the yields of gas, tar and char and the composition of liquid sample have been studied on the hot test units. The gas and liquid yields are increased with the temperature and the liquid yield is decreased with the increase of coal particle meter. The light oil yield can reach 7.5 % and the cresol yield can reach 2–3 % at 660 °C with a coal particle meter less than 0.28 mm. Then this device with a capacity of 30 kg/h was built. In 2009, pilot plant of coal topping with a capacity of 10 t/d coal and tar hydrogenation device with a capacity of 700 kg/d tar was built in Langfang, Hebei.

6.2.4 The Other Technology Based on Internally Circulating Fluidized Bed

This system is made up of two fluidized beds. Coal is pyrolyzed in one fluidized bed and the char is fired in another fluidized bed. Huazhong University of Science and Technology, South Korea's Science and Technology college and Granfield in Britain have done a lot of research about this technology (Kim et al. 1997; Jaber and Probert 1999). Take Huazhong University of Science and Technology for example, this technology is introduced in detail.

System chart of the coal staged conversion process by Huazhong University of Science and Technology is shown in Fig. 6.9. The coal is sent to the fluidized bed pyrolyzer through a screw and undergoes pyrolysis, getting gas, tar and char. The char from the pyrolysis is sent to combustion furnace through overflow (which is made up of tube and L-valve). The char is fired in the boiler for combustion. The ash is sent to the pyrolyzer through next overflow.

This technology has the advantages as follows:

(1) This construct is simple and compact. The size of the device is greatly reduced compared with outer circulation.
(2) The heat loss is greatly reduced and thermal efficiency is increased.

The technologies proposed and developed by South Korea's Science and Technology College and Granfield in Britain are similar to that by Huazhong University of Science and Technology.

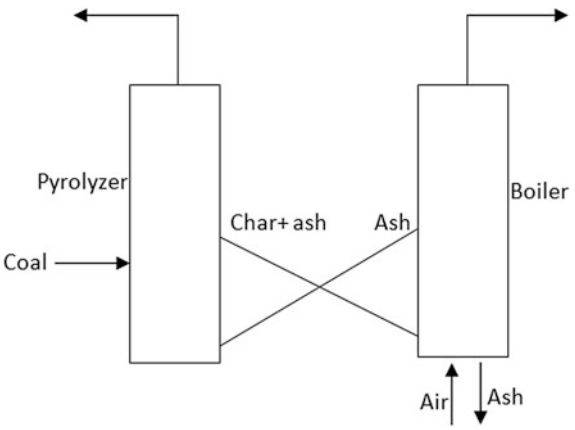

Fig. 6.9 System chart of the coal staged conversion process by Huazhong University of Science and Technology

6.3 Industry Application and System Analysis for the Coal Staged Conversion Process

Coal staged conversion process is rapidly developed in recent year, which has been in the stage of industrialization demonstration application. The technology by Zhejiang University is the most recognized and industrialized technology among all coal staged conversion technologies. 12 MW and 40 t/h facility have been built and successfully operated. 300 MW facility is under construction. So in this paper, take the technology by Zhejiang University for example, the brief introduction of industrial application based on coal staged conversion technology is given to show the progress in industrialization. In addition, to evaluate the feasibility of the large-scale coal staged conversion, 2 × 300 MW coal staged conversion system is analyzed.

6.3.1 The Industrial Application

6.3.1.1 12 MW Poly-generation Industrial Facility Based on Coal Staged Conversion Process

The scheme of 12 MW poly-generation industrial application is proposed based on a large number of experiments in 1 MW poly-generation facility. The device consisted of two parts: one is CFB boiler for power generation, the other part is fluidized bed pyrolyzer for coal pyrolysis.

The design fuel is Huainan bituminous coal and coal analysis is summarized in Table 6.1. It can be seen, from Table 6.1, that Huainan bituminous coal is an inferior coal with high ash and high fusion temperature, but the volatile matter and tar content are high. The design parameters of this facility are listed in Table 6.2.

The schematic flow of this device is shown in Fig. 6.10 (Wang et al. 2002; Fang et al. 2011; Guo 2015). The device is coupling a circulating fluidized bed boiler for combustion and a fluidized bed for pyrolysis. The CFB boiler is fluidized by air and the operating temperature range is 900–950 °C. The fluidized pyrolyzer is fluidized by circulating coal gas and the operating temperature range is 500–750 °C.

The coal which is crushed to 0–8 mm is fed to the pyrolyzer, and then mixed with high temperature circulating materials from CFB boiler. The coal incurs pyrolysis in the pyrolyzer at about 600 °C, getting gas, tar and char. The char is sent

Table 6.1 Coal analysis of Huannan bituminous coal

	Car (%)	Har (%)	Oar (%)	Nar (%)	Sar (%)	Aar (%)	Mt (%)	Var (%)	Qnet, ar (KJ/kg)	DT (°C)	Tar* (%)
Huainan	50.44	3.62	4.69	0.81	0.18	31.06	9.2	21.92	4637	>1400	8.2

*Obtained from Gray-King assay

Table 6.2 The design parameter of 12 MW poly-generation facility

	Poly-generation	Single CFB combustion
Main steam evaporation (t/h)	75	75
Main steam temperature (°C)	450	450
Main steam pressure (MPa)	3.82	3.82
Feed water temperature (°C)	150	150
Boiler pollutant discharge rate (%)	2.5	2.5
Outlet flue gas temperature of the boiler furnace (°C)	907	910
Pyrolysis temperature (°C)	600	–
Coal feed rate (t/h)	18.5	11.7
Particle size (mm)	0–8	0–8
Thermal efficiency of the boiler (%)	89	89.8
Primary air (Nm^3/h)	45341	44960
Secondary air (Nm^3/h)	29018	28774
Flue gas (Nm^3/h)	89440	92390
The flue gas temperature of air preheater outlet (°C)	135	137
Gas yield (Nm^3/h)	1700	–
Tar yield (t/h)	1.5	–

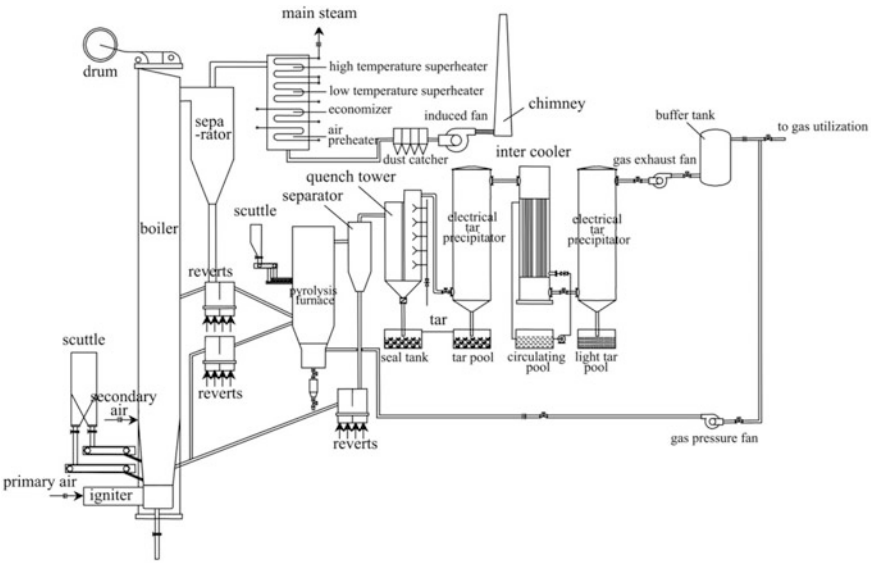

Fig. 6.10 12 MW coal staged conversion process

to the boiler for combustion through recycle device along with circulating materials. A great deal of high temperature material together with high temperature flue gas enters high temperature cyclones through furnaces outlet. The high temperature flue gas is separated from high temperature materials in the cyclones and then introduced to back pass heat exchangers, i.e., superheater, reheater, economizer and air preheater. The high temperature circulating materials enters recycle device through a vertical pipe. Part circulating materials are sent to the pyrolyzer by controlling the high temperature valve and the rest high temperature materials are sent to the CFB boiler.

Raw coal gas and fine particles in the pyrolyzer enters the cyclones of the pyrolyzer through pyrolyzer outlet. The fine particles (char and ash) are sent to the boiler for combustion through recycle device. The raw coal gas is cooled and purified when it passes through quench tower, electrical tar precipitator, intercooler, filter and gas exhauster. Part gas as fluidizing medium for the pyrolysis is sent to the bottom of the pyrolyzer after pressurization through fan. The rest gas is further purified to clean gas after desulfurization. The tar can be used for other purposes.

The connection between the boiler and pyrolyzer is realized through non-mechanical valves, including the materials recycled from the vertical tube of CFB cyclone to the pyrolyzer, from the pyrolyzer to the boiler and from cyclone of the pyrolyzer to the boiler. This three recycle devices are fluidized by the steam, so it can avoid blow-by of the air, flue gas and coal gas, thereby solving the problems of local slagging, the lower yields and heat values for the coal gas and oil due to blow-by. The recycle device is fluidized by the air when the pyrolyzer does not run.

The poly-generation facility can run in two models: (1) The poly-generation system is in normal operation, the boiler and pyrolyzer are both put into operation. In this case, the coal is fed to the pyrolyzer and the recycle devices are fluidized by the steam. (2) Only CFB boiler is in normal operation and the pyrolyzer does not run. In this case, the coal is fed to boiler and the recycle device is fluidized by the air. So this device can generate electricity normally when the pyrolyzer is in maintenance. In both two models, rated electricity generation can be obtained and the only difference is the operation temperature of the boiler is different.

The typical results of 12 MW facility is shown in Tables 6.3 and 6.4. The gas yield and oil yield reach 1250 Nm^3/h and 1.1 t/h respectively. H_2 and CH_4 are the main components in the coal gas. The operation results of the 12 MW CFB coal

Table 6.3 The typical result of 12 MW facility

	Boiler	Pyrolyzer
Temperature (°C)	925	600
Coal feed rate (t/h)	3	11
Gas yield (NM^3/h)	–	1250
Tar yield (t/h)	–	1.1
Tar yield (%)	–	10 %
Steam capacity (t/h)	70	–

Table 6.4 The composition of coal gas for 12 MW facility

H_2 (%)	O_2 (%)	N_2 (%)	CH_4 (%)	CO (%)	CO_2 (%)	C_2H_4 (%)	C_2H_6 (%)	H_2S (mg/Nm^3)	Heat value (MJ/m^3)
27.30	0.11	6.56	36.9	7.1	10.1	3.54	6.9	8000	23

staged conversion poly-generation plant showed that this system is stable, the adjustment is convenient and the operation is safe and reliable. It realized the production of a variety of high value products with coal as the resource in an organic integrated system. It mainly manifested in those aspects as follows:

(1) The circulating fluidized bed combustion furnace and the fluidized bed pyrolyzer can be coordinately operated. The material circulation between the pyrolyzer and the furnace is stable and reliable, and the operation is convenient.
(2) The semi coke from the pyrolysis furnace could be fired stably in the circulating fluidized bed boiler. At the same time, the load of the pyrolyzer and the furnace can be adjusted respectively according to the operation requirements.
(3) The system can stabilize the production of tar and gas. Gas purification and tar recovery system operates normally.
(4) The automatic operation control of the multi generation system is realized, and the operation is convenient.
(5) The efficiency of the electric precipitator for dust removal at high temperatures is high, and this ensured the low dust content of tar.

6.3.1.2 40 t/h Poly-generation Industrial Facility Based on Coal Staged Conversion Process

In 2009, Zhejiang University cooperating with Xiaolongtan Mine Bureau and State Power Xiaolongtan Power Plant reformed the existing 300 MWe lignite circulating fluidized bed boiler of Xiaolongtan Power Plant to a 300 MWe circulating fluidized bed coal staged conversion poly-generation device. This is designed to be constructed in two phases. At present, the first step, the construction of 40 t/h coal staged conversion process has been completed. The second step is that 300 MW coal staged conversion process is under construction. The reformed poly-generation facility consists of a 300 MW CFB boiler and a fluidized bed pyrolyzer with a capacity of 40 t/h coal. The schematic flow of this facility is shown in Fig. 6.11 and the picture of this reformed device is shown in Fig. 6.12.

The Xiaolongtan lignite is fed to the pyrolyzer, and then mixed with high temperature circulating materials from CFB boiler. The coal undergoes pyrolysis in the pyrolyzer at about 700 °C, getting gas, tar and char. The char is sent to the boiler through recycle device along with circulating materials. The char is burnt in the boiler to heat steam, circulating materials and flue gas. A great deal of high

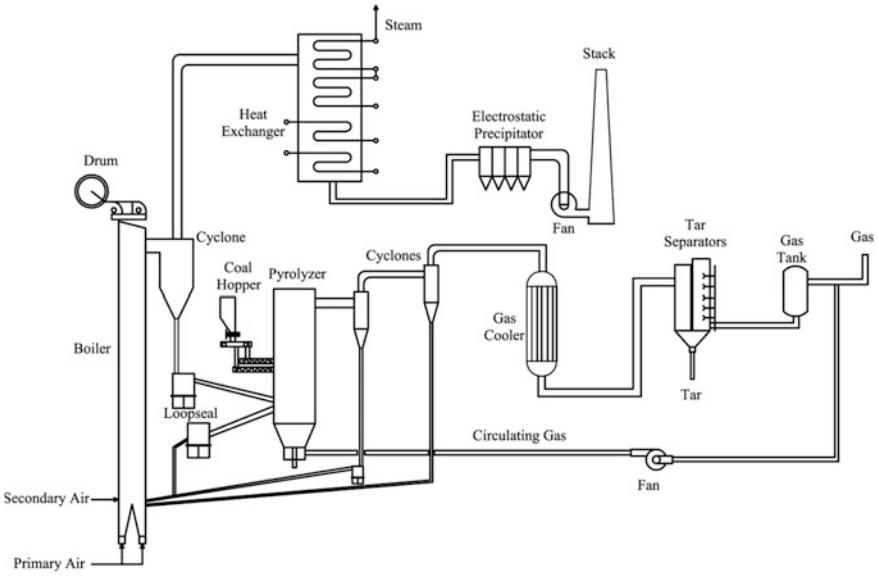

Fig. 6.11 40 t/h coal staged conversion process

Fig. 6.12 The device of 40 t/h coal staged conversion process

temperature material together with high temperature flue gas enters high temperature cyclones through furnaces outlet. The high temperature flue gas is separated from high temperature materials in the cyclones and then introduced to back pass

heat exchangers, i.e., superheater, reheater, economizer and air preheater. The high temperature circulating materials enters recycle device through a vertical pipe. Part circulating materials are sent to the pyrolyzer by controlling the high temperature valve and the rest high temperature materials are sent to the boiler.

Raw coal gas and fine particles from the pyrolysis enters the one-stage cyclones of the pyrolyzer for separation. The raw coal gas after separation enters two-stage cyclone for further separation from fine particles. The fine particles (char and ash) from one-stage cyclone are sent to the boiler for combustion through recycle device. The fine particles from two-stage cyclone are sent to the boiler for combustion through pneumatic transportation.

The raw coal gas passes through quench scrubber to remove dust. The gas with a temperature about 600 °C is cooled to 100 °C after mixing with circulating washing water at 80 °C from circulating water pump and the dust from the coal gas is scrubbed and removed in this process. Then the coal gas is further cooled to 45 °C after mixing with circulating washing water at 30 °C from circulating water pump and the tar is scrubbed and separated from the coal gas in this process. The coal gas is then pressurized to 10 kPa through pressurization fan. The gas is split two parts, one as fluidized medium is sent to pyrolyzer and the other is fired through torches.

The typical results are shown in Table 6.5. Operation and test results showed that the system is stable and easy to operate, and the amount of coal feed for the pyrolyzer is 40 t/h lignite without drying. The gas production is 8200 Nm^3/h (the effective components such as H_2, CO and CH_4 were more than 75 %). Tar yield is 1.1 t/h. The parameters reach or exceed the design requirements.

This facility also can run in two models: (1) The poly-generation system is in normal operation, the boiler and pyrolyzer are both put into operation. In this case, the coal is fed to the pyrolyzer and the recycle devices are fluidized by the steam. (2) Only 300 MW CFB boiler is in normal operation and the pyrolyzer does not run. In this case, the coal is fed to boiler and the recycle device is fluidized by the air. In both two models, rated electricity generation can be obtained and the only difference is the operation temperature of the boiler is different.

On the basis of the existing work, the coal staged conversion process has the ability for industrial application. Zhejiang University is now working with the Dongfang Boiler Co., Ltd. to carry out the design the device of 350 MWe super-critical circulating fluidized bed and scheme design of 600 MWe device, laying the foundation for the next large-scale industrial application.

Table 6.5 The typical result of 40 t/h coal staged conversion process

	Designed parameters	Operation parameters
Coal feed rate for the pyrolyzer (t/h)	40	40
Coal gas yield (Nm^3/h) >	7800	8200
(H_2 + CO + CH_4) % >	65	74
Tar yield (t/h)	1	1.1

6.3.2 System Analysis

To evaluate the feasibility of a large-scale poly-generation plant, a new 2 × 300 MW coal staged conversion process system is built, which couples 2 × 300 MW CFB boiler, two fluidized pyrolyzer, steam turbine, the cooling and purification of coal gas, methane reforming, methanol synthesis, tar hydrogenation and IGCC. The flow diagram of this system is shown in Fig. 6.13. Coal is fed to the pyrolyzer for pyrolysis at 700 °C by the interaction with circulating ash from CFB boiler. The char is sent to the boiler for combustion at 900 °C to generate electricity. Raw coal gas is cooled during flowing through heat exchanger unit. In this process, most tar is collected and some H_2S, NH_3 are also removed from the coal gas. Tar is upgraded for liquid oil through two-stage hydrogenation. Some gas is returned to the pyrolyzer as fluidized medium. H_2S of the other gas is further removed using Selexol/Claus technology. Then hydrogen-rich gas is obtained in the PSA unit and this part is used for hydrogenation. The other gas is synthesized to

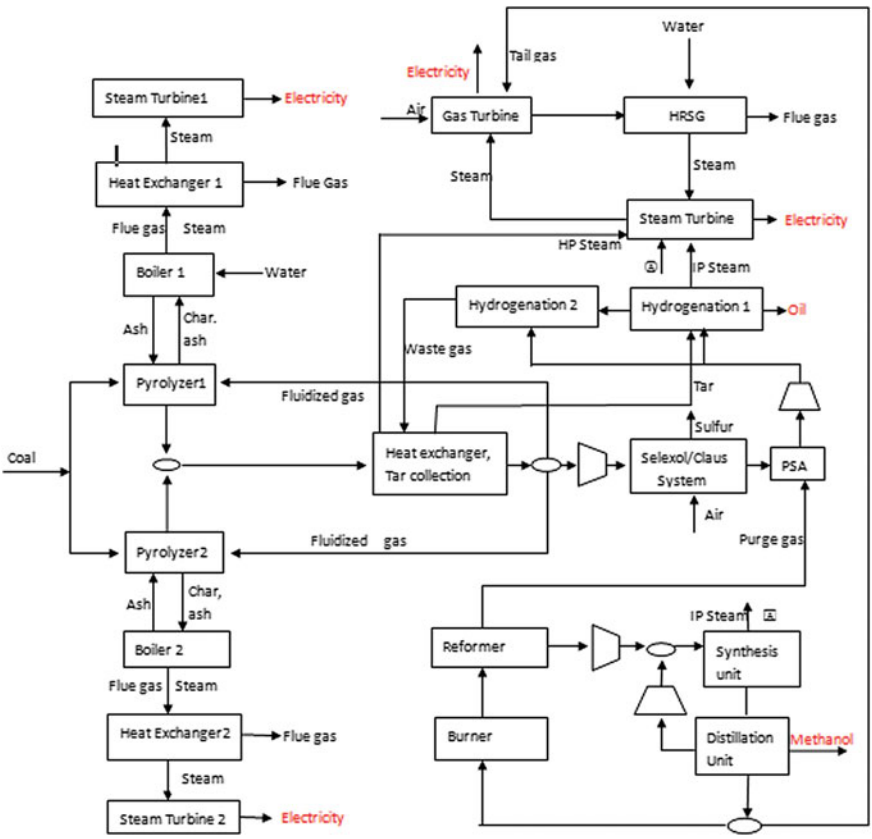

Fig. 6.13 Flow diagram of coal staged conversion process

methanol after methane reforming. The un-recycling tail gas of the systhesis tower is sent to gas turbine and heat recovery boiler for electric generation. Therefore, this system can realize poly-generation of electricity, methanol and liquid fuel.

Steady state models of poly-generation and traditional CFB plant are established respectively by using Aspen Plus software. The details of unit modeling parameters and method are presented in Guo's study (Guo et al. 2014). The thermodynamic and economic parameters are compared between poly-generation and traditional CFB plant.

The thermodynamic characteristics of the optimum design and the economic analysis of poly-generation plant are compared with those of conventional 300 MW CFB power plant are summarized in Tables 6.6 and 6.7 respectively.

To ensure the same electricity for the coal staged conversion plant and traditional CFB plant. The coal fed to the staged conversion process is 239.50 kg/s, while this value is 142.16 kg/s for traditional plant. The increase of coal in the poly-generation plant is converted into extra power (172.16 MW), methanol (16.86 kg/s), and oil (2.95 kg/s). The energy efficiency (43.2 %) and exergy efficiency (43.24) is far high than that for traditional CFB plant (34.9 and 34.2 %). This showed the coal staged conversion process system is a high efficiency technology.

According to the economic results, on one hand, higher total capital cost for polygeneration plant are needed compared with traditional CFB plant, which is mainly ascribed to the costs of building extra facilities, like pyrolyzers, gas clean-up and tar utilization facilities. But the extra facilities account for only 36 % of the total fixed capital cost of the polygeneration plant. But on the other hand, the profit of poly-generation is four times than that of traditional CFB plant, which is mainly due to the profit of fuel oil and methanol. The internal rate of return (IRR) for coal staged conversion process is also higher than that for the traditional CFB plant. The price fluctuation can greatly influence the profit of the coal staged conversion process. The increase of the price of coal, oil and methanol is in favor of the coal staged conversion process, but the increase of electricity price is not in favor of the coal staged conversion process. The coal staged conversion process has a good economic feature within a wide range of price fluctuations. Overall, at the market condition in 2011, building a polygeneration plant is much attractive to the government and investors.

The thermodynamic and economic analysis of coal staged conversion process coupling a 2 × 300 MW CFB power plant and 2 atmospheric pressure fluidized bed pyrolyzers are summarized below.

In thermodynamic aspects, the energy and exergy efficiency for coal staged conversion process is 43.2 % and 43.24 % respectively, which is far higher than that for traditional CFB plant (34.9 and 34.2 %). In economic aspect, Polygeneration plant, which has larger IRR (24.07 %) and shorter DPP (6.36 years) and SPP (5.29 years), shows more profitable than CFB power plant at the prices in 2011. The price of the coal, oil, electricity and methanol can greatly affect the poly-generation profit. The increase of the price of coal, oil and methanol is in favor of the coal staged conversion process, but the increase of electricity price is not in

Table 6.6 Thermodynamic performance of CFB power plant and poly-generation plant

Items	CFB power plant	Poly-generation plant
Energy input		
Coal input (kg/s)	143.16	239.50
Coal energy input(MW)	1739.39	2909.93
Coal external input (MW)	1776.66	2972.12
Energy output		
Oil yields (kg/s)	0.00	2.95
Methanol yields (kg/s)	0.00	16.86
Oil energy output (MW)	0.00	127.76
Oil external output (MW)	0.00	127.70
Methanol external output (MW)	0.00	377.56
Methanol energy output (MW)	0.00	349.50
Power generation in steam turbine (MW)	0.00	180.93
Power generation in gas turbine (MW)	0.00	66.16
Power generation in boiler side (MW)	649.73	649.73
Auxiliary power consumption		
Grinding (MW)	0.04	0.07
GTCC steam cycle pump (MW)	0.00	1.31
Tar pump (KW)	0.00	65.54
Primary air fan (MW)	14.13	17.54
Introduced draft fan (MW)	11.78	10.12
Circulating gas fan (MW)	0.00	1.78
Secondary air fan (MW)	0.46	0.56
Recycling tail gas compressor (MW)	0.00	17.54
Boiler side steam cycle pump (MW)	15.59	15.58
Gas compressor after reformer (MW)	0.00	17.8
Hydrogen compressor (MW)	0.00	2.26
Gas compressor before Selexol (MW)	0.00	31.72
Glaus compressor (MW)	0.00	1.01
Total consumption (MW)	41.93	117.35
Summary		
Net power output (MW)	607.80	779.96
Total energy output (MW)	607.80	1257.22
Total external output (MW)	607.80	1294.36
System energy efficiency (%)	34.9	43.20
System external efficiency (%)	34.2	43.24

favor of the coal staged conversion process. Overall, the coal staged conversion process has a good economic feature within a wide range of price fluctuations.

According to the system analysis, the coal staged conversion process has obvious advantages, thus enjoying a broad prospect of market application.

Table 6.7 Comparison of economic results

Items	CFB power plant	Poly-generation plant
Fuel and material consumption (million tone/year)		
Coal	3.71	6.21
Limestone	0.22	0.66
Water	10.92	15.68
Product (tone/year)		
Electricity (KWh/year)	4.38×10^9	5.62×10^9
Crude oil	0.00	76,412.16
Methanol (99.85 % purity)	0.00	437,025.16
Sulfur	0.00	41,741.36
Capital cost ($10^6$$)		
Coal power plant (including boiler, steam turbine, fans, grinding and other facilities	542.63	541.42
Fluidized bed pyrolyzer and gas cleanup	–	63.89
Gas turbine	–	28.06
Heat exchanger	–	27.78
H_2 PSA	–	12.63
Selexol	–	43.49
Steam turbine and cycle	–	40.16
Glaus/SCOT	–	29.64
Methanol reactor	–	7.18
Methanol reforming	–	19.28
HRSC	–	15.49
Product upgrading	–	3.66
Tar hydrogenation	–	2.36
Pumps	–	0.07
Coal gas compressor	–	21.85
Fixed capital cost	542.63	856.97
Interest during installation	53.18	83.98
Total capital cost	595.81	940.95
O&M ($10^6$$/year)	21.71	34.28
Summary		
Material and fuel cost (10^6 $/year)	147.48	245.44
Annual output value (10^6 $/year)	243.88	565.74
Annual profit (10^6 $/year)	74.69	286.02
IRR (%)	10.63	24.07
SPP (years)	9.98	5.29
DPP(years)	14.49	6.36

6.4 Summary

The coal staged conversion process technology coupling circulating fluidized bed combustion and coal pyrolysis can make effective use of hydrogen-rich volatile matter and can solve the environment pollution in an economic way, thus realizing high-value utilization for coal. The principle is: the coal incurs pyrolysis in the pyrolyzer obtaining gas, tar and the char. The heat for the pyrolysis is provided by the circulating materials. The char in the pyrolyzer is forward to the boiler for combustion. Poly-generation of gas, tar, heat and power can be realized in this system.

According to the bed type of the pyrolyzer, the coal staged conversion technology is mainly divided into three types: the technology based on fluidized bed pyrolysis, moving bed pyrolysis and downer bed pyrolysis. Those three types of technologies are introduced in detail.

Take the technology by Zhejiang University for example, industrial application and system analysis are discussed in detail. 12 MW and 40 t/h coal staged conversion process is established and successfully operated. The test results show a satisfactory process performance and co-produce heat, electricity, gas and tar. When the pyrolyzer does not run, the circulating fluidized bed can still be normally operated. This shows that the coal staged conversion technology can be applied to industrial scale poly-generation plant. To evaluate the feasibility of the coal staged conversion technology in a large scale, thermodynamic and economic of poly-generation system coupling 2 × 300 MW circulating fluidized bed and 2 fluidized bed pyrolyzers is analyzed. The result shows that the coal staged conversion process system has higher energy and exergy efficiency and more profitable than single traditional 2 × 300 MW CFB plant. This justified the coal staged conversion process in a large scale is an efficient and economic technology.

References

Fang, M. X., Luo, Z. Y., Li, X. T., Wang, Q. H., Ni, M. J., & Cen, K. F. (1998). A multi-product cogeneration system using combined coal gasification and combustion. *Energy, 23*, 203–212.

Fang, M. X., Zeng, W. Q., Cen, J. M., Wang, Q. H., & Luo, Z. Y. (2011). Development and application of CFB-based polygeneration technique of coal grading conversion. *Guangdong Electric Power, 24*, 1–7.

Guo, Z., Wang, Q., Fang, M., Luo, Z., & Cen, K. (2014). Thermodynamic and economic analysis of polygeneration system integrating atmospheric pressure coal pyrolysis technology with circulating fluidized bed power plant. *Applied Energy, 113*, 1301–1314.

Guo, Z. H. (2015). Research on key issues of lignite pyrolysis-based staged conversion polygeneration technology. Doctor, Zhejiang University.

Jaber, J. O., & Probert, S. D. (1999). Environmental-impact assessment for the proposed oil-shale integrated tri-generation plant. *Applied Energy, 62*, 169–209.

Kim, Y. J., Lee, J. M., & Kim, S. D. (1997). Coal gasification characteristics in an internally circulating fluidized bed with draught tube. *Fuel, 76*, 1067–1073.

Liang, P., Wang, Z., & Bi, J. (2008). Simulation of coal pyrolysis by solid heat carrier in a moving-bed pyrolyzer. *Fuel, 87*, 435–442.

Lin, X. B., Zhao, W. B., Wu, S. H., & Qin, M. (2007). Mixing behavior between coal and hot circulating ash in down flow of coal topping process. *Coal Conversion, 30*, 36–40.

Lv, Q. G., Liu, Q., Fan, X. X., Song, G. L., Na, Y. J., & He, J. (2008). Experimental Study on Gasification in Dual Fluidized Beds. *Journal of Engineering Thermophysics, 29*, 1435–1439.

Qu, X. (2010). *Fundamental research and pilot development of the poly–generation process of circulating fluidized combustion combined with coal pyrolysis.* Doctor, Institute of Coal Chemistry.

Qu, X. A., Liang, P., Wang, Z. F., Zhang, R., Sun, D. K., Gong, X. K., et al. (2011). Pilot development of a polygeneration process of circulating fluidized bed combustion combined with coal pyrolysis. *Chemical Engineering and Technology, 34*, 61–68.

Wang, Q. H., Luo, Z. Y., Fang, M. X., Ni, M. J., & Cen, K. F. (2002). Development of a 12 MW multi-generation of gas, steam and power. *Fuel Chemistry and Technology, 30*, 141–146.

Wang, Q. H., Luo, Z. Y., Li, X. T., Fang, M. X., Ni, M. J., & Cen, K. F. (1999). A mathematical model for a circulating fluidized bed (CFB) boiler. *Energy, 24*, 633–653.

Zhang, Y., Zhao, Y. H., & He, X. F. (2014). The pyrolysis process of coal topping simulation using Aspen Plus. *Computers and Applied Chemistry, 31*, 774–778.

Zhang, Z. F., Ren, J., Li, Z. H., & Chen, G. (2010). General description for coal pyrolysis poly-generation technology. *Chemical Fertilizer Design, 48*(11–15), 21.

Chapter 7
Energy Polygeneration Systems and CO_2 Recycle

Qun Yi, Yan-Hong Hao, Ji-Long Zhang and Wen-Ying Li

Abstract Polygeneration, typically involving co-production of chemicals or liquid fuels and electricity, is an advanced and promising energy conversion technology which provides opportunities for high energy utilization efficiency, high economic benefits and low/zero emissions. This chapter firstly gives a brief introduction of characteristics, development and patterns on energy polygeneration systems. Basis and theories of the integration and optimization of energy polygeneration system are then introduced. Polygeneration is a highly flexible and integrated system which couples with different technologies and energy fuels feedstocks, and it is pointed out that the integration and optimization of an energy polygeneration system should comply with the principle of cascade utilization of energy and classify conversion of element to obtain the optimal comprehensive performance in terms of energy, economy and environment. Finally, the feasibility and development prospect of CO_2 conversion and use in polygeneration system are discussed. The everlasting pursuit in energy conversion and utilization is improving the utilization and efficiency of resources, whilst reducing CO_2 emissions. Polygeneration has inherent advantages of CO_2 recycle and reduction, and CO_2 recycle and use coupled with

Q. Yi (✉) · W.-Y. Li
Key Laboratory of Coal Science and Technology (Taiyuan University of Technology), Ministry of Education and Shanxi Province, Taiyuan 030024, People's Republic of China
e-mail: yiqun@tyut.edu.cn

Q. Yi · W.-Y. Li
Training Base of State Key Laboratory of Coal Science and Technology Jointly Constructed by Shanxi Province and Ministry of Science and Technology, Taiyuan 030024, People's Republic of China

Y.-H. Hao
Environmental Engineering Department, Shanxi University, Taiyuan 030013 People's Republic of China

J.-L. Zhang
Institute of Mineral Resources Utilization, College of Mining Technology, Taiyuan University of Technology, Taiyuan 030024, People's Republic of China

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_7

different energy polygeneration systems are analyzed. All of those present excellent sustainability in their energy use and environmental protection. Polygeneration with CO_2 recycle has a huge potential in CO_2 reduction, while shows low cost involved with CO_2 reduction.

Nomenclature

Capital Letters

C1	CO_2 emission caused by energy consumption
CALCOR	Reforming process combined with CO_2 recovery and CO purification
CANMET	Canada Centre for Mineral and Energy Technology
CAPEX	Capital expense
CCR	CO_2 capture and recycle
CCS	CO_2 capture and storage
C_{coal}	Carbon cost in coal
C_e	The total energy cost
C_{ne}	The total non-energy cost
C-tax	CO_2 tax
DMC	Dimethyl carbonate
DME	Dimethyl ether
D-PL	Dual gas polygeneration system
D-PL-CR	Dual gas polygeneration system with CO_2 recycle
DOE	Department of Energy
ERC	Electrochemical reduction of CO_2
FT	Fischer-Tropsch
GHG	Greenhouse gas
Gt C/y	Gigaton carbon per year
GTL	GAS-TO-LIQUID
GTR	Gas-to-chemical resources
GW	Gigawatt
GWh	Gigawatt hour
HZSM-5	Hydrogen zeolite Socony mobile-five
IGCC-CCS	Integrated gasification combined cycle with CO_2 capture and storage
IGCC-CRS	Integrated gasification combined cycle with CO_2 recycle and storage
IPCC	International Panel on Climate Change
IRR	Internal rate of return
JOGMEC	Japan Oil, Gas and Metals National Corporation
KOGAS	Korea gas corporation
MeOH	Methanol
Mt/y	Megaton per year
MWth	Megawatt thermal
NASA	National Aeronautics and Space Administration Technical Memorandum
OPEX	Operating expense
PEC	Photoelectrocatalytic

PIS	Capital cost saving ratio
PL-CCS	Polygeneration system with CO_2 capture and storage
PL-CRS	Polygeneration system with CO_2 recycle and storage
PV	Photovoltaic
RCO_2	Renewable CO_2
RD&D	Research, development and demonstration
SPARG	Sulfur passivated reforming
SPC	Solar sulphur–iodine cycle
t-CO_2/GWh	Ton-carbon dioxide per gigawatt hour
UV	Ultraviolet

Lowercase Letters

bbl	Barrel
k	Distribution coefficient of carbon in products

Greek Letters

λ	The mass flow of CO_2 recycled to the reforming unit over the mass flow of all separated CO_2 from distillation unit
€/t	Euro per ton

Acronyms

BFG	Blast-furnace gas
BGL	British gas Lurgi
CAS	Chinese Academy of Sciences
CDCL	Coal direct chemical looping
CG	Converter gas
CLC	Chemical looping combustion
CLP	Calcium looping process
CLR	Chemical-looping reforming
COE	Cost of energy
COG	Coke oven gas
CTL	Coal to liquid
DRI	Desert Research Institute
EAGLE	Energy Application for Gas, Liquid & Electricity
ESR	Energy saving ratio
FT	Fischer-Tropsch synthesis
GT	Gas turbine
HRSG	Heat recovery steam generator
IFRF	International Flame Research Foundation
IGCC	Integrated gasification combined cycle
IHI	Ishikawajima-Harima Heavy Industries
LCA	Life cycle assessment
LPG	Liquefied petroleum gas

METI	Ministry of Economy, Trade and Industry
MFEPS	Multi-feedstock energy polygeneration system
MP	Methanol plant
NG	Nature gas
RWGS	Reverse water-gas shift
SCL	Syngas chemical looping
ST	Steam turbine
THU	Tsinghua University
TYUT	Taiyuan University of Technology
UNG	Unconventional nature gas
WGS	Water gas-shift

7.1 Polygeneration

7.1.1 Definition and Characteristics

The term 'polygeneration' is reported first appearing in the open engineering literature in 1982 through studies at National Aeronautics and Space Administration (NASA) and General Electric (Burns et al. 1982). 'Polygeneration' has had many meanings, however, the chemical engineering community has defined it a thermochemical process which simultaneously produces at least two different products in non-trivial quantities, but is not a petroleum refining process, a co-generation process, or a tri-generation process, and at least one product is a chemical or fuel, and at least one is electricity (Adams and Ghouse 2015). The syngas is almost the primary intermediate in these polygeneration processes while the feeds, products, technologies, and pathways have various options. Nowadays, polygeneration has developed to be a multi-input and multi-output energy system that coproduce electricity and synthetic liquid fuels or chemicals, which as one of the most promising technologies is intended to be a highly flexible and optimal integration of resource/energy/environment systems meant to benefit all of society, while also breaking down the existing separation of different industrial sectors (Serra et al. 2009).

A conceptual diagram of polygeneration proposed by The United States Department of Energy (DOE) is shown in Fig. 7.1. It can be seen that the polygeneration system is a complex system that comprises many units and pieces of equipment, however, the synergism of different processes realize the cascade utilization of energy and step by step conversion of components. Different types of feedstock, including coal, biomass (e.g., agricultural residue, sugarcane bagasse), natural gas, municipal solid waste, petroleum coke, black liquor from paper and pulp mills etc. are being used for gasification on an industrial scale. All these have the benefit of enhancing energy security by reducing the dependence on imported

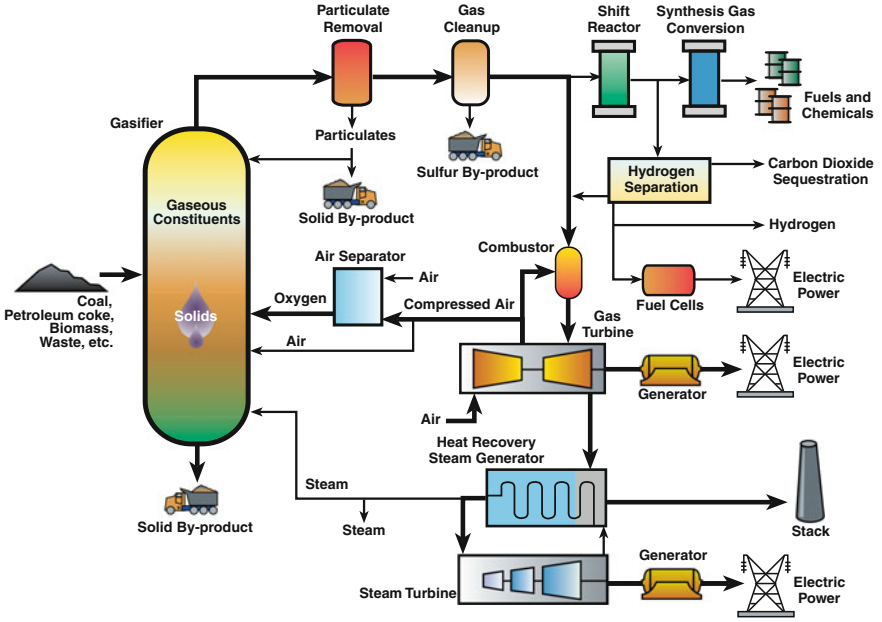

Fig. 7.1 Conceptual diagram of polygeneration (Breault 2010)

oil. What's more, there is product flexibility since the product ratio can be changed according to market fluctuations. The energy conversion efficiency of a polygeneration plant is typically higher than that of conventional stand-alone plants which produce the same products, as a result of more efficient energy utilization (according to energy quality) and higher degree of process integration (Liu et al. 2010a). Further, polygeneration also provides opportunities to realize a low/(almost) zero emissions system by integrating with CO_2 capture and sequestration (CCS) technologies (Liu et al. 2010a). Besides, polygeneration shares public facilities thus reduces the capital investment and operation costs due to the high degree of integration between power generation and chemical synthesis, and production cost also can be reduced by co-production of high value-added chemicals and electricity. In a word, polygeneration is a system taking advantages of higher element utilization, more effective energy conversion, more emission reduction, and profitable economic benefit over conventional systems.

As the concept of polygeneration is first proposed, polygeneration system has experienced a rapid development. The typical concepts are "Vision 21" and "syngas park", which were put forward by DOE and Shell company, respectively. The former plant was aimed to realize coal utilization with near-zero emission (Ni et al. 2000). The system takes the advantages of high efficiency of energy utilization efficiency reaching 50–60 %, generating efficiency over 60 %, and coproduction H_2 efficiency approaching 75 %, near-zero pollutants emission with CO_2 emission reduction by 40–50 %, and low cost of 10 % more economical benefit than modern

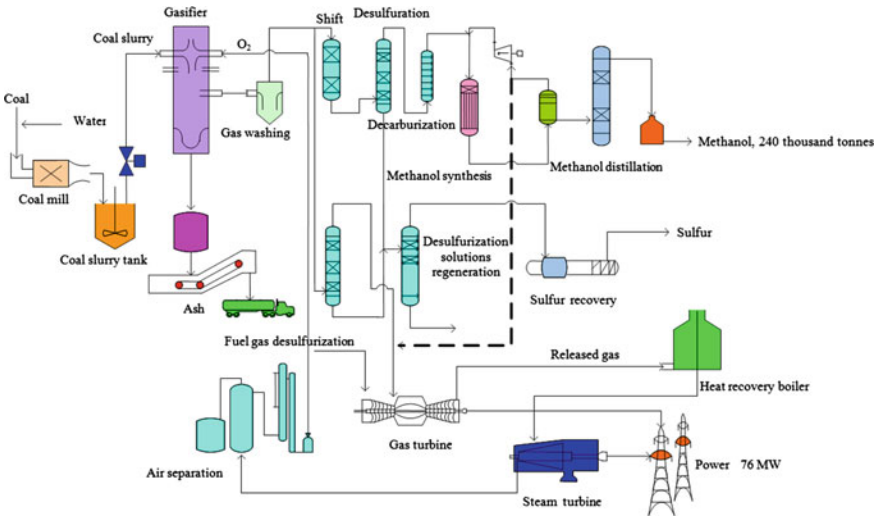

Fig. 7.2 Polygeneration demonstration project of Yankuang Group Co. Ltd

pulverized coal gasifier power plant (Wang 2001). In the Shell park, coal or residual oil is gasified. The syngas produced can be used to produce electricity by IGCC, and methanol and fertilizer in a "once-through" process, and can also be used as town gas. The first polygeneration demonstration plant in the world with 60 MW capacity and 0.24 million tons methanol per year was successfully established in China (Xiao 2008). The process flow diagram is shown as Fig. 7.2. Between April 2006 to August 2008, the polygeneration plant produced methanol and power of 8.26×10^5 t and 9.2×10^8 kW h, respectively. System availability reached to 93 % and system reliability is 97 %.

7.1.2 Development Model of Polygeneration System

Figure 7.3 lists the main modes of polygeneration such as "Vision 21" in the USA, the Syngas Park in Holland, EAGLE in Japan, and polygeneration integrated with IGCC and methanol in China etc. Most of these polygeneration system use coal as feedstock. In order to solve the energy security and environmental problem brought by fossil fuels use, nowadays, new polygeneration patterns based on renewable energy has attracted much attention, and the typical types are biomass based polygeneration and coal/biomass based polygeneration and so on. Besides, multi-energy (coal/solar, natural gas/biomass, coal/natural gas/biomass, coal/nuclear/biomass etc.) polygeneration systems already have been extensively investigated (Piatkowski and Steinfeld 2008; Sahoo et al. 2015: Wong 2012; Ruth et al. 2014). Polygeneration begin to pay more attention on its sustainability and

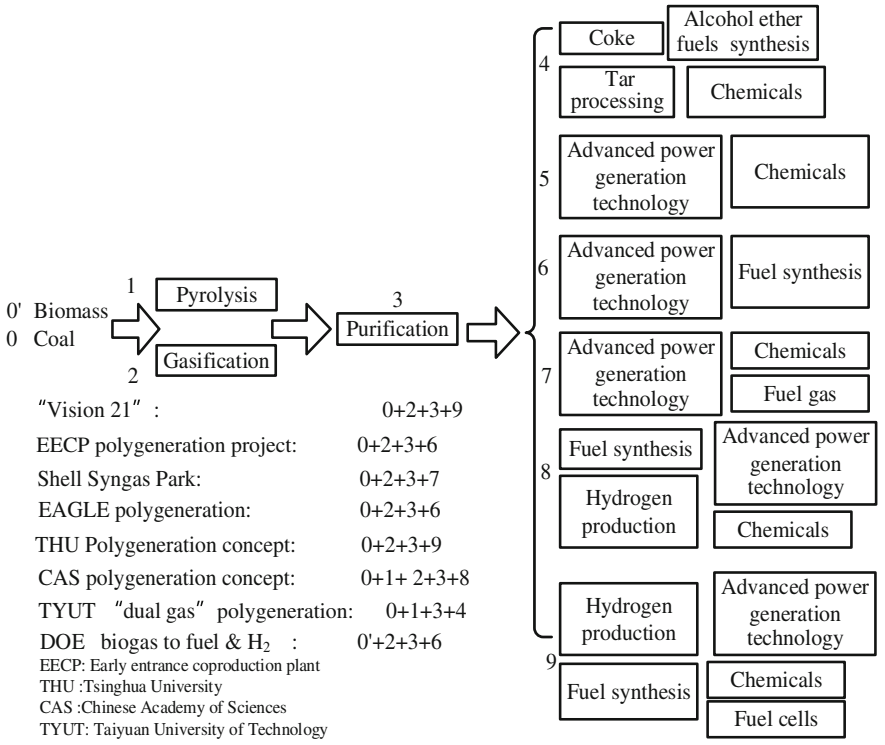

Fig. 7.3 The main modes of polygeneration

high efficiency. Currently, most of the polygeneration are still in the stage of concept and commercial demonstration, however, the successful development of IGCC provides a significant technical support, and it can be expected to be a bright prospect for polygeneration in the near future.

7.2 Basic Types of Polygeneration Systems

The combination of chemical engineering and power generating is the main feature of polygeneration. Polygeneration is also an important form of multifunction energy system, but the two are not equal. Polygeneration system is not a simple superposition of chemical system and power system, while it is an integrated and optimized coproduction system based on a special principle. Generally, polygeneration systems can be classified into three types: parallel type, series type and parallel-series.

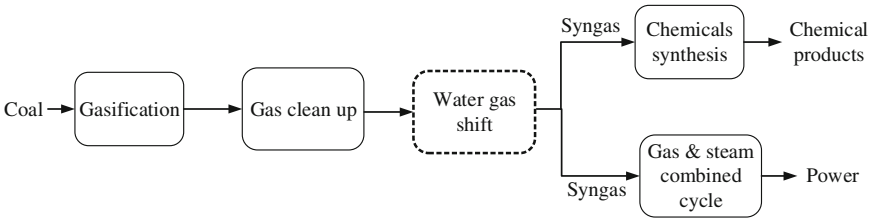

Fig. 7.4 The simplified flow diagram of parallel polygeneration system

7.2.1 Parallel Polygeneration System

Parallel system refers that the chemical process and power system are connected in parallel, and feedstock (such as syngas) is supplied to chemical process and power system respectively. The system basically maintains the inherent structure of the original production processes, and the chemical process and power system are still independent with each other. Figure 7.4 is a simplified flow diagram of a parallel polygeneration system. Syngas from coal gasification is divided into two parts: one part is sent for chemicals synthesis, and the rest is sent for power generation. The allocation proportion of syngas determines the distribution of chemical products and power. Since chemical unit and power unit of the system are independent from each other, so the improvement of energy utilization of the whole system is limited, and the energy saving is lower than 2 % (Lin et al. 2006).

Integrated parallel polygeneration system, which integrates the two energy systems (chemical process and power generation process) based on a simple parallel system, can obtain a more reasonable synthetic cascade utilization of physical energy (thermal integration). Compared with a simple parallel polygeneration system, the integrated parallel polygeneration system emphasizes on comprehensive optimization of chemical engineering and power. It simultaneously meets the process matching and the thermal integration of the two systems. The energy demand of all chemical processes is supplied by the power subsystem (Gao 2005). Ma et al. found that the energy saving rate of the integrated parallel polygeneration coproduction of methanol and power, which by the use of waste heat boiler, and the recovery of released gas and waste heat, can be reached to nearly 7 % (Ma et al. 2004).

7.2.2 Series Polygeneration System

Series polygeneration system is that chemical unit and the power unitare connected in series. Figure 7.5 is a simplified flow diagram of a series polygeneration system based on gasification. Syngas is firstly sent for chemical synthesis, then part of unreacted gas is recycled to chemical synthesis unit to increase chemical products

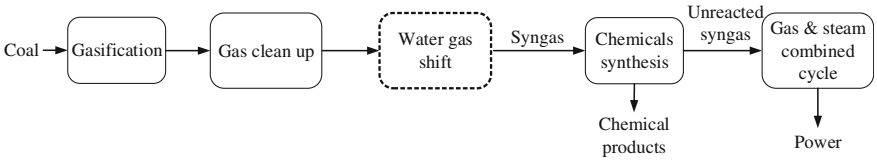

Fig. 7.5 The simplified flow diagram of series polygeneration system

according to the need, and the rest unreacted gas as fuel is sent to combined cycle for power generation. Compared with the parallel system, the most prominent feature of the series system is to break the basic structure of the separated production processes. The circulating ratio of unreacted gas in series system, which is similar to the allocation proportion of syngas in a parallel system, determines the distribution of chemical products and power (Wang 2006).

Series polygeneration system refers to simple series system and integrated series system. The main features of the simple series system are "no adjustment of syngas" and "syngas gas once-through". Syngas gas sent to chemical subsystem is not all converted to chemical products, and the water gas shift (WGS) process for composition adjusting of the syngas is avoided. Ma et al. (2004) studied on a simple series polygeneration system for producing methanol and power, and found that the energy consumption of the syngas preparation and the crude product distillation can be effectively reduced, and the relative energy saving of the system is about 7 %. Integrated series polygeneration system is an optimization and integration process on the basis of simple series polygeneration system, and the main characteristics are "no adjustment of syngas" and "proper circulation of unreacted gas". It simultaneously realizes composition adjustment of synthesis gas and improvement of effective gas utilization, and the relative energy saving ratio of the system increase by 12–15 % (Wang 2006). Gao (2005) compared the performance of the simple series system and integrated series system with production of methanol and power, and the results showed that the energy saving ratio of the latter is 15 %, indicating that energy saving effect of the latter is obviously superior to the former.

7.2.3 Parallel-Series Polygeneration System

Figure 7.6 is a schematic diagram of a series-parallel polygeneration system, which is integration and optimization of techniques involving allocations of series system and parallel system, composition reasonable adjustment, unreacted syngas partial circulation and energy cascade utilization and so on. The series-parallel polygeneration system realizes the complementary advantages of the series system and parallel system, presenting more potential for improvement of comprehensive performance. Syngas from coal gasification is divided into two parts: one part is sent for chemical synthesis and the other part is sent for power generation. The

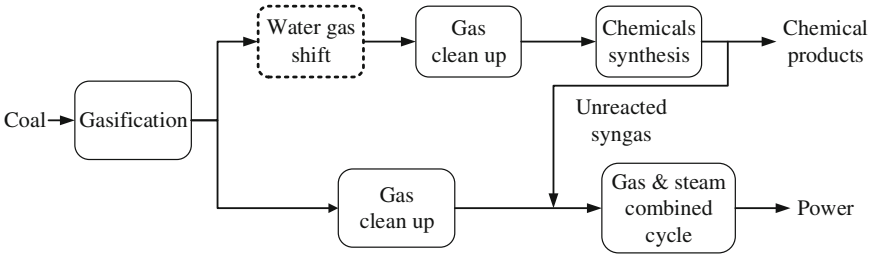

Fig. 7.6 Schematic diagram of series-parallel polygeneration system

syngas adopts once-through or partial circulation scheme for chemical synthesis, and unreacted syngas mixed with part of gasification gas as fuel gas are sent to power generation subsystem. Series-parallel system has high energy saving and high flexibility that series system and parallel system possess, respectively (Gao 2005).

7.3 Basis and Theories of the Integration and Optimization of Energy Polygeneration System

The new system integration must comply with the principle of cascade utilization of energy and classify conversion of element. Effects of process configuration, technology type and process conditions on system feature and performance are the guide of the design and optimization of the integrated system. Optimization process should reflect the feature of energy and materials conversion, especially in carbon, hydrogen and oxygen element conversion, which can help reveal the principle and mechanism of energy saving. The ultimate embodiment of optimization is that the polygeneration system can obtain the best operational conditions, technology configuration and production scales etc., on the basis of which the optimal total performance of the integrated system can be acquired.

7.3.1 Basic Principle for the Integration

For chemical production process, the major concern is the achievement of the highest conversion between raw material and product. Comparatively, energy utilization in the chemical production processes usually submitted to the pursuance of chemical conversion. This constraint has set the limitation on the energy utilization technologies adopted by the chemical production process. Thus, this is the main obstacle for the improvement of energy utilization (Cai et al. 2010). Otherwise, thermal energy utilization holds the dominant position in traditional energy

research. Polygeneration system is the integration that combines different types of energy resources (such as biomass, natural gas (NG), solar, wind etc.) and integrates chemical production and power generation into one system to obtain high energy efficiency, high economic benefits, high element conversion and low pollutant emissions. The basic principles for system integration mainly covers the four parts of the following (Cai et al. 2010):

- Coupling among different types of fuels

Multi-fuel feedstock input is one of the main features of the polygeneration system. The coupling with different types of fuels can be made according to their characteristics, advantages, and disadvantages. For example, those fuels being rich hydrogen, such as biomass, can be used to co-gasification with coal, and the syngas can avoid the WGS process to adjust H/C ratio. More importantly, different types of fuels can be utilized in a cascade way according to their energy level and chemical characteristics. This can promise the effective utilization of the fuels' chemical energy.

- Step by step conversion of effective components

The raw material is usually composed of different types of effective components for chemical production. The concept of the step by step conversion of effective components is the regulation of each component's utilization. Step by step conversion tries to meet the requirements of different chemical production processes and utilizes the different components selectively and orderly. This is achieved by relying on system integration instead of complex composition adjustment processes and recycle scheme.

- Cascade utilization of both chemical and thermal energies

Physical energy should be cascade-utilized according to temperature level. More importantly, in the integrated system, heat produced in the chemical production process can be effectively utilized in power generation according to its energy level. At the same time, a power subsystem can provide heat for the chemical production process with a matching energy level. Otherwise, an important point emphasized here is that chemical energy also should be cascade-utilized according to its energy level.

- Integration of energy utilization and pollutant control

Recovering pollutants before or within their production process is a promising technology which will lead to collaboration between energy utilization and environmental protection. Polygeneration system, coupling with chemical and power production processes, can bring a breakthrough for the joint solution of energy and environmental problems. The desired performance is that pollutants can be concentrated and removed without extra energy consumption accompanying with products being in production. With the production of excellent clean fuel such as

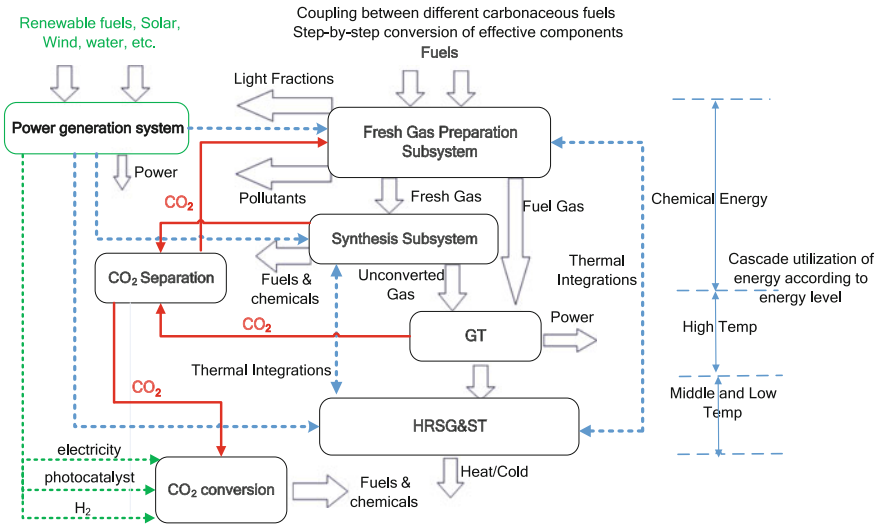

Fig. 7.7 Basic principles for integration of energy polygeneration system

methanol, DME, or hydrogen, the polygeneration system can recover CO_2 with reduced (or even without) energy penalty.

Figure 7.7 illustrates the basic principles for integration of energy polygeneration system. Carbon or hydrogen resource such as coal, biomass, NG, unconventional natural gas (UNG) and COG etc. may be converted into fresh gas in a fresh gas preparation subsystem. The primary function of the fresh gas preparation subsystem is to produce qualified syngas gas from raw materials. In the subsystem, some methods such as partial-gasification or co-gasification/pyrolysis, and partial-reforming or dual-fuel reforming can be used to produce the syngas. Thus, the syngas with proper compositions can be sent to the chemical synthesis unit. At the same time, the part of the syngas that is difficult to be converted for chemical production can be used as fuel for power generation. Non-fossil fuels (or renewable fuels) such as solar or wind can provide the heat, electricity and hydrogen needed in chemical conversion processes and power generation process. More importantly, CO_2 can be recovered with less (or even without) energy penalty from synthesis subsystem and flue gas. On the one hand, the captured CO_2 can recycle to fresh gas preparation subsystem to be converted into fresh gas or fuel gas by methanation, reforming and gasification etc. On the other hand, CO_2 also can be converted into fuels and chemicals by electrochemical conversion, photocatalytic conversion and hydrogenation reduction which are supported by renewable fuels.

7.3.2 Optimization of Objective Functions

The optimization process can be described as follows: given a set of feedstock and desired final products with specifications, it is desirable to generate a systematic framework with various objectives (such as energy, economy and environment etc.) for the extraction of the best configuration and unit operation with the optimal technical conditions among an astronomically large number of promising alternatives (Yuan and Chen 2012). However, multi-objectives brings a new challenge to system analysis and evaluation, which urgently requires to establish multi-objective comprehensive analysis and evaluation methods (Tan et al. 2014; Wang et al. 2014; Li et al. 2015; Ahmadi et al. 2014). In general, the three aspects are: energy/exergy, economy and environment. These need to be studied for energy system evaluation (Li et al. 2014; Yi et al. 2014).

For a given set of feedstock, several promising technologies can be utilized to convert them to final products. For the device level of a given energy conversion system, there are probably multiple alternatives for each operating unit, for example, fluidized bed reactor, moving bed reactor, and fixed bed reactor can be selected for the solid sorbent based carbon capture process. Of course, different unit will generate various capital costs. Apparently, the decision on the operating unit selection will significantly affect the whole flow sheet. For a given energy system, there usually exist several conflict objectives such as environmental requirement and overall process cost. As an example, for a coal-based power plant with CO_2 capture system, adding the CO_2 capture device will improve the capital cost and reduce the efficiency of the power plant. In addition, the process flow sheet will vary with the changes in multi-objectives trade-offs. All of the above aspects demonstrate that the selection of the optimal technology and corresponding operating unit to systematically deal with the above conflicts based on the system engineering viewpoint requires a systematic optimization methodology. Majority of investigations for operation optimization of energy conversion systems focused on single objective optimization, such as cost or energy minimization or economic optimization as objective (Dimopoulos et al. 2013; Ratlamwala et al. 2012; Yi et al. 2012a, b). However, the optimization of energy conversion system should cover multicriteria that is integrated from different fields of science: energy, economy, environment and even society. Multicriteria optimization and assessment is the most important to the system's reliability, feasibility and sustainability. Hence, most of recent studies for operation optimization of energy conversion system establish multi-objective optimization models to obtain comprehensive optimal operation strategies considering cost saving, energy saving and emission reduction. As above three aspects often conflict with each other, how to optimize these objectives in a comprehensive way is the main problem of multi-objective optimization. Usually, the weighted method (Wang et al. 2009; Ren et al. 2010; Jing et al. 2012), genetic algorithm method (Abdollahi and Meratizaman 2011; Ebrahimi et al. 2012; Jabbari et al. 2013), lexicographically stratified method (Wu et al. 2012) and evolutionary algorithm method (Abdollahi and Sayyaadi 2013; Kavvadias and Maroulis 2010)

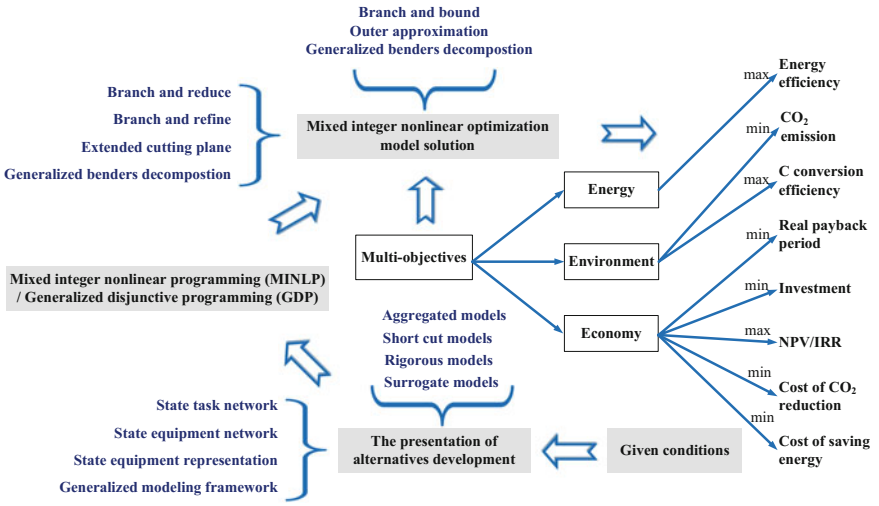

Fig. 7.8 The brief flow sheet of the mathematical programming-based approach

are used in multi-objective optimization. Among these technologies, mathematical programming-based approach is the most widely used approaches for the optimization of both specific subsystems and total process flow sheets. Figure 7.8 demonstrates the brief flow sheet of the mathematical programming-based approach.

7.4 CO_2 Conversion and Utilization in Polygeneration System

7.4.1 The Possible Routes for CO_2 Conversion and Utilization

CO_2 is inert and its chemical transformations are thermodynamically highly unfavorable (Zangeneh et al. 2011). Its inertness is due to its being the most oxidized state of carbon. There are various possibilities to activate and convert CO_2, but it is necessary to overcome a thermodynamic barrier. Consequently a substantial input of energy, optimized reaction conditions and (almost invariably) active catalysts are required for any transformation of CO_2 into useful products. As a result, the strategies for CO_2 conversion and utilization should be based on the unique physical and chemical properties, some strategic considerations for developing technologies for CO_2 conversion and utilization are listed below, which will be further elaborated in subsequent sections.

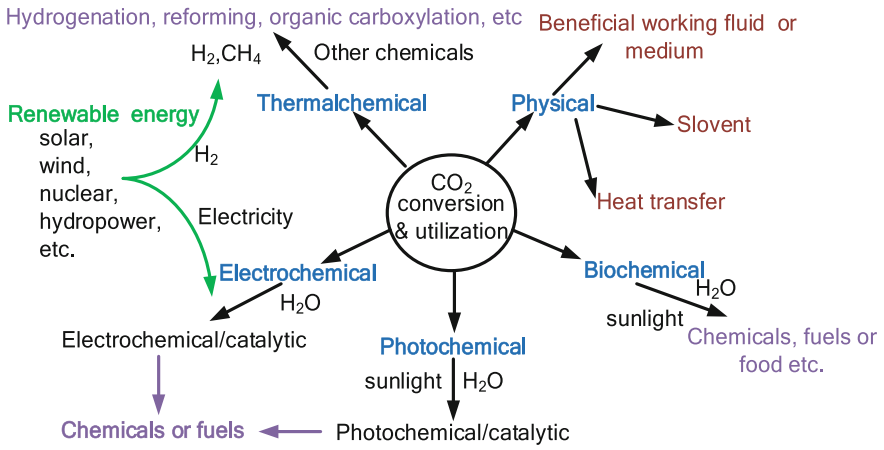

Fig. 7.9 Possible routes for CO_2 conversion

- Use CO_2 as C-source to produce industrially useful chemicals and materials that adds value to the products.
- Use CO_2 as a beneficial fluid for processing or as a medium for energy recovery, contaminant removal, and emission reduction.
- Use CO_2 recycling involving renewable sources of energy such as biomass and solar energy or 'waste' energy to conserve C sources for sustainable development.

Therefore, there are two major pathways for CO_2 conversion (Fig. 7.9) (Olajire 2013), chemical way and physical way. CO_2 generally is used as a storage medium for renewable energy or a feedstock for various chemicals. CO_2 can be converted via thermal-chemical, biochemical, electrochemical and photochemical processes to other energy storage chemicals, such as syngas, formic acid, methane, ethylene, methanol, and DME (Jhong et al. 2013; Olah et al. 2008). The entire portfolio of commodity chemicals are currently manufactured from a few primary building blocks or platform chemicals in the fossil-based chemical industry. CO_2 can also be used as a source material that can be converted into a similar suite of building block chemicals. Insertion of CO_2 into epoxides to manufacture various polymeric materials is an exciting technology as it does not only utilize CO_2, but also avoids using fossil feedstock and creating CO_2 emissions. It was estimated that the various chemical conversion pathways could consume approximately 0.3–0.7 Gt/y of CO_2 (Aresta and Dibenedetto 2004). Besides, there are huge potential in the aspect of the physical way for CO_2 utilization as well, for instance, CO_2 as a beneficial working medium to replace N_2 will effectively increase the generating efficiency of gas turbine.

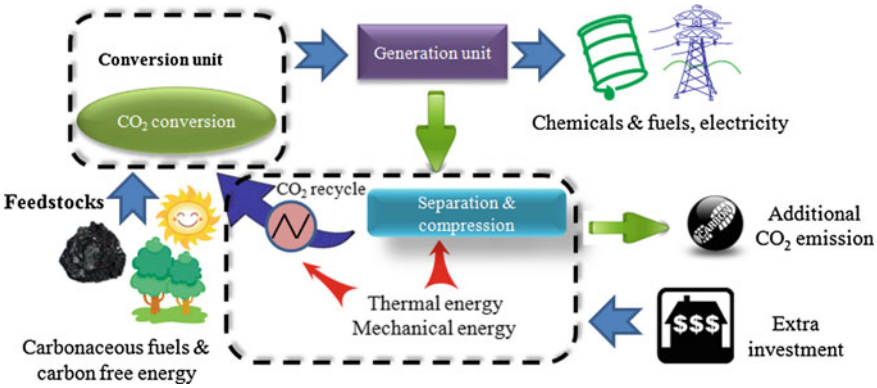

Fig. 7.10 Diagram of CO_2 recycle in energy polygeneration system

7.4.2 *The Relationship Among Carbon Conversion, Cost and Carbon Cycle*

7.4.2.1 Carbon Cost and Carbon Cycle

Many products can be obtained from energy system. Different production processes of chemical or fuel products accompany with different carbon emission intensities. In other words, the carbon in product is at the cost of some other carbon being emitted. CO_2 recycle can improve the carbon value in CO_2, but it is predicted that the generated CO_2 recycle use will cause additional energy consumption, CO_2 emission and cost, as shown in Fig. 7.10. CO_2 is a highly stable molecule. Consequently, a substantial input of energy, effective reaction conditions, and often active catalysts, are necessary for chemical conversion of CO_2.

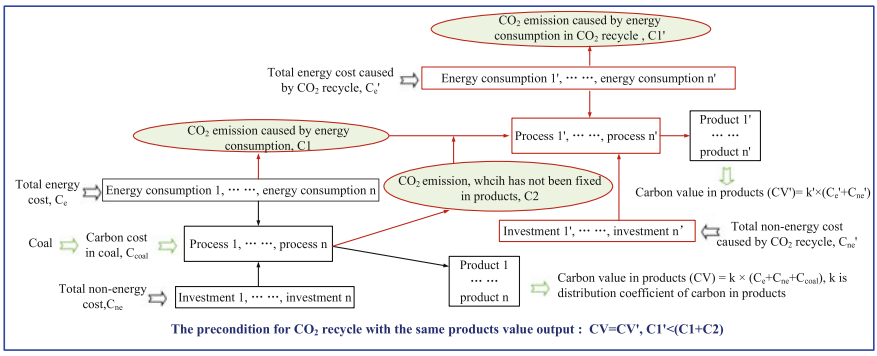

Fig. 7.11 Carbon value in the carbon conversion process and recycle

The total CO_2 emission includes: (1) CO_2 emission caused by energy consumption in each production unit; and (2) CO_2 emission from conversion of a part of C in carbonaceous energy. Each chemical product has its own value which is determined by production cost that is comprised of energy cost and non-energy cost (capital cost and operation & maintenance cost), as can be seen in Fig. 7.11. According to the use value of C, the carbon value in CO_2 is zero, as a result, the value of the carbon in chemical products is the sum of the total value of all carbon input the system. It is required using CO_2 recycle to produce the same product should present lower CO_2 emission, energy consumption and production cost compared with that of production without CO_2 recycling, or else CO_2 recycling makes no sense. Technology and final products selection, and investment of CO_2 recycle process determine the final value of the carbon of the recycle CO_2. Although a small bit more products are produced, the carbon cycle has no future as many studies concluded if no carbon tax is adopted, because the input of the capital and energy outweighs the output revenue. Therefore, there exists a balance among the carbon cycle, emission and other treating scenarios when mature carbon policies are made. Integration and optimization of coal conversion coupled with CO_2 recycles are considerably significant and necessary.

7.4.2.2 Carbon Cycle and Carbon Tax

As more and more fossil fuels are used annually, energy saving and emission reduction is becoming a hot topic of the related researching fields (Fang et al. 2013). The problems of how to harmonize the relationship between necessary energy consumption and emission controlling have attracted wide attention around the world (Niu et al. 2011). Low carbon economy or green economy requires more attention to be paid to the CO_2 reduction when developing economically. Traditionally, the Pigouvian fee on CO_2 and other greenhouse gas via carbon taxes or emissions caps and permit trading are regarded as the economically optimal policies to address the climate change-related problems (Jenkins 2014). Carbon tax is an incentive-based policy instrument for controlling GHG emissions. Many economic studies indicate that low-rate carbon tax policy can effectively reduce GHG emissions without having significant negative impacts on the future economic growth (Lu et al. 2010). Nevertheless, carbon tax is only the restrictive or punitive measure for CO_2 reduction, but won't get to the root of the problem. As there is not yet any practice about carbon tax in most countries, the pricing for carbon tax is highly volatile. A summary of international policies on carbon tax is listed in Table 7.1 (Zhou et al. 2012).

The carbon policies in different countries indicate that in most cases carbon taxes are not greater than $30 per ton of CO_2 emissions. As a result, an unwilling result will be presented that when the cost of CO_2 reduction is higher than carbon tax, the CO_2 producer would pay for carbon rather than reduce CO_2. For instance, the production cost of methanol is today 0.08 €/kg, in the best case considering the capital expense (CAPEX) and operating expense (OPEX), methanol can be

Table 7.1 Summary of international policies on carbon tax

Country	Current carbon tax ($/t-$CO_2$)	Year started	Policy details
Denmark	≈18	2002	Levied on hydrocarbon fuels according to the level of pollution each source emits
Ireland	20	2010	Apply to kerosene, marked gas oil, liquid petroleum gas, fuel oil, and natural gas
Netherlands	19.53	1992 and 1996	Collected on hydrocarbon fuels only for energy use; a 50/50 carbon/energy tax which is assessed partly on carbon content and partly on energy content
Sweden	140	1997	Levied on transport, space heating, and non-combined heat and power generation
Norway	21	1991	Levied on hydrocarbon fuels
Switzerland	11.41	2008	Collected on hydrocarbon fuels such as coal, oil and natural gas, unless they are used for energy
Canada	15–30	2007–2008	Implemented in Quebec, British Columbia, Alberta
United States	≈5	2006–2010	Implemented in Colorado, California, Maryland
South Africa	–	2010	Apply at the time of sale; $9.72 will be added to the price for every gram of CO_2 per kilometer the vehicle emits over 120 g/km
India	$1.07/t-coal	2010	Levied on produced or imported coal into India
Costa Rica	3.5 % of fuel price	1997	Levied on hydrocarbon fuels

produced from electrolytic hydrogen at a cost of 0.3 €/kg, that is circa four times higher than current price. What may make the balance positive for industry is the fact that if CO_2 tax (C-tax) can be at least 160–200 €/t (Centi and Perathoner 2009). This price for CO_2 tax is unpractical. More importantly, CCS has been received a lot of attention in recent years for reducing CO_2 emissions. As an effective and advanced technology, CCS has huge potential in CO_2 reduction, however, it has always been controversial as to whether CCS should be implemented, due to its high cost and big risks. Low carbon tax is not attractive enough to promote the development and application of CCS. Fossil fuel runs out gradually nowadays, and the storage of CO_2 means a waste of C resource. CO_2 as a main carbon feedstock to replace a considerable part of fossil fuels will be an indisputable fact. Therefore, reducing CO_2 should comprehensively consider the sustainable use of the resource and economy.

CO_2 recycle as a technology of CO_2 conversion can convert CO_2 into high valuable products. These value-added products may thus offset costs associated with carbon management, and also make use of C resource thus reduce CO_2 emission. Under the condition of low carbon tax or even having no extra subsidies with respect to CO_2 reduction, CO_2 recycle is more reliable and feasible from the

viewpoint of economy and resource use compared with CCS, and it is a sustainable technology for CO_2 reduction and control.

7.4.3 *The Utilization of the CO_2 Recycle in Energy Polygeneration System*

Building upon energy depletion and CO_2 emissions challenges, various innovative with CO_2 recycle have been developed for the production of cost-competitive fuels, chemicals. The following different types of energy polygeneration systems with CO_2 recycle have arisen much concern by both the industry and academia due to their sustainability.

7.4.3.1 CO_2 Recycle as a Gasifying Agent in Coal-Based Polygeneration System

Gasification is a frequently applied processing technique in coal conversion process. Many high valued products are achieved from gasification. One of the methods to increase the carbon conversion efficiency is to reuse the CO_2 produced during the process as a gasifying agent and put them back into the gasifier. However, during the process of CO_2 forming—conversion—reforming, the C value does not increase, but keeping cycling from high to low (here assuming the amount of carbon cycling is fixed, may be easier to understand).

CO_2 has been proposed as a candidate for gasifying agent since its utilization has the potential to further reduce CO_2 emission and can reduce steam consumption (Chaiwatanodom et al. 2014). Thus, valuable steam can be saved, resulting in a

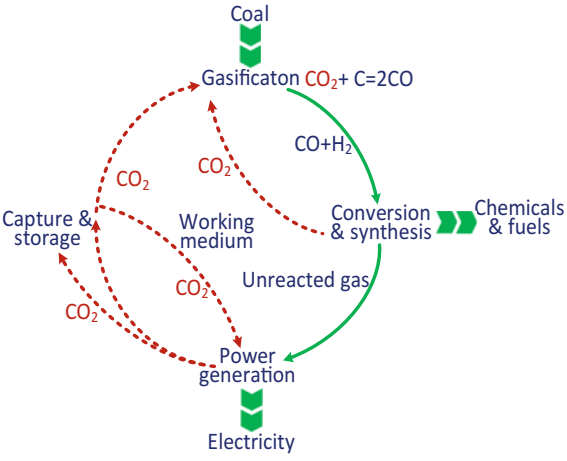

Fig. 7.12 Coal gasification with CO_2 recycles in coal-based polygeneration system

reduction of phenol water produced in the gasification process, especially in fixed-bed gasifier (Lurgi, BGL etc.) (Chaiwatanodom et al. 2014; Wang et al. 2011) as well as a reduction in the load of wastewater to be treated. CO_2 can react with C at temperatures above 1073 K ($C + CO_2 = 2CO$). Generally, the temperature of the gasifier is higher than 1273 K. The higher the temperature is, the higher the chemical activity of CO_2 will be. Increasing the amount of CO by recycling CO_2 into gasifier is a smart way to improve carbon element utilization and downstream chemicals output. The recycled CO_2 proves to increase syngas production. A scheme design based on coal-based polygeneration system with CO_2 recycle is presented in Fig. 7.12. The CO_2 generated from chemicals or fuels production process as a gasifying agent are recycled into gasifier. Besides, the CO_2, from power production process, also can be recycled and utilized for gasification of coal to increase fuel utilization and to decrease the CO_2 emissions into the atmosphere, or as a working medium are recycled to gas turbine to increase power output (Yi et al. 2012b, 2013; Jillson et al. 2009). In this way, a considerable amount of can be used and fixed into chemicals or fuels finally, which can improve carbon element utilization efficiency and energy efficiency due to avoidance of the amount of CO_2 capture and storage.

A variety of systems based on this CO_2 recycle have been reported (Yi et al. 2012b; Jillson et al. 2009; Yang et al. 2013). Kidoguchi et al. (2011) have researched a new system that combines an oxygen-CO_2 blown coal gasifier that uses CO_2 in exhaust gas with a closed gas turbine. It has been experimentally clarified that char gasification reaction is promoted by enriched CO_2. A high net thermal efficiency (40 % or more) could be expected because of a more simplified system without WGS unit or CO_2 separation unit. In terms of the whole system performance. Yi et al. (2013) have proposed a novel system in which CO_2 is recycling into gasifier and gas turbine with a high exergy efficiency (46.3 %) and CO_2 emission (0.47 t/(t-coal)). Yang et al. (2013) have done the similar works by combining dry reform reaction with CO_2 recycling into gasifier in a polygeneration system to produce chemicals as well as generate power. The same results show that the new process is promising since it reduces the CO_2 emission and increases the carbon efficiency and the energy efficiency. Otherwise, the process with CO_2 capture and utilization has advantages in both technical and environmental aspects and even in economic aspect with introducing carbon tax (Yi et al. 2013, 2014; Man et al. 2014).

CO_2 recycling into gasifier is a promising way for CO_2 reduction and reuse in coal coal-based polygeneration system, however, the key technology is to control the CO_2 recycle ratio. Too much CO_2 recycling mass flow leading to a low gasification temperature and a low H/C ratio would give negative influences on the gasification efficiency, the gasifier slag-off performance and the downstream products synthetic ratio if the recycling ratio of CO_2 is above a certain quantity, though a high portion of CO_2 get into gasifier could substitute for steam and improve carbon element utilization (Yi et al. 2012b, 2013). Relative research over coal gasification with CO_2 such as coal rank, pressure, temperature, gas composition, etc., and the kinetics and the reaction rate equations for coal-car gasification

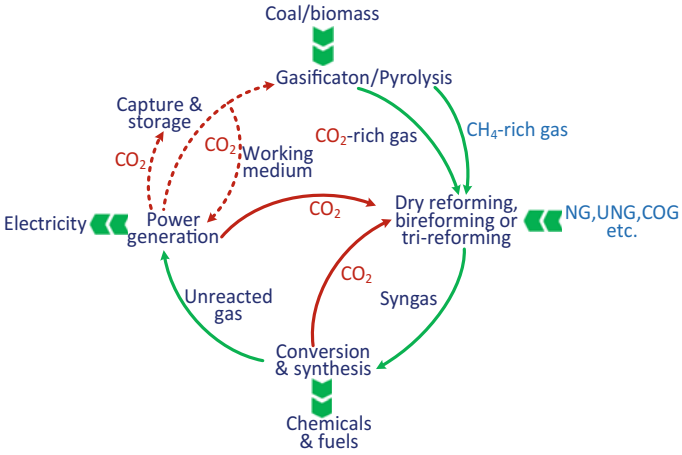

Fig. 7.13 Scheme of CO_2 conversion and recycle coupled with coal/biomass and CH_4-rich gas based polygeneration system

both in the reaction kinetic control region (low temperature) and the diffusion control region (high temperature) is discussed in literatures (Irfan et al. 2011), which will be of great help to the development and application of this technology.

7.4.3.2 CO_2 Recycle Coupled with Coal/Biomass and CH_4-Rich Gas Based Polygeneration System

Two of the most important GHG, CO_2 and CH_4, are consumed into useful and valuable product, syngas (hydrogen and carbon monoxide) by reforming (dry reforming, bireforming or tri-reforming) that retards GHG emission and increases the C element utilization efficiency (Theofanidis et al. 2016; Itkulova et al. 2014; Luu et al. 2015). The synthesis gas with a proper mole ratio of CO/H_2 is preferentially used for the production of liquid hydrocarbons or chemicals synthesis, and the conventional WGS process can be avoided. CO_2 and CH_4 can therefore be chemically transformed into a valuable carbon source of the future allowing environmentally neutral use of carbon fuels and derived hydrocarbon products. These advantages will undoubtedly result in high energy utilization efficiency and low CO_2 emission for the whole production system. An integrated scheme in terms of CO_2 recycle coupled with coal/biomass and CH_4-rich gas based polygeneration system is proposed in Fig. 7.13.

Due to the favorable environmental and energy feature of coal/biomass and CH_4-rich gas based polygeneration system integrated with CO_2 recycle, the integrated energy conversion systems have attracted significant attention in recent years. A polygeneration system with COG and coal gasified gas, in which dry reforming was used to convert $CH_4 + CO_2$, was proposed by Xie et al. and chemical energy,

internal rate of return and power output increase by 11.5 %, 1.3 % and 8.4 % respectively, and the CO_2 emission reduces by 33.8 % (Yi et al. 2012a; Xie et al. 2010). Lim et al. (2012) have proposed a combined steam reforming process with dry methane reforming process instead of the removal of the former and then give a conclusion that the combined system could reduce more net CO_2 emission by 67 % than a single steam reforming process. Yang et al. (2013) proposed a coal-to-olefins system that recycled CO_2 into both gasifier and dry reforming unit considering the CO_2 recycling distribution between the two paths further. They get conclusions that the new system is 2 times higher at carbon efficiency, 1.5 times lower at CO_2 emission rate, and 1.3 times higher at energy efficiency than those of old processes. Song and Pan (2004) proposed a novel conception called tri-reforming of methane using CO_2 in the fuel gases without CO_2 separation. A comparative energy analysis by calculation indicates that tri-reforming is more desired for producing syngas with H_2/CO ratios of 1.5–2.0 compared to CO_2 reforming and steam reforming of methane, in terms of less amount of energy required and less net amount of CO_2 emitted in the whole process for producing synthesis gas with H_2/CO ratio of 2.0. Besides, the similar studies are conducted with respect to NG, COG or CH_4-rich gas reformed with CO_2 or exhaust gas for CO_2 recycle to produce methanol, DME, olefins etc. are also widely investigated, and all that presented excellent performance in CO_2 emission reduction, products output increase and energy saving compared with that without CO_2 recycle or with CCS processes (Lin et al. 2014; Adams II and Barton 2011; Bermúdez et al. 2013; Man et al. 2014b; Zhou et al. 2008).

In summary, CO_2 recycle coupled with coal/biomass and CH_4-rich gas based polygeneration system is a good option for dealing with deficiencies such as low efficiency and high investment cost typically associated with CO_2 control and reduction. However, CH_4/CO_2 reforming (no matter dry reforming, bi-reforming or tri-reforming) reaction is strongly endothermic. It is therefore high-energy-consuming to use this reaction for CO_2 mitigation. As the proportion of CO_2 for the CH_4/CO_2 reforming reaction increase, the amount of CO_2 conversion is spontaneously increased but more energy is consumed at the same time. Thus, it is important to select an appropriate CO_2 recycling ratio for best trade-off of energy and resource utilization (Yang et al. 2013). More importantly, high activity and stability catalyst for CO_2/CH_4 reforming reaction still needs to be developed, which is the key for industrialization of this recycle technology.

7.4.3.3 CO_2 Conversion and Recycle in Coal-Biomass Based Polygeneration System

From the viewpoint of lifecycle, biomass, a renewable energy source indirectly comes from solar energy, incorporated with coal to produce chemicals or electricity can reduce coal consumption thus results in the decrease of CO_2 emission. The C element will be released into the atmosphere in the form of CO_2 which will be fixed into biomass by photosynthesis finally, which completes the carbon cycle. A closed

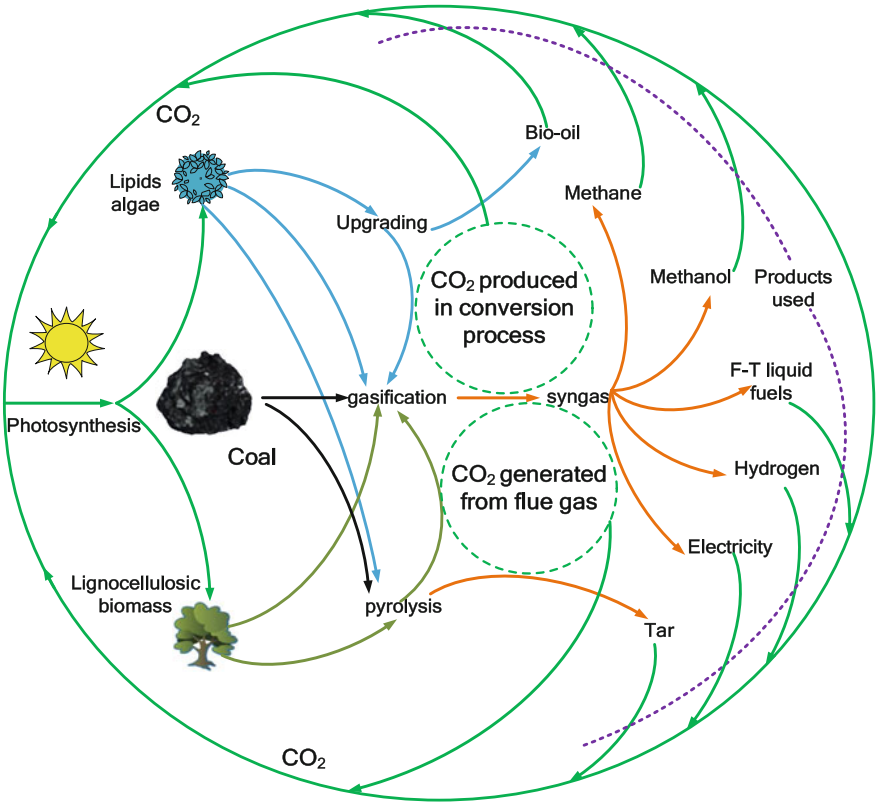

Fig. 7.14 CO_2 conversion and recycle in coal-biomass based polygeneration system

carbon circulation can be found in polygeneration system of coal integrated with biomass production in Fig. 7.14.

With a number of advantages based on co-production of coal integrated with biomass, extensive attentions have been attracted to the design and optimization of a newly integrated system with different feasible technologies and routs to produce chemicals or fuels such as F-T hydrocarbons, methanol, hydrogen and electricity etc. (Trop et al. 2014; Liu et al. 2010b; Ruhul Kabir and Kumar 2012; Cormos 2014; Williams et al. 2009). Especially, it was estimated that the coal/biomass co-generation system with products of F-T fuels and power can even obtain zero fossil CO_2 emission, and the co-generation of liquid fuel and hydrogen from coal and municipal solid waste, whose hydrogen is generated from cellulosic-waste gasification, is economically competitive at crude oil price above $66/bbl in the base coal/biomass to liquid fuels plant compared to nowadays about above $80–100/bbl (Wu et al. 2012b). The results presented clearly demonstrate the economic competitiveness of coal and biomass refineries which produce liquid transportation fuels while simultaneously reducing the life-cycle of the greenhouse

gas emissions. In Muresan et al.'s LCA study (Muresan et al. 2014), the CO_2 emission for the cases of co-gasification of coal and biomass are greatly lower than the cases of coal gasification alone, and the acidification and eutrophication potentials are lower for the co-gasification cases. That is because biomass as a feedstock in energy generation still releases CO_2 back to the atmosphere. There would be an overall reduction in emissions per unit energy as the CO_2 is being recycled by the biomass producing the extra energy per unit of CO_2 emissions. It can be assumed that the release of CO_2 during biomass co-production is recycled and is reused by the biomass. Essentially the net CO_2 emissions produced by biomass are zero so the net emissions remain the same while the energy output is significantly increased.

Coal-biomass co-production potentially makes a significant contribution to CO_2 recycle through biological system, and it has many advantages over conventional carbon sequestration methods as the CO_2 is being utilized by photosynthesis to produce renewable biomass which can be recycled into the production system for energy production. It is a more desirable method of carbon sequestration compared with geological storage where CO_2 is pumped into aquifers, because even despite the concerns over leaks in the longer term it has a high cost associated with the separation of CO_2 from exhaust gases without any economic gains from the process. However, there are many obstacles to be solved. Major challenges include cost and CO_2 emission reduction through breakthrough processes (culture, harvesting, drying/extraction and separation), no mature commercial operation experience for widely application of coal coupled with biomass conversion, as well as integration and optimization of co-production system to obtain the better whole performance.

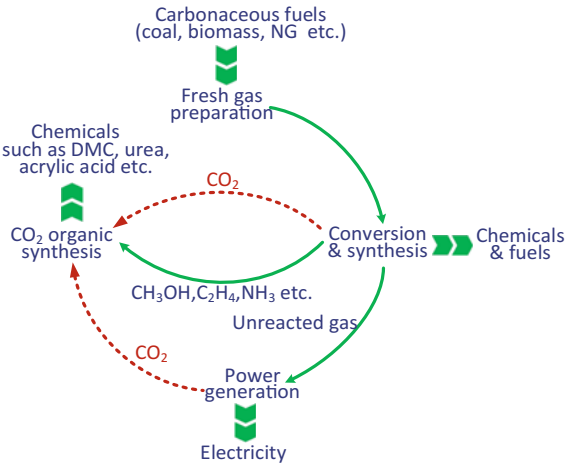

Fig. 7.15 A concept of CO_2 recycling as a feedstock for organic synthesis in energy polygeneration system

7.4.3.4 CO_2 as a Feedstock Recycling for Organic Synthesis in Energy Polygeneration System

CO_2 is a nontoxic, non-flammable, abundant and economical C1 feedstock. Currently, carbon dioxide is used industrially for the synthesis of urea, salicylic acid, inorganic carbonates, methanol, cyclic carbonates and polycarbonates (Centi and Perathoner 2011; Peters et al. 2011; Olajire 2013). Major emphasis is on replacing phosgene-based routes by introducing CO_2 into organic substrates for synthesis of various chemicals (Aresta et al. 2013). These successful cases of CO_2 conversion for organic synthesis have provided huge potential in aspect of CO_2 recycle and reduction for coal chemical production. A concept of energy polygeneration system with CO_2 recycles as a feedstock for organic synthesis is shown in Fig. 7.15. Chemicals such as CH_3OH, C_2H_4 and NH_3 etc. are produced from carbonaceous energy and materials through a series process in terms of gasification/pyrolysis, clean up, reforming/WGS and synthesis etc. The CO_2 generated from chemical production process and power plant can both serve as raw material of organic synthesis to react with those chemicals to produce new chemicals or fuels. Due to the maturity of CO_2 to urea, recently, many studies focused on CO_2 recycled from chemical production or power generation processed to produce urea incorporating with other liquid fuels or chemicals production (Kaggerud et al. 2006; Meerman et al. 2011, 2012; Li et al. 2003).

7.4.3.5 CO_2 Recycle in Multi-feedstock Energy Polygeneration System

CO_2 produced from energy system can be converted into chemicals and fuels by hydrogenation, electrochemical and photochemical processes. Generally, hydrogen sources and electricity for the chemical recycling of CO_2 could be generated from carbon-free energy (such as solar, wind, tide, geothermal, nuclear etc.) by splitting

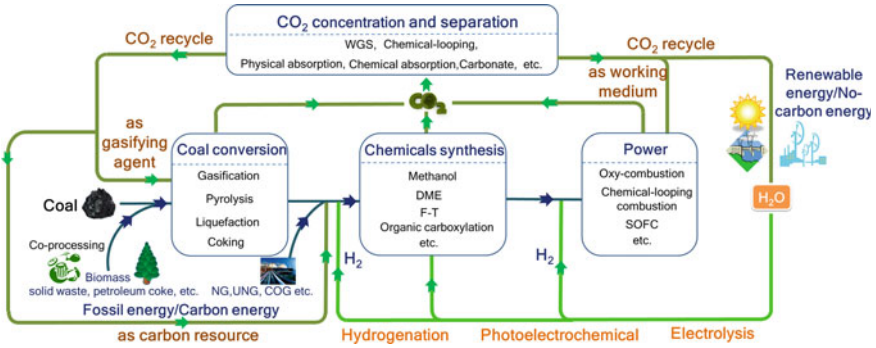

Fig. 7.16 Diagram of MFEPS with CO_2 recycle

water (by electrolysis or other cleavage) or generating power. An integration of CO_2 recycle in MFEPS was presented in Fig. 7.16. The most striking feature for this energy integration system is using carbon-free renewable energy to realize CO_2 recycle, and as a result, that non-carbon renewable energy is stored in the synthesized chemicals or fuels in the form of chemical energy that is much easier to transportation and use. Meanwhile, CO_2, from carbonaceous fuel, is almost totally fixed into hydrocarbons, and the whole carbon resource conversion process presents near-zero CO_2 emission. The net carbon neutral cycle can be achieved by recycling CO_2. Further improvement in carbon neutral cycle could be possible by utilizing CO_2 from atmosphere in addition to recycling CO_2 from source.

By synthetically combining different carbonaceous energy and renewable energy, the using of advanced CO_2 control and conversion technologies and the co-production of alternative fuels and power, MFEPS can present more opportunities to simultaneously achieve higher efficiency, lower investment, and less environmental impact as compared to traditional energy systems. For this type of MFEPS, according to energy level and chemical characteristics, some given processes of energy and elements conversion can be coupled together. Moreover, different advanced CO_2 control and conversion technologies (such as chemical-looping, CO_2/CH_4 reforming, photo/electrolysis of CO_2, and hydrogenation of CO_2 etc.) coupled together in MFEPS can provide a better platform for CO_2 conversion, concentration, separation and capture with low or even no energy consumption. MFEPS with characteristic of multi-products can supply with more opportunities to achieve energy saving and economic way for CO_2 recycle. Therefore, it is more efficient and economical for CO_2 recycle in MFEPS compared with traditional single-feed single product energy system. A series of existing research reports have proclaimed that the MFEPS including multiple CO_2 conversion and recycle routes would be the best solution to the sustainable cost-efficient energy system development (Perathoner and Centi 2013, 2014; Boretti 2013; Centi and Perathoner 2011; Goeppert et al. 2014; Graves et al. 2011; Tahir and Amin 2013; Budzianowski 2010). Although MFEPS showed preferable sustainability, since material, energy, and capital flows are crossing and coupling through MFEPS, it is much more complex than systems with single-feed single product without CO_2 recycle. Except for improving CO_2 conversion and utilization in unit process, configuration optimization and assessment should be conducted on MFEPS, which will be a topic of future work.

There are already many possible technological pathways reported by literatures for CO_2 recycling into fuels such as methanol, DME and F-T synthesis etc. and the most apparent differences between them focus on developing and optimizing of three stages—usage of different energy, H_2O and CO_2 dissociation, and synthesize various fuel (as shown in Fig. 7.17). Graves et al. (2011) proposed possible technological pathways for recycling CO_2 into sustainable hydrocarbon fuels using renewable or nuclear energy in a non-biological process. Initially, CO_2 captured

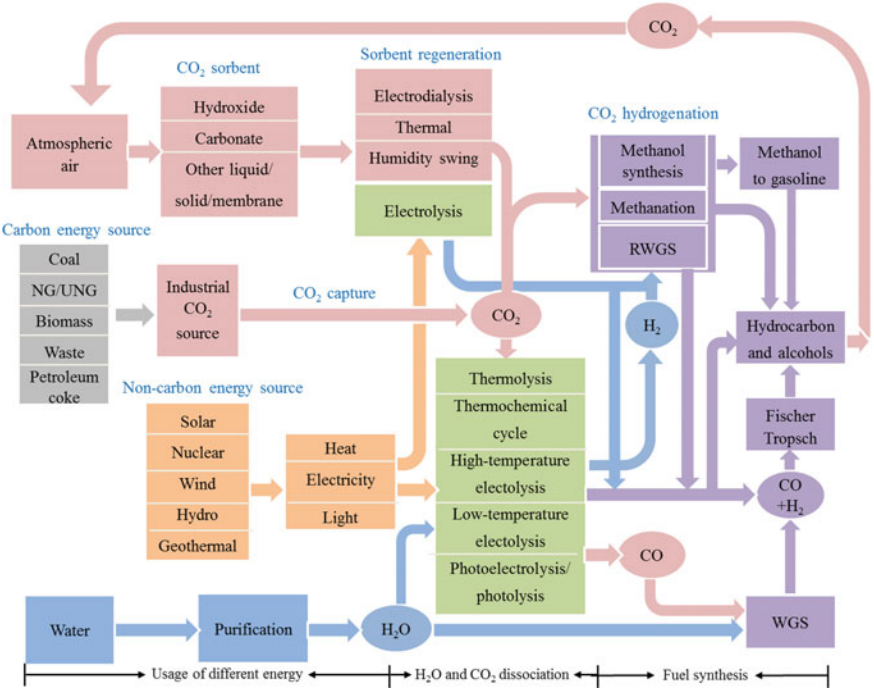

Fig. 7.17 The possible pathways of CO_2 recycle coupled with non-carbon energy for hydrocarbon fuels or chemicals synthesis (Modified from Graves et al. (2011))

from large industrial sources (e.g. aluminum plants) could be utilized. In the long term, the capture of CO_2 from the atmosphere would enable a closed-loop hydrocarbon fuel cycle. Typically, "Methanol Economy" with carbon recycling has been proposed by Olah (Dibenedetto et al. 2014). Any available energy source (alternative energies such as solar, wind, geothermal, and atomic energy) can be used for the production of needed hydrogen for chemical conversion of CO_2 that is air-captured or by low-temperature electrolysis (as shown in Fig. 7.18).

Although these techniques are not yet technologically mature, they offer potentially a feasible and environmentally benign way to convert both carbon-free energy and CO_2 into chemical energy stored liquid fuels or chemicals implying that would have wide application for CO_2 recycle in the future due to excellent performance in the aspects of CO_2 control and reduction. However, this technology seems to be difficult to realize commercialization in the short or mid-term. This is because most of the cost of products is higher than that of traditional production processes, and this will be a major challenge to using renewable wind or solar technologies. Otherwise, even the energy demand and CO_2 emissions for some production are higher than existing technologies in lifecycle. Water is also an important restrictive condition for the development and application of the technologies. These drawbacks due to operation only for a fraction of a day or year

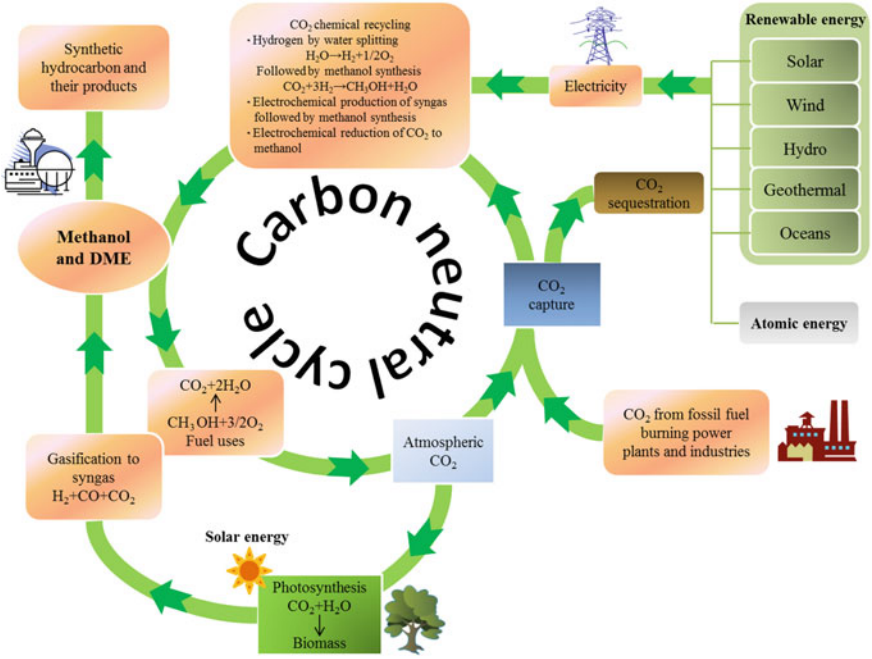

Fig. 7.18 CO_2 recycling in the methanol economy (Modified from Goeppert et al. (2014))

(especially for solar and wind) which thus result in much higher fixed-capital costs and frequent switch on/shut off, with a consequent decrease in the life time of the equipment and components and increase in the operation cost. Moreover, the low integration in energy coupling, as well as small scale capacity for photo-electrolysis device and low catalytic efficiency will all lead to the increase of the cost per unit product. As a result, in order to introduce this CO_2 recycle into coal chemical engineering and then become industrialization in the future, some improvements in energy store device, large capacity reactor as well as high efficient and environmental photocatalyst and electrode materials should be acquired.

7.4.4 Typical Cases Study on CO_2 Recycle in Energy Polygeneration System

7.4.4.1 A Coal-Based Polygeneration System with CO_2 as Gasifying Agent and Working Medium Recycle Use

Figure 7.19 depicts the flow diagram of the polygeneration system with CO_2 recycle and capture (PL-CRS) (Yi et al. 2013). A portion of the CO_2 is recycled into

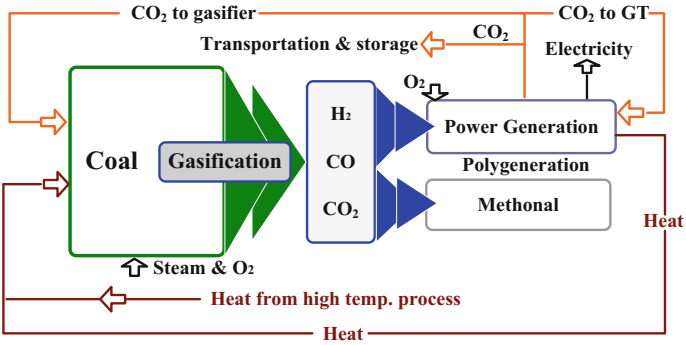

Fig. 7.19 Flow diagram of the PL-CRS system

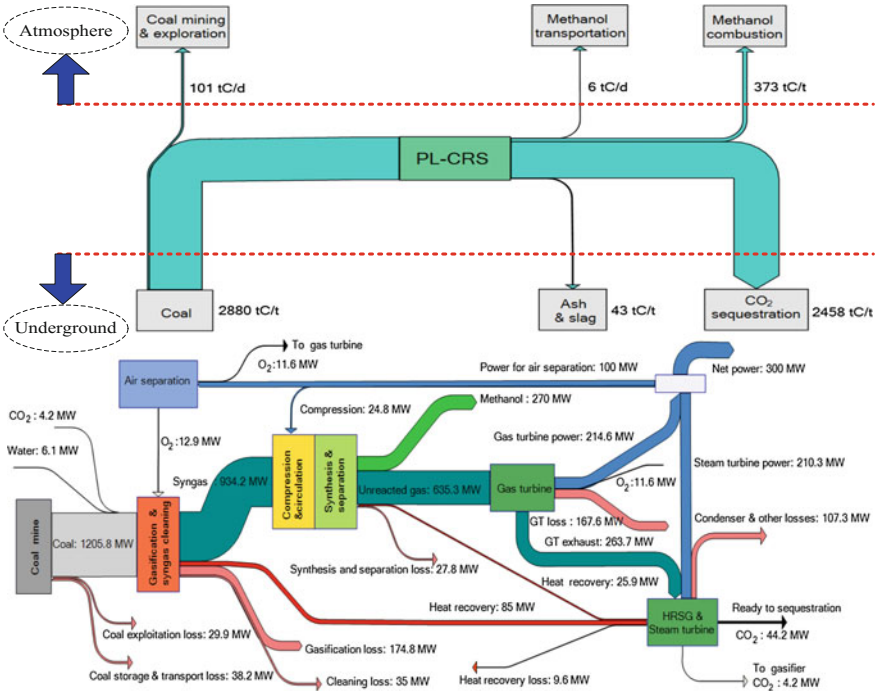

Fig. 7.20 Equivalent-carbon flow and energy flow chart for the PL-CRS

gasifier and gas turbine, while the rest is sequestered. The power generation system supplies the power and the steam that are required in the former processes. Syngas production can be enhanced by using of CO_2 recycling into the gasifier, and a portion of CO_2 can be used to replace H_2O as the gasification agent. Not only can water be conserved and carbon element utilization improved, but also the gasifier temperature can be controlled by a suitable ratio of CO_2/H_2O (kg/kg). At an equal

methanol output, the exergy efficiency of the system increases with CO_2 recycling, but, at a certain point, further CO_2 recycling will lead to a decrease in the system exergy efficiency, primarily because the compression and cycling of syngas, unreacted gas and CO_2 consumes a great deal of energy. The optimal ratio of CO_2 mass flow to coal mass flow-CO_2/coal (kg/kg) was determined to be 0.2, meaning that CO_2 that was equal to 20 % of the coal mass was recycled into the gasifier, while the remaining CO_2 undergoes sequestration. The design requirements of a certain gasifier for temperature and coal-water-slurry concentration (58–70 %) always limit CO_2/H_2O (kg/kg) ratio. It follows that a suitable ratio of recycling for CO_2 is a key to the whole system performance.

Figure 7.20 details the exergy distribution during various phases of the production process (Yi et al. 2013). The design scale for PL-CRS is 3×10^5 t-methanol/y and a net power output of 300 MW. The system experiences an exergy destruction of approximately 635.8 MW in total. The carbon flows for this system include carbon-equivalent GHG emission flows associated with activities upstream and downstream of the conversion facility. In this system, coal mining and exploration will cause equivalent carbon emissions of 101 t-C/d (ton per day) to be released into the atmosphere. As the gross flow rate of equivalent carbon of the entire system is 2981 t-C/d, and 2458 t-C/d is captured by sequestration, the system yields a net flow of 480 t-C/d of equivalent carbon emissions to the atmosphere, accounting for only 16.1 % of the carbon in the whole production process.

Performance comparison between single production systems and the PL-CRS system was found that the overall exergy efficiency of the novel systems is approximately 46.3 %, which is between 2–10 % higher than the individual systems. At the same output and CO_2 capture ratio (85 %), the energy saving ratio based on IGCC-CCS and single methanol plant (MP) (ESR1) of the PL-CRS system is 15.8 %, and the energy saving ratio based on IGCC-CRS and MP (ESR2) of the PL-CRS system is 11.1 %. What's more, in comparison with these two single production systems, the PL-CRS system achieves savings in capital cost of 30.4 % and 27.1 %, respectively. Furthermore, the cost of CO_2 avoidance of the PL-CRS is lower than the other systems by −18.93 $/(t-$CO_2$). The internal rate of return of PL-CRS (14.76 %) is lower than that of single methanol plant system (18.52 %). However, when the single methanol plant includes CO_2 capture and storage, the IRR will decrease to 11.12 %. As a result, taking into account the energy, CO_2 emission and economic performance for the different systems, the overall performance of PL-CRS is more favorable in all areas.

7.4.4.2 A Coal-Based Polygeneration System Integrating with CO_2 Recycle Reforming with CH_4-Rich Gas

Figure 7.21 is coal-based polygeneration system coupling with CO_2 recycle reforming with COG (D-PL-CR) (Yi et al. 2012b). Coal is gasified with steam and oxygen to produce coal gasified gas. After clean up, the clean gas is mixed with purified COG and recycled CO_2 before entering the CH_4/CO_2 reforming unit. The

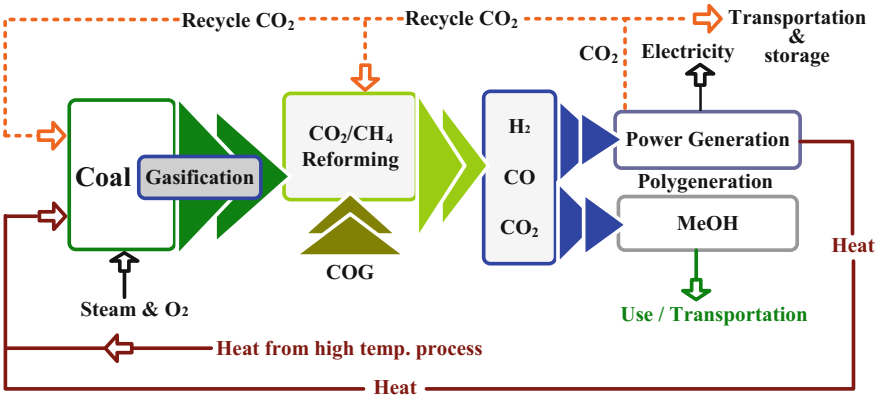

Fig. 7.21 Schematic of coal-based polygeneration system coupling with CO_2 recycles reforming with COG

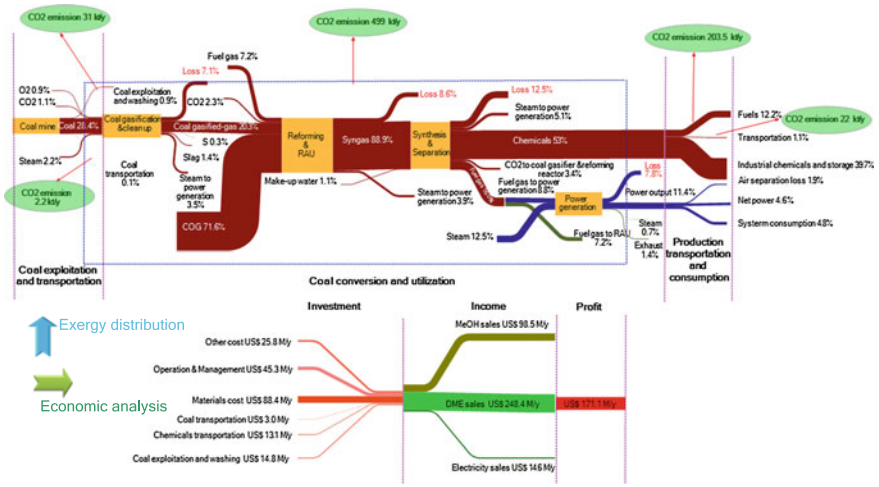

Fig. 7.22 The comprehensive performance of the D-PL-CR system

resulting clean syngas is sent for methanol synthesis. After processing in the synthesis reactor, CO_2, methanol (MeOH) and water are separated out from the unreacted gas. A portion of the separated CO_2 (95 wt%) is recycled back into the gasifier, and the remainder is used in the reforming unit.

The two CO_2 recycle technologies have different effects on the system. In order to find the best proportion of recycled CO_2 with these two technologies to satisfy the optimal performance of the whole system referring to energy, environment and economy we introduce another sensitivity variable λ (λ is the mass flow of CO_2 recycled to the reforming unit over the mass flow of all separated CO_2 from distillation unit) to investigate the effect of the CO_2 utilized proportion on the system. λ is

selected as 0.6 for the system study, which means that the recycling of 60 % CO_2 to reforming unit and 40 % to gasifier will bring better performance to the system.

Figure 7.22 showed the detailed exergy distribution and CO_2 emission during different production process, and the numbers in the Fig. express the proportion of the system energy input. The recycled system can reduce about 200 kt/y in the whole life cycle compared with the system without CO_2 recycle (Yi et al. 2012b). In the respect of energy use, great changes have been aroused due to the adoption of CO_2 recycles scheme. Because of CO_2 recycle, the exergy destruction of the recycled system in the reforming process and synthesis process increases by 1.4 % and 2.6 % respectively, while the combustion exergy destruction decreases from 12.7 to 7.8 %, the chemical exergy in chemical production increases from 41.5 to 53 %. The new system contributes 8.4 % power output from the system itself to realize the 11.5 % chemical energy increase, 33.8 % CO_2 emission reduction, and the total effective energy of the system increases from 54.5 to 57.6 %. As the CO_2 concentration in exhaust gas from the distillation tower is more than 95 wt%, this system doesn't require CO_2 separation unit. The increased chemical exergy (11.5 %) in chemical production and the exergy decrement (4.9 %) of combustion in gas turbine can almost offset most of the system exergy loss (4.9 %) and power output decrease (8.4 %) aroused by CO_2 recycles. So the system can keep thermal efficiency as high as that of no CO_2 recycle system. Additionally, the CO_2 recycled system shows good economy performance with IRR 18.1 %. In the CO_2 recycled system, the ratio of steam/water to coal was reduced by 0.10 kg/kg, which resulting in water saving as well as a reduction in the load of wastewater to be treated. According to the research results, for D-PL-CR, the system can conserve coal at a rate of 5.0×10^4 t/y and COG at a rate of 3.0×10^7 m^3/y, and reduce steam by 1.0×10^5 t/y and CO_2 by 1.3×10^5 t/y.

The system showed promising results from many of the component schemes, from CO_2 recycling to the gasifier and reforming unit, reducing the amount of steam required by gasification and realizing the conversion and utilization of CH_4 and CO_2, rendering the conventional water gas shift process unnecessary. This not only reduces required fuel input, waste water treatment load and greenhouse gas emissions, but also increases the utilization efficiency of carbon resources and energy. This technology, which realizes the integration among energy utilization, CO_2 emission and economy benefits, is highly instructive for development of new CO_2 emission control technology in energy system.

7.5 Challenges and Opportunities for Energy Polygeneration System with CO_2 Recycle

Actually, CO_2 is not just a greenhouse gas, but also an important source of carbon for making organic chemicals, materials, and carbohydrates. With CCS becoming key elements in worldwide efforts to control/minimize its emissions, it can be anticipated that large amounts of CO_2 from energy system will become available as

feedstock for innovative conversions to synthetic fuels or chemicals. As being discussed above, recycling of CO_2 in energy polygeneration system could be fruitful to produce various chemicals and fuels, which should be a sustainable way in the long term when renewable sources of energy such as biomass, solar, wind, wave and nuclear energy are used as energy input for the chemical processing.

However, each must overcome similar challenges before implementing this on an industrial scale. The challenges and the barriers for CO_2 conversion and utilization, as outlined below, are primarily due to the lack of value-based driving forces: (1) Costs of CO_2 capture, separation, purification, and transportation to user site; (2) Energy requirements of CO_2 chemical conversion (plus source and cost of H_2 and/or other co-reactants if involved). (3) Market size limitations, little investment-incentives and lack of industrial commitments for enhancing CO_2-based chemicals. The technologies relevant to CO_2 recycle technologies are still in laboratory research, pilot scale or demonstration stage, from which some knowledge and experience have been gained, however, uncertainties still exist. Especially, energy polygeneration system with CO_2 recycles that integrated different energy resources and production processes into one system are lack of design and operation experience. (4) Lack of socio-economical and policy driving forces for enhanced CO_2 utilization. Without support of the policy, CO_2 recycle technological research, market development and financial support are very hard to realize. Due to little pressure on carbon emission taxes, the enterprises have no will to reduce CO_2 emissions due to increase of investment. They will balance between paying emission taxes and investing in emission taxes, if emission taxes are set up. To develop successful CO_2 conversion projects, these technical and commercial barriers need to be overcome.

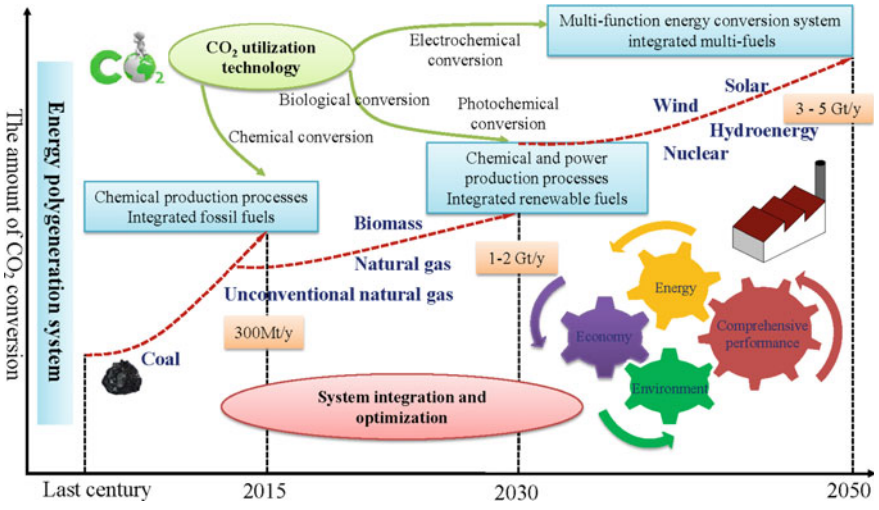

Fig. 7.23 The development roadmap of energy polygeneration system with CO_2 recycle

The possible development route of energy polygeneration system with CO_2 recycle is presented in Fig. 7.23. This technique roadmap is devised to serve as a blueprint describing development of CO_2 recycle in energy polygeneration system based on the technology maturity, with short and long-term vision of a low-carbon economy. In a short-to-medium term, CO_2 recycle use in chemical engineering mainly based on coal or fossil-C energy sources via chemical conversion of CO_2 due to its maturity, which has been investigated since last century. Within the same timeframe, CO_2 recycling can become an important component of the strategy portfolio necessary for curbing CO_2 emissions: with very optimistic assumptions, CO_2 recycle can offer a potential reduction equivalent of 250–350 million tons per year in the short-to-medium term, mostly driven by methanol, DME, and DMC syntheses etc. This corresponds to about 10 % of the total reduction required globally, and is at a level comparable to the expected impact of CCS technologies. Coal integrated biomass conversion to fuels is perhaps the most intensively pursued route, which is predicted to become extensively adopted in the near future, not only to mitigate CO_2 emissions, but also to save fossil fuel resource. More importantly, it is significant to integrate and optimize CO_2 recycle system based on fossil fuels and renewable energy sources, with combining chemical production and power generation, in which electrochemical and photochemical conversion of CO_2 are used to transform CO_2 into chemicals or fuels. After up scaling to commercial scale projects, the early commercial deployment should be immediately followed. Given that early CO_2 recycle projects will pave the way for large-scale deployment, and such projects will provide the early experience needed to facilitate CO_2 recycle for industrial sectors, it is important in this phase to continue to invest in research development and demonstration (RD&D) for both CO_2 conversion technologies and system integration and optimization, and to explore some early low-cost CO_2 recycle opportunities in simple polygeneration systems.

In the long term, CO_2 recycle could be proved to be an enabling technology for the emerging low-carbon economy in energy polygeneration system from previous demonstration. By that time, the MFEPS with advanced CO_2 conversion and recycle technologies integrated with different multi-fuels and production processes would show huge potential of CO_2 conversion and use, which was predicted to be 3-5 Gt/y, as CO_2 is defined as an excellent energy-vector which stores renewable energy (Centi and Perathoner 2011). And synthesis of chemicals and fuels from recycled CO_2 offers a valuable opportunity to introduce renewable energy into the existing energy and chemical infrastructure with the benefits of lowering the carbon footprint of the whole CO_2 conversion process. And it is a promising way to replace the fossil-derived fuels and simultaneously accelerate the adjustment of the current energy structure. Nevertheless, these bright prospects brought by CO_2 recycle depend on the technology and engineering breakthrough such as the high selectivity and activity catalyst, high CO_2 conversion efficiency, low production cost of hydrogen and system integration and optimization and so on.

Acknowledgments We thank Yi Huang, Minhui Gong and Guosheng Wu for contributions to the research described here. For funding in support of this work, the authors gratefully

acknowledge the financial support from the National Natural Science Foundation of China (51404164), the National Natural Science Foundation of China (51276120) and Shanxi Province Outstanding Youth Natural Science Foundation (201601D201004).

References

Abdollahi, G., & Meratizaman, M. (2011). Multi-objective approach in thermoenvironomic optimization of a small-scale distributed CCHP system with risk analysis. *Energy and Buildings, 43*, 3144–3153.

Abdollahi, G., & Sayyaadi, H. (2013). Application of the multi-objective optimization and risk analysis for the sizing of a residential small-scale CCHP system. *Energy and Buildings, 60*, 330–344.

Adams II, T. A., & Barton, P. I. (2011). Combining coal gasification and natural gas reforming for efficient polygeneration. *Fuel Processing Technology, 92*, 639–655.

Adams, T. A., & Ghouse, J. H. (2015). Polygeneration of fuels and chemicals. *Current Opinion in Chemical Engineering, 10*, 87–93.

Ahmadi, P., Dincer, I., & Rosen, M. A. (2014). Thermoeconomic multi-objective optimization of a novel biomass-based integrated energy system. *Energy, 68*, 958–970.

Aresta, M., & Dibenedetto, A. (2004). The contribution of the utilization option to reducing the CO_2 atmospheric loading: research needed to overcome existing barriers for a full exploitation of the potential of the CO_2 use. *Catalysis Today, 98*, 455–462.

Aresta, M., Dibenedetto, A., & Angelini, A. (2013). Catalysis for the valorization of exhaust carbon: from CO_2 to chemicals, materials, and fuels. Technological use of CO_2. *Chemical Reviews, 114*, 1709–1742.

Bermúdez, J. M., Ferrera-Lorenzo, N., Luque, S., Arenillas, A., & Menéndez, J. A. (2013). New process for producing methanol from coke oven gas by means of CO_2 reforming. Comparison with conventional process. *Fuel Processing Technology, 115*, 215–221.

Boretti, A. (2013). Renewable hydrogen to recycle CO_2 to methanol. *International Journal of Hydrogen Energy, 38*, 1806–1812.

Breault, R. W. (2010). Gasification processes old and new: A basic review of the major technologies. *Energies, 3*, 216–240.

Budzianowski, W. M. (2010). Thermal integration of combustion-based energy generators by heat recirculation. *Rynek Energii, 91*, 108–115.

Burns, R. K., Stagier, P. J., & Donovan, R. M. (1982). Integrated gasifier combined cycle polygeneration system to produce liquid hydrogen.

Cai, R., Jin, H., Gao, L., & Hong, H. (2010). Development of multifunctional energy systems (MESs). *Energy, 35*, 4375–4382.

Centi, G., & Perathoner, S. (2009). Catalysis: role and challenges for a sustainable energy. *Topics in Catalysis, 52*, 948–961.

Centi, G., & Perathoner, S. (2011). CO_2-based energy vectors for the storage of solar energy. *Greenhouse Gases: Science and Technology, 1*, 21–35.

Chaiwatanodom, P., Vivanpatarakij, S., & Assabumrungrat, S. (2014). Thermodynamic analysis of biomass gasification with CO_2 recycle for synthesis gas production. *Applied Energy, 114*, 10–17.

Cormos, C.-C. (2014). Techno-economic and environmental analysis of hydrogen and power co-generation based on co-gasification of coal and biomass/solid wastes with carbon capture. *Chemical Engineering, 37*, 139–144.

Dibenedetto, A., Angelini, A., & Stufano, P. (2014). Use of carbon dioxide as feedstock for chemicals and fuels: Homogeneous and heterogeneous catalysis. *Journal of Chemical Technology and Biotechnology, 89*, 334–353.

Dimopoulos, G. G., Stefanatos, I. C., & Kakalis, N. M. P. (2013). Exergy analysis and optimisation of a steam methane pre-reforming system. *Energy, 58*, 17–27.

Ebrahimi, M., Keshavarz, A., & Jamali, A. (2012). Energy and exergy analyses of a micro-steam CCHP cycle for a residential building. *Energy and Buildings, 45*, 202–210.

Fang, G., Tian, L., Fu, M., & Sun, M. (2013). The impacts of carbon tax on energy intensity and economic growth: A dynamic evolution analysis on the case of China. *Applied Energy, 110*, 17–28.

Gao, L. (2005). *Investigation of coal-based polygeneration systems for production of power and liquid fuel [D]*. Beijing: Chinese Academy of Sciences.

Goeppert, A., Czaun, M., Jones, J.-P., Prakash, G. S., & Olah, G. A. (2014). Recycling of carbon dioxide to methanol and derived products–closing the loop. *Chemical Society Reviews, 43*, 7995–8048.

Graves, C., Ebbesen, S. D., Mogensen, M., & Lackner, K. S. (2011). Sustainable hydrocarbon fuels by recycling $CO_2$and H_2O with renewable or nuclear energy. *Renewable and Sustainable Energy Reviews, 15*, 1–23.

Irfan, M. F., Usman, M. R., & Kusakabe, K. (2011). Coal gasification in CO_2 atmosphere and its kinetics since 1948: A brief review. *Energy, 36*, 12–40.

Itkulova, S. S., Zakumbaeva, G. D., Nurmakanov, Y. Y., Mukazhanova, A. A., & Yermaganbetova, A. K. (2014). Syngas production by bireforming of methane over Co-based alumina-supported catalysts. *Catalysis Today, 228*, 194–198.

Jabbari, B., Tahouni, N., Ataei, A., & Panjeshahi, M. H. (2013). Design and optimization of CCHP system incorporated into kraft process, using Pinch Analysis with pressure drop consideration. *Applied Thermal Engineering, 61*, 88–97.

Jenkins, J. D. (2014). Political economy constraints on carbon pricing policies: What are the implications for economic efficiency, environmental efficacy, and climate policy design? *Energy Policy, 69*, 467–477.

Jhong, H.-R. M., Ma, S., & Kenis, P. J. A. (2013). Electrochemical conversion of CO_2 to useful chemicals: Current status, remaining challenges, and future opportunities. *Current Opinion in Chemical Engineering, 2*, 191–199.

Jillson, K. R., Chapalamadugu, V., & Erik Ydstie, B. (2009). Inventory and flow control of the IGCC process with CO_2 recycles. *Journal of Process Control, 19*, 1470–1485.

Jing, Y.-Y., Bai, H., & Wang, J.-J. (2012). Multi-objective optimization design and operation strategy analysis of BCHP system based on life cycle assessment. *Energy, 37*, 405–416.

Kaggerud, K. H., Bolland, O., & Gundersen, T. (2006). Chemical and process integration: Synergies in co-production of power and chemicals from natural gas with CO_2 capture. *Applied Thermal Engineering, 26*, 1345–1352.

Kavvadias, K. C., & Maroulis, Z. B. (2010). Multi-objective optimization of a trigeneration plant. *Energy Policy, 38*, 945–954.

Kidoguchi, K., Hara, S., Oki, Y., Kajitani, S., Umemoto, S., & Inumaru, J. (2011). Development of oxy-fuel IGCC system with CO_2 recirculation for CO_2 capture: Experimental examination on effect of gasification reaction promotion by CO_2 enriched using bench scale gasifier facility. In *ASME 2011 Power Conference collocated with JSME ICOPE 2011* (pp. 485–492). American Society of Mechanical Engineers.

Li, S., Jin, H., Gao, L., & Zhang, X. (2014). Exergy analysis and the energy saving mechanism for coal to synthetic/substitute natural gas and power cogeneration system without and with CO_2 capture. *Applied Energy, 130*, 552–561.

Li, Y., Liao, S., & Liu, G. (2015). Thermo-economic multi-objective optimization for a solar-dish Brayton system using NSGA-II and decision making. *International Journal of Electrical Power and Energy Systems, 64*, 167–175.

Li, Z., Ni, W., Zheng, H., & Ma L. (2003). Polygeneration energy system based on coal gasification. *Energy for Sustainable Development, 7*, 57–62.

Lim, Y., Lee, C.-J., Jeong, Y. S., Song, I. H., Lee, C. J., & Han, C. (2012). Optimal design and decision for combined steam reforming process with dry methane reforming to reuse CO_2 as a raw material. *Industrial and Engineering Chemistry Research, 51*, 4982–4989.

Lin, R., Jin, H., & Gao, L. (2006). Chemical-power polygeneration system and its integrated optimization mechanism. *Journal of Engineering for Thermal Energy and Power, 21*, 331–337.

Lin, H., Jin, H., Gao, L., & Zhang, N. (2014). A polygeneration system for methanol and power production based on coke oven gas and coal gas with CO_2 recovery. *Energy, 74*, 174–180.

Liu, G., Larson, E. D., Williams, R. H., Kreutz, T. G., & Guo, X. (2010a). Making Fischer–Tropsch fuels and electricity from coal and biomass: Performance and cost analysis. *Energy and Fuels, 25*, 415–437.

Liu, P., Pistikopoulos, E. N., & Li, Z. (2010b). A multi-objective optimization approach to polygeneration energy systems design. *AIChE Journal, 56*, 1218–1234.

Lu, C., Tong, Q., & Liu, X. (2010). The impacts of carbon tax and complementary policies on Chinese economy. *Energy Policy, 38*, 7278–7285.

Luu, M. T., Milani, D., Bahadori, A., & Abbas, A. (2015). A comparative study of CO_2 utilization in methanol synthesis with various syngas production technologies. *Journal of CO_2 Utilization, 12,* 62–76.

Ma, L., Ni, W., Li, Z., & Ren, T. (2004). Analysis of the polygeneration system of methanol and electricity based on coal gasification (1). *Power Engineering, 24*, 451–456.

Man, Y., Yang, S., Xiang, D., Li, X., & Qian, Y. (2014a). Environmental impact and techno-economic analysis of the coal gasification process with/without CO_2 capture. *Journal of Cleaner Production, 71*, 59–66.

Man, Y., Yang, S., Zhang, J., & Qian, Y. (2014b). Conceptual design of coke-oven gas assisted coal to olefins process for high energy efficiency and low CO_2 emission. *Applied Energy, 133*, 197–205.

Meerman, J. C., Ramírez, A., Turkenburg, W. C., & Faaij, A. P. C. (2011). Performance of simulated flexible integrated gasification polygeneration facilities. Part A: A technical-energetic assessment. *Renewable and Sustainable Energy Reviews, 15*, 2563–2587.

Meerman, J. C., Ramírez, A., Turkenburg, W. C., & Faaij, A. P. C. (2012). Performance of simulated flexible integrated gasification polygeneration facilities, Part B: Economic evaluation. *Renewable and Sustainable Energy Reviews, 16*, 6083–6102.

Muresan, M., Cormos, C., & Agachi, P. (2014). Comparative life cycle analysis for gasification-based hydrogen production systems. *Journal of Renewable and Sustainable Energy, 6*, 013131.

Ni, W., Li, Z., & Xue, Y. (2000). Polygeneration energy system based on coal gasification integrated optimization and sustainable development of resources, energy and environment. *Engineering Sciences, 2*, 59–68.

Niu, S., Ding, Y., Niu, Y., Li, Y., & Luo, G. (2011). Economic growth, energy conservation and emissions reduction: A comparative analysis based on panel data for 8 Asian-Pacific countries. *Energy Policy, 39*, 2121–2131.

Olah, G. A., Goeppert, A., & Prakash, G. K. S. (2008). Chemical recycling of carbon dioxide to methanol and dimethyl ether: From greenhouse gas to renewable, environmentally carbon neutral fuels and synthetic hydrocarbons. *The Journal of Organic Chemistry, 74*, 487–498.

Olajire, A. A. (2013). Valorization of greenhouse carbon dioxide emissions into value-added products by catalytic processes. *Journal of CO_2 Utilization, 3–4,* 74–92.

Perathoner, S., & Centi, G. (2013). New energy sources and CO_2 treatment. *Global Change, Energy Issues and Regulation Policies.* Springer.

Perathoner, S., & Centi, G. (2014). CO_2 recycling: A key strategy to introduce green energy in the chemical production chain. *ChemSusChem, 7*, 1274–1282.

Peters, M., Köhler, B., Kuckshinrichs, W., Leitner, W., Markewitz, P., & Müller, T. E. (2011). Chemical technologies for exploiting and recycling carbon dioxide into the value chain. *ChemSusChem, 4*, 1216–1240.

Piatkowski, N., & Steinfeld, A. (2008). Solar-driven coal gasification in a thermally irradiated packed-bed reactor. *Energy and Fuels, 22*, 2043–2052.

Ratlamwala, T. A. H., Dincer, I., & Aydin, M. (2012). Energy and exergy analyses and optimization study of an integrated solar heliostat field system for hydrogen production. *International Journal of Hydrogen Energy, 37*, 18704–18712.

Ren, H., Zhou, W., Nakagami, K. I., Gao, W., & Wu, Q. (2010). Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects. *Applied Energy, 87*, 3642–3651.

Ruhul Kabir, M., & Kumar, A. (2012). Comparison of the energy and environmental performances of nine biomass/coal co-firing pathways. *Bioresource technology, 124,* 394–405.

Ruth, M. F., Zinaman, O. R., Antkowiak, M., Boardman, R. D., Cherry, R. S., & Bazilian, M. D. (2014). Nuclear-renewable hybrid energy systems: Opportunities, interconnections, and needs. *Energy Conversion and Management, 78*, 684–694.

Sahoo, U., Kumar, R., Pant, P., & Chaudhury, R. (2015). Scope and sustainability of hybrid solar–biomass power plant with cooling, desalination in polygeneration process in India. *Renewable and Sustainable Energy Reviews, 51*, 304–316.

Serra, L. M., Lozano, M.-A., Ramos, J., Ensinas, A. V., & Nebra, S. A. (2009). Polygeneration and efficient use of natural resources. *Energy, 34*, 575–586.

Song, C., & Pan, W. (2004). Tri-reforming of methane: A novel concept for catalytic production of industrially useful synthesis gas with desired H_2/CO ratios. *Catalysis Today, 98*, 463–484.

Tahir, M., & Amin, N. S. (2013). Recycling of carbon dioxide to renewable fuels by photocatalysis: Prospects and challenges. *Renewable and Sustainable Energy Reviews, 25*, 560–579.

Tan, Z.-F., Zhang, H.-J., Shi, Q.-S., Song, Y.-H., & Ju, L.-W. (2014). Multi-objective operation optimization and evaluation of large-scale NG distributed energy system driven by gas-steam combined cycle in China. *Energy and Buildings, 76*, 572–587.

Theofanidis, S. A., Batchu, R., Galvita, V. V., Poelman, H., & Marin, G. B. (2016). Carbon gasification from Fe–Ni catalysts after methane dry reforming. *Applied Catalysis, B: Environmental, 185*, 42–55.

Trop, P., Anicic, B., & Goricanec, D. (2014). Production of methanol from a mixture of torrefied biomass and coal. *Energy, 77*, 125–132.

Wang, Q. (2001). "Vision 21" Coal based energy plant: Co-production, high efficiency, zero-emission. *China Coal, 9*, 5–8.

Wang, L. (2006). *Coal energy industry ecology*. Beijing: Chemical Industry Press.

Wang, J.-J., Jing, Y.-Y., Zhang, C.-F., & Zhao, J.-H. (2009). Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and Sustainable Energy Reviews, 13*, 2263–2278.

Wang, W., Han, H., Yuan, M., Li, H., Fang, F., & Wang, K. (2011). Treatment of coal gasification wastewater by a two-continuous UASB system with step-feed for COD and phenols removal. *Bioresource technology, 102*, 5454–5460.

Wang, L., Yang, Y., Dong, C., Morosuk, T., & Tsatsaronis, G. (2014). Multi-objective optimization of coal-fired power plants using differential evolution. *Applied Energy, 115*, 254–264.

Williams, R. H., Larson, E. D., Liu, G., & Kreutz, T. G. (2009). Fischer-Tropsch fuels from coal and biomass: Strategic advantages of once-through ("polygeneration") configurations. *Energy Procedia, 1*, 4379–4386.

Wong, A. (2012). Comparative economics of using biomass and natural gas for co-generation of steam and power for prospective industrial enterprise in belarus. In *11th International Scientific Conference: Engineering for rural development*, *Jelgava, Latvia*, *24–25 May, 2012*. Latvia University of Agriculture.

Wu, J.-Y., Wang, J.-L., & Li, S. (2012a). Multi-objective optimal operation strategy study of micro-CCHP system. *Energy, 48*, 472–483.

Wu, J., Wang, J., Cheng, Q., & Devallance, D. (2012b). Assessment of coal and biomass to liquid fuels in central Appalachia, USA. *International Journal of Energy Research, 36*, 856–870.

Xiao, Y. (2008). Coal gasification based coproduction technology innovation. *China Coal, 34*, 11–15.

Xie, K., Li, W., & Zhao, W. (2010). Coal chemical industry and its sustainable development in China. *Energy, 35*, 4349–4355.

Yang, S., Yang, Q., Man, Y., Xiang, D., & Qian, Y. (2013). Conceptual design and analysis of a natural gas assisted coal-to-olefins process for CO_2 reuse. *Industrial and Engineering Chemistry Research, 52*, 14406–14414.

Yi, Q., Feng, J., & Li, W. Y. (2012a). Optimization and efficiency analysis of polygeneration system with coke-oven gas and coal gasified gas by Aspen Plus. *Fuel, 96*, 131–140.

Yi, Q., Lu, B., Feng, J., Wu, Y., & Li, W. (2012b). Evaluation of newly designed polygeneration system with CO_2 recycle. *Energy and Fuels, 26*, 1459–1469.

Yi, Q., Fan, Y., Li, W., & Feng, J. (2013). CO_2 capture and use in a novel coal-based polygeneration system. *Industrial and Engineering Chemistry Research, 52*, 14231–14240.

Yi, Q., Feng, J., Wu, Y., & Li, W. (2014). 3E (energy, environmental, and economy) evaluation and assessment to an innovative dual-gas polygeneration system. *Energy, 66*, 285–294.

Yuan, Z., & Chen, B. (2012). Process synthesis for addressing the sustainable energy systems and environmental issues. *AIChE Journal, 58*, 3370–3389.

Zangeneh, F. T., Sahebdelfar, S., & Ravanchi, M. T. (2011). Conversion of carbon dioxide to valuable petrochemicals: An approach to clean development mechanism. *Journal of Natural Gas Chemistry, 20*, 219–231.

Zhou, L., Hu, S., Li, Y., & Zhou, Q. (2008). Study on co-feed and co-production system based on coal and natural gas for producing DME and electricity. *Chemical Engineering Journal, 136*, 31–40.

Zhou, Z., Liu, P., Zhang, J. Y., & Li, Z. (2012). Evaluating the impact of carbon taxes on the optimal design of distributed energy systems. *Advanced Materials Research, 524*, 2420–2424.

Part II
Power and Transport Systems

Chapter 8
Evaluating the Contribution of Energy Storages to Support Renewable Integrations

Qixin Chen, Peng Zou, Qing Xia and Chongqing Kang

Abstract Energy storage systems (ESSs) are one kind of advances in energy systems engineering and of great value to realize energy management and to support renewable generation. The combined operation of ESSs and renewables is one way to achieve output levelling and to improve the integration of sustainable energy. However, in a market-based environment, ESSs would make strategic decisions on self-schedules and arbitrage in energy and ancillary service markets, maximizing the overall profits. Will the strategic operation of ESSs promote renewable generation integration? To explicitly answer this question, this chapter proposes a multi-period Nash-Cournot equilibrium model for joint energy and ancillary service markets to evaluate the contribution of the ESSs for supporting renewable generation. Then, a reformulation approach based on the potential function is proposed, which can transform the bi-level equilibrium model into an integrated single-level optimization problem to enhance the computation efficiency. Numerical examples are implemented to validate the effectiveness of the reformulation technique. The results of the case study indicate that the ESSs indirectly but substantially provide improved flexibilities while pursuing individual profit maximization.

Recent past years have witnessed a continuously changing generation mix around the world. The proportion of renewables is gradually growing, and the conventional fossil energy is undergoing a declining share in the primary inputs to power generation (The BP Energy Outlook 2015). The ever-increasing wind farms and solar stations produce more clean and eco-friendly power, but put great pressure on the operation of power systems and the electricity markets due to their intermittent and stochastic nature (Woo et al. 2013; Jiang et al. 2013). Meanwhile, emerging energy storage systems (ESSs) are tentatively used to provide a practical solution for more flexibility and to prevent renewable energy from being curtailed. Therefore, several

Q. Chen (✉) · P. Zou · Q. Xia · C. Kang
Department of Electrical Engineering, State Key Lab of Power Systems,
Tsinghua University, Beijing 100084, China
e-mail: qxchen@tsinghua.edu.cn

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_8

papers have concentrated on the combined operation of ESSs and renewable generations in power markets, which are always price-takers, to make the systems more controllable and to obtain more profits.

Maria et al. (2012) analyzes the combined planning and operation of wind power and ESSs in the electricity market. Based on the forecasted prices, a procedure is proposed to determine the hourly production profile of this combined system. Usaola (2011) proposes a technology that combines the concentrating solar power and the ESS as dispatch-able to smooth the production profile and obtain more profit in the spot energy market. Ali et al. (2011) proposes a strategy for the integrated operation of a wind farm and pumped-hydro storage to increase the benefits from both day-ahead energy and ancillary service markets. Thatte et al. (2013) uses a risk measure-based, robust bidding strategy for a wind farm in combination with ESS in the day-ahead energy market. An optimal bidding strategy is considered for independently operated ESSs participating in the energy and reserve markets with high penetration of wind power based on a stochastic programming approach (Akhavan-Hejazi and Mohsenian-Rad 2014). However, the ESSs are regarded as price-takers, and the effects of ESSs to support renewables integration are not discussed in this chapter.

In fact, in a practical market-based environment, there will be increasing independently invested ESSs and renewable generators that belong to different companies. In this case, the ESSs tend to make strategic decisions on self-schedules and arbitrage in energy and ancillary service markets, maximizing their own profits. Will the strategic operation and behaviors of ESSs still promote renewable generation integration? Specifically, will the ESSs spontaneously or indirectly facilitate more renewable energy to be accommodated while maximizing individual profits in joint energy and ancillary service (AS) markets? Why and how do these things happen? What factors will influence the strategic behaviors of ESSs and the effects of renewable generation accommodation?

To answer the above questions, this chapter establishes a multi-period Nash-Cournot equilibrium model for joint energy and AS markets to accurately study the impacts of the strategic behaviors of independently operated ESSs on the integration of large-scale renewable generations. Multi-scenario settings of wind generation profile and a deterministic model for each scenario are applied to reflect the stochastic nature of the renewables (Wang et al. 2013; Aigner et al. 2012; Yousefi et al. 2013).

In fact, the Cournot model and the supply function model both have their advantages and disadvantages and which one is more appropriate depends on the topic and the purpose of the research. The Cournot competition model may seem less attractive than the supply function equilibrium model in the aspect of reflecting the participants' real strategic bidding behaviors. However, according to (Klemperer and Meyer 1989), when facing certain demand distributions, a market participant would exactly know its equilibrium residual demand as a result and it has only one choice to maximize its profit, that is, it has no difference to choose either a

strategic price or a strategic quantity. In other words, there is no difference to use either the Cournot equilibrium model or the supply function equilibrium model when the uncertainties of demand or renewables are ignored in the model. Although the bidding prices cannot be obtained by using the Cournot model, considering the advantages in computation efficiency, it is still widely used in numerous applications (Li et al. 2011; Willems et al. 2009), including the market power analysis (Tamaschke et al. 2005), hydrothermal coordination (Molina et al. 2011), transmission congestion (Jian et al. 2007) and co-optimization of energy and ancillary service markets (Deb 2004).

At present, two traditional way to obtain the Nash-Cournot equilibrium is to simultaneously solve the multi-individual profit-maximization problems with either a nonlinear complementarity problem (NCP) approach or diagonalization methods (Gabriel et al. 2013). Specifically, the diagonalization methods consist of the Jacobi and Guass-Seidel (GS) algorithms, which iteratively solve each generator's profit-maximization model until a stationary point is obtained. The NCP method collects all of the optimal KKT conditions of each generator's individual profit-maximization models and then solves them together. However, these two conventional methods are computation intractable and lack robustness (Carlos et al. 2012). These issues are studied in this chapter.

Overall, the main contributions of this chapter are as follows:

(1) A multi-period optimization model is constructed to study the Nash-Cournot equilibrium and the strategic interactions among various generators, including the ESSs, thermal units, hydro units, large-scale wind farms and solar stations in joint energy and AS markets. The pay-for-performance mechanism is also considered in the regulation market.
(2) The ESSs are considered as price-makers, considering their remarkable potential in the future, instead of price-takers to precisely describe their strategic behaviors in the energy market.
(3) A novel reformulation technique supported by introducing the potential function is proposed to transform the conventional Nash-Cournot equilibrium model into an integrated single-level optimization problem, to improve the computation efficiency and identify more optimal solutions.

Specifically, the remainder of this chapter is organized as follows. The operation pattern of joint energy and AS markets is briefly introduced in Sect. 8.1. The Nash-Cournot model is derived, and the multi-individual optimization problems are obtained in Sect. 8.2. The reformulated single-level optimization model supported by the potential function is deduced in Sect. 8.3. Numerical examples are tested to validate the reformulation approach and to evaluate the contribution of ESSs for supporting large-scale renewable generation in Sect. 8.4. Finally, the conclusions are summarized in Sect. 8.5.

8.1 Co-optimization of Energy and Ancillary Service Markets

Without loss of generality and rationality, multi-period energy and AS markets with one shot game are considered in this chapter. Just like the rules implemented in the electricity markets in North America, renewable generators, including wind farms and solar stations, are only permitted to participate in the energy market because of their stochastic outputs, whereas ESSs have the rights to enter the three markets at the same time, as do the thermal power and conventional hydropower.

8.1.1 Energy Market

In the pool-based energy market, it is assumed that various generators compete in the Cournot manner with complete information, that is, each generator strategically determines its energy quantity to produce, considering the possible behaviors of r3als, and submits this offer to the market operator (MO) before the gate closure. The MO clears the energy market considering the power supply-demand balance constraints and calculates the time-varying marginal clearing price (MCP) to make settlements based on the linear inverse demand function received from the consumer side.

8.1.2 Regulation and Reserve Markets

Because the regulation and reserve services are mainly established to guarantee secure operation of power systems, the AS demands are usually not considered as price-sensitive but are dependent on the system conditions. In this chapter, the regulation and reserve markets are reasonably assumed to have abundant resources and to be relatively competitive. In this case, participants in the AS markets are regarded as price-takers for the purpose of simplicity and the regulation and reserve prices are assumed to be constants. The reserve market implements capacity payments, whereas the performance-based regulation market includes not only capacity payments, but also performance payments to ensure the fast response resources, such as hydropower and ESSs, more incentives and compensation. This mechanism is referenced from the “pay for performance” scheme in the regulation market of PJM (Xiao et al. 2014).

8.1.3 Co-optimization of Energy and AS Markets

When generators are simultaneously providing energy and AS services, the capacity coupling effects of generators should be precisely considered to achieve the optimization of power resource allocation. Therefore, the co-optimization of energy and AS markets is implemented in several major electricity markets operated by the independent system operators (ISO) or regional transmission organizations (RTO) in North America, such as the markets administrated by PJM (short for Pennsylvania, New Jersey and Maryland), ERCOT (short for Electric Reliability Council of Texas) and CAISO (short for California ISO). Accordingly, the objective of the MO is extended to minimize the total production costs of the energy and AS services. Various generators' technical constraints, especially their capacity coupling constraints for providing energy and AS services, are included to jointly clear the multi-markets.

8.2 Modeling of the Nash-Cournot Equilibrium in Joint Energy and AS Markets

The Nash-Cournot equilibrium can be formulated as a bi-level optimization model. The upper-level model represents the individual profit-maximization problems of various participants, whereas the lower-level model indicates the clearing conditions of the energy and AS markets, which are shared by each generator. Then, the multi-individual optimization problems can be obtained by substituting the common lower-level model into each upper-level model. The multi-individual optimization problems need to be simultaneously solved to get the Nash-Cournot equilibrium. In this chapter, the potential function is introduced to transform the multi-individual optimization problems into an integrated single-level optimization problem to improve the computation efficiency. The specific flowchart of the Nash-Cournot equilibrium modeling is presented as Fig. 8.1.

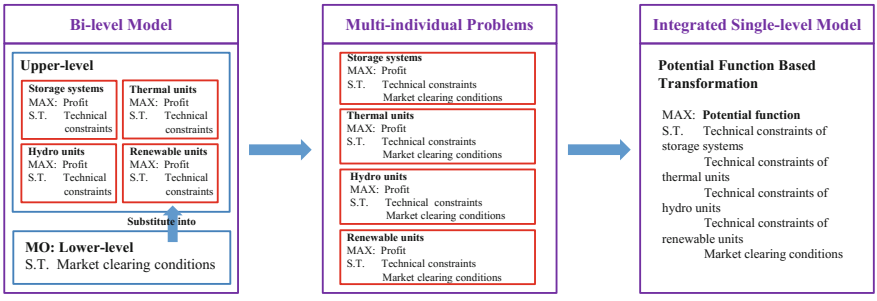

Fig. 8.1 The flowchart of the Nash-Cournot equilibrium modeling

It should be noted here that there is no objective function in the lower-level model as the classic Cournot model does (Drew and Jean 2013). But the equilibrium model presented in this chapter is still called bi-level optimization model because it has the same mathematical structure as bi-level models.

8.2.1 Profit Maximization Models for Various Generators

The objective functions of thermal units, hydro units and ESSs are to maximize the payments from both the energy and AS markets, whereas for wind and solar power, the revenues are only from the energy market. The explicit variable costs of all generators, except thermal power, are reasonably assumed as zero (Energy Information Administration 2013). The operation and maintenance costs of ESSs have been implicitly considered in the form of round-trip efficiencies.

(1) *Energy Storage Systems*

The objective function of the individual profit-maximization model of ESSs is:

$$\begin{aligned}&\max \sum_{t=1}^{T} \pi_{Si}(t) \\ &= \max \sum_{t=1}^{T} \begin{cases} \lambda^E(t) q_{Si}^E(t) + \lambda^R(t) q_{Si}^R(t) \\ + \lambda_C^F(t) q_{Si}^F(t) + \theta_{Si} \lambda_P^F(t) q_{Si}^F(t) \end{cases}\end{aligned} \tag{8.1a}$$

where $\lambda^E(t)$ is the MCP of the energy market in period t, $\lambda^R(t)$ is the clearing price of the reserve market in period t, $\lambda_C^F(t)$ is the capability price of the regulation market in period t, $\lambda_P^F(t)$ is the performance price of the regulation market in period t, θ_{Si} is the regulation mileage ratio of ESSs, and $q_{Si}^E(t)$, $q_{Si}^R(t)$ and $q_{Si}^F(t)$ are, respectively, the generation output, reserve capacity and regulation capacity determined by the ESS Si in period t.

The charging/discharging states determination equations are: $\forall t \in T$,

$$\begin{cases} q_{Si}^E(t) = q_{Si}^{dis}(t) + q_{Si}^{cha}(t) \\ 0 \le q_{Si}^{dis}(t) \le b_{Si}^{dis}(t) q_{Si\max} \\ -b_{Si}^{cha}(t) q_{Si\max} \le q_{Si}^{cha}(t) \le 0 \\ b_{Si}^{dis}(t) + b_{Si}^{cha}(t) = 1 \\ b_{Si}^{dis}(t), b_{Si}^{cha}(t) \in \{0, 1\} \end{cases} \tag{8.1b}$$

where $q_{Si}^{dis}(t)$ and $q_{Si}^{cha}(t)$ are, respectively, the discharging and charging power of the ESS *Si*, $q_{Si\max}$ is the maximum capacity limit, and $b_{Si}^{dis}(t)$ and $b_{Si}^{cha}(t)$ are the binary variables representing the working status of ESSs.

The capacity constraints of ESS *Si* are: $\forall t \in T$,

$$\begin{cases} q_{Si}^{E}(t) + q_{Si}^{R}(t) + q_{Si}^{F}(t) \le q_{Si\max} \\ q_{Si}^{E}(t) - q_{Si}^{F}(t) \ge -q_{Si\max} \\ q_{Si}^{R}(t), q_{Si}^{F}(t) \ge 0 \end{cases} \tag{8.1c}$$

The available energy constraints necessarily kept to guarantee the continuous provision of regulation and reserve services for a certain period are: $\forall t \in T$,

$$\begin{cases} q_{Si}^{E}(t) \cdot h + q_{Si}^{R}(t) \cdot \frac{h}{2} + q_{Si}^{F}(t) \cdot \frac{h}{4} \le E_{Si}(t) \\ E_{Si}(t) - q_{Si}^{E}(t) \cdot h + q_{Si}^{F}(t) \cdot \frac{h}{4} \le E_{Si\max} \end{cases} \tag{8.1d}$$

where h equals one hour here, $E_{Si}(t)$ is the stored energy in period t, and $E_{Si\max}$ is the maximum energy that the ESS can store. The first constraint of (8.1d) means that the ESS in period t can continuously discharge for one hour at the $q_{Si}^{E}(t)$ level, provide reserve service for 30 min at the $q_{Si}^{R}(t)$ level and regulate for 15 min at the $q_{Si}^{F}(t)$ level. These coefficients are referenced from PJM markets and related papers (Masiello et al. 2014; He et al. 2015). Obviously, the longer the duration for providing the AS services means less regulation and reserve capacities that the ESS can bid considering its limited stored energy.

The energy conversion relations are as below: $\forall t \in T$,

$$\begin{cases} E_{Si}|_{end} - E_{Si}|_{Init} = 0 \\ E_{Si}(t+1) = E_{Si}(t) - q_{Si}^{dis}(t)/\eta_{Si}^{dis} - \eta_{Si}^{cha} q_{Si}^{cha}(t) \end{cases} \tag{8.1e}$$

where η_{Si}^{dis} and η_{Si}^{cha} are the efficiencies of discharging and charging. The first equation of (8.1e) ensures that the final state of charge equals the initial state of charge on a daily cycle.

The self-discharge of ESS is ignored in the model because many kinds of storages usually have no daily self-charge, such as pumped hydro storage, hydrogen-based energy storage system, compressed air energy storage, zinc-bromine flow battery, polysulphide-bromide flow battery, sodium-sulphur battery (Díaz-González et al. 2012), and this assumption does not imposes significant effects on the results.

(2) *Thermal Unit*

The objective function of the individual profit-maximization model of thermal unit is formulated as the following:

$$\begin{aligned} &\max \sum_{t=1}^{T} \pi_{Ti}(t) \\ &= \max \sum_{t=1}^{T} \begin{cases} \lambda^{E}(t) q_{Ti}^{E}(t) + \lambda^{R}(t) q_{Ti}^{R}(t) + \lambda_{C}^{F}(t) q_{Ti}^{F}(t) \\ + \theta_{Ti} \lambda_{P}^{F}(t) q_{Ti}^{F}(t) - \{a_i [q_{Ti}^{E}(t)]^2 + b_i q_{Ti}^{E}(t) + c_i\} \end{cases} \end{aligned} \tag{8.1f}$$

where θ_{Ti} is the regulation mileage ratio of thermal power, $q_{Ti}^{E}(t)$, $q_{Ti}^{R}(t)$ and $q_{Ti}^{F}(t)$ are, respectively, the generation output, reserve capacity and regulation capacity, that is, the decision variables determined by thermal unit Ti in period t, and $a_i, b_i, c_i \geq 0$ are the coefficients of the total generation cost.

The capacity constraints of the thermal units are: $\forall t \in T$,

$$\begin{cases} q_{Ti}^{E}(t) + q_{Ti}^{R}(t) + q_{Ti}^{F}(t) \leq q_{Ti\max} \\ q_{Ti}^{E}(t) - q_{Ti}^{F}(t) \geq q_{Ti\min} \\ q_{Ti}^{E}(t), q_{Ti}^{R}(t), q_{Ti}^{F}(t) \geq 0 \end{cases} \tag{8.1g}$$

where $q_{Ti\min}$ and $q_{Ti\max}$ are the minimum and maximum generation output limits.

The ramp constraints of thermal unit are: $\forall t \in T$,

$$\begin{cases} [q_{Ti}^{E}(t+1) + q_{Ti}^{R}(t+1) + q_{Ti}^{F}(t+1)] - [q_{Ti}^{E}(t) - q_{Ti}^{F}(t)] \leq q_{Ti}^{RMPU} \\ [q_{Ti}^{E}(t) + q_{Ti}^{R}(t) + q_{Ti}^{F}(t)] - [q_{Ti}^{E}(t+1) - q_{Ti}^{F}(t+1)] \leq q_{Ti}^{RMPD} \end{cases} \tag{8.1h}$$

where q_{Ti}^{RMPU} and q_{Ti}^{RMPD} are the ramp up and down rate limits of thermal unit Ti.

The response time constraints of the reserve and regulation services are: $\forall t \in T$,

$$\begin{cases} q_{Ti}^{R}(t) \leq q_{Ti}^{RMPU} \cdot t^{R} \\ q_{Ti}^{F}(t) \leq q_{Ti}^{RMPU} \cdot t^{F} \end{cases} \tag{8.1i}$$

where t^R is the spinning reserve response time limit, which is assumed to be 10 min, and t^F is the regulation response time limit, which is assumed to be 5 min in this chapter.

(3) *Hydropower*

The objective function of the individual profit-maximization model of hydro unit is:

$$\begin{aligned} &\max \sum_{t=1}^{T} \pi_{Hi}(t) \\ &= \max \sum_{t=1}^{T} \begin{cases} \lambda^{E}(t) q_{Hi}^{E}(t) + \lambda^{R}(t) q_{Hi}^{R}(t) \\ + \lambda_{C}^{F}(t) q_{Hi}^{F}(t) + \theta_{Hi} \lambda_{P}^{F}(t) q_{Hi}^{F}(t) \end{cases} \end{aligned} \tag{8.1j}$$

where θ_{Hi} is the regulation mileage ratio of hydropower, and $q_{Hi}^{E}(t)$,$q_{Hi}^{R}(t)$ and $q_{Hi}^{F}(t)$ are, respectively, the generation output, reserve capacity and regulation capacity determined by hydro unit Hi in period t.

The capacity constraints of hydropower are: $\forall t \in T$,

$$\begin{cases} q_{Hi}^{E}(t) + q_{Hi}^{E}(t) + q_{Hi}^{F}(t) \leq q_{Hi\max} \\ q_{Hi}^{E}(t) - q_{Hi}^{F}(t) \geq q_{Hi\min} \\ q_{Hi}^{E}(t), q_{Hi}^{E}(t), q_{Hi}^{F}(t) \geq 0 \end{cases} \tag{8.1k}$$

where $q_{Hi\max}$ and $q_{Ti\max}$ are the minimum and maximum generation output limits.

The available energy constraint considering the water resource limit is:

$$\sum_{t=1}^{T} q_{Hi}^{E}(t) \le E_{Hi\max} \tag{8.1l}$$

where $E_{Hi\max}$ is the maximum energy hydropower can provide in an entire day.

The constraints that guarantee the continuous provision of regulation and reserve services for a certain duration are ignored here for simplicity. Moreover, the response time constraints of the reserve and regulation services are also ignored considering the high flexibility of hydro units.

(4) *Renewable Unit*

The objective function of the individual profit-maximization model of wind farms or solar stations is to maximize the payments only from the energy market:

$$\max \sum_{t=1}^{T} \pi_{Ri}(t) = \sum_{t=1}^{T} \lambda^{E}(t) q_{Ri}^{E}(t) \tag{8.1m}$$

where $q_{Ri}^{E}(t)$ is the generation output decided by the renewables Ri in period t.

The time-varying weather-dependent capacity constraints of renewable energy are: $\forall t \in T$,

$$q_{Ri\min}(t) \le q_{Ri}^{E}(t) \le q_{Ri\max}(t) \tag{8.1n}$$

where $q_{Ri\max}(t)$ and $q_{Ri\min}(t)$ are the minimum and maximum generation output limits.

8.2.2 Market Clearing Conditions

The energy market clearing conditions incorporate the multi-period power balance constraints and the linear inverse demand function. For $t = 1, 2, \ldots T$, we have:

$$\begin{cases} \sum\limits_{i}^{N} q_{Xi}^{E}(t) = \sum\limits_{i=1}^{TP} q_{Ti}^{E}(t) + \sum\limits_{i=1}^{HP} q_{Hi}^{E}(t) \\ \quad + \sum\limits_{i=1}^{RE} q_{Ri}^{E}(t) + \sum\limits_{i=1}^{ES} q_{Si}^{E}(t) \\ \sum\limits_{i}^{N} q_{Xi}^{E}(t) = q_{D}(t) \\ \lambda^{E}(t) = \alpha(t) - \beta(t) q_{D}(t) \end{cases} \tag{8.2a}$$

where $N = TP + HP + RE + ES$ is the total number of various generators participating in the energy market, $q_{Xi}^{E}(t)$ represents a certain kind of generation units,

$q_D(t)$ is the electricity demand in period t, and $\alpha(t)$ and $\beta(t)$ are the coefficients of the inverse demand function in period t. Based on the decreasing inverse demand function, the generators collectively and strategically determine the energy prices considering the possible reactions of the rivals to maximize its individual profit.

The reserve market clearing conditions are the multi-period reserve supply-demand balance constraints:

$$q_D^R(t) = \sum_{i=1}^{TP} q_{Ti}^R(t) + \sum_{i=1}^{HP} q_{Hi}^R(t) + \sum_{i=1}^{ES} q_{Si}^R(t) \tag{8.2b}$$

where $q_D^R(t)$ is the reserve demand in period t, which is determined by three parts: the load reserve, emergency reserve and the extra reserve caused by the stochastic outputs of renewable power.

The regulation market clearing conditions are the multi-period regulation supply-demand balance constraints:

$$q_D^F(t) = \sum_{i=1}^{TP} q_{Ti}^F(t) + \sum_{i=1}^{HP} q_{Hi}^F(t) + \sum_{i=1}^{ES} q_{Si}^F(t) \tag{8.2c}$$

where $q_D^F(t)$ is the regulation demand in period t, which is often a certain ratio of the peak load, for example, 0.7 % in the PJM regulation market (Xiao et al. 2014).

Considering that the contributions of ESSs to support large-scale renewables are mainly studied from the system-wide perspective, the transmission constraints are not included for simplicity.

8.2.3 Multi-individual Optimization Problems

By substituting the energy market clearing conditions (8.2a) into the objective functions of the individual profit-maximization models (8.1a), (8.1f), (8.1j) and (8.1m) for each generator, respectively, the compact objective functions can be expressed as:

$$\begin{aligned} &\max_{Xi=Ti,Hi,Ri,Si} \sum_{t=1}^{T} \pi_{Xi}(t) \\ &= \max \sum_{t=1}^{T} \left\{ \begin{array}{l} [\alpha(t) - \beta(t) \sum_{i=1}^{N} q_{Xi}^E(t)] q_{Xi}^E(t) + \lambda^R(t) q_{Xi}^R(t) \\ + \lambda_C^F(t) q_{Xi}^F(t) + \theta_{Xi} \lambda_P^F(t) q_{Xi}^F(t) - C_{Xi}[q_{Xi}^E(t)] \end{array} \right. \end{aligned} \tag{8.2d}$$

where the generation cost function $C_{Xi}[q_{Xi}^E(t)]$ is as below: $\forall t \in T$,

$$C_{Xi}[q_{Xi}^{E}(t)] = \begin{cases} a_i[q_{Ti}^{E}(t)]^2 + b_i q_{Ti}^{E}(t) + c_i, & \text{for } Xi = Ti \\ 0, & \text{for } Xi = Hi, Ri \text{ or } Si \end{cases} \tag{8.2e}$$

Moreover, the renewables do not participate in the AS markets; thus, the $q_{Xi}^{R}(t)$ and $q_{Xi}^{F}(t)$ in (8.2d) for wind farms and solar stations are always equal to zero.

Thus, the multi-individual profit-maximization model of each generator can be obtained by incorporating its own technical constraints, that is, (8.1b), (8.1c), (8.1d) and (8.1e) for the ESSs; (8.1g), (8.1h) and (8.1i) for thermal power; (8.1k) and (8.1l) for hydropower; and (8.1n) for the renewables. In addition, the reserve and regulation supply-demand balance conditions are also needed as the extra constraints shared by each generator.

8.3 Reformulation

By introducing the potential function, the conventional Nash-Cournot equilibrium model with multi-individual optimization problems can be equivalently transformed into an integrated single-level optimization model. The optimal results of this new model are strictly equal to the original equilibrium point.

8.3.1 Principle of the Potential Function

If $\forall i \in N$, the objective functions (8.2d) are continuously differentiable. Then, $\hat{P}[q_{Xi}^{E}(t), q_{Xi}^{R}(t), q_{Xi}^{F}(t)]$ is the potential function, if and only if (Kulkarni and Shanbhag 2014; Monderer and Shapley 1996; Voorneveld 2000): $\forall t \in T$,

$$\begin{cases} \dfrac{\partial \hat{P}[q_{Xi}^{E}(t), q_{Xi}^{R}(t), q_{Xi}^{F}(t)]}{\partial q_{Xi}^{E}(t)} = \dfrac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^{E}(t)} \\ \dfrac{\partial \hat{P}[q_{Xi}^{E}(t), q_{Xi}^{R}(t), q_{Xi}^{F}(t)]}{\partial q_{Xi}^{R}(t)} = \dfrac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^{R}(t)} \\ \dfrac{\partial \hat{P}[q_{Xi}^{E}(t), q_{Xi}^{R}(t), q_{Xi}^{F}(t)]}{\partial q_{Xi}^{F}(t)} = \dfrac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^{F}(t)} \end{cases} \tag{8.3a}$$

Considering the decision variables $q_{Xi}^{E}(t), q_{Xi}^{R}(t), q_{Xi}^{F}(t)$ of each individual profit-maximization model, it can be seen from (8.3a) that when a participant adjusts its generation output, reserve capacity or regulation capacity, the change of its own profit can be reflected by either the original individual objective function (8.2d) or the potential function $\hat{P}[q_{Xi}^{E}(t), q_{Xi}^{R}(t), q_{Xi}^{F}(t)]$, that is, the generator's profit difference of its own objective function and that of the potential function are exactly identical when it changes the bidding strategy in the electricity market. In short, the potential function can precisely capture the payoff deviation of each generator when it alters its capacity allocations considering the possible reactions of the rivals.

Therefore, taking the potential function as the new objective function and incorporating the technical constraints of all generators (8.1b)–(8.1e), (8.1g)–(8.1i), (8.1k)–(8.1l) and (8.1n), a reformulated Nash-Cournot equilibrium model can be obtained. The specific potential function $\hat{P}[q_{Xi}^E(t), q_{Xi}^R(t), q_{Xi}^F(t)]$ is derived as below.

8.3.2 Procedures to Derive the Potential Function

First, take the derivative of the single-period function (8.2d) with respect to $q_{Xi}^E(t)$, $q_{Xi}^R(t)$ and $q_{Xi}^F(t)$. We then have:

$$\begin{cases} \frac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^E(t)} = \alpha(t) - 2\beta(t) q_{Xi}^E(t) - \beta(t) \sum_{j \neq i}^{N} q_{Xj}^E(t) - \frac{dC_{Xi}[q_{Xi}^E(t)]}{dq_{Xi}^E(t)} \\ \frac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^R(t)} = \lambda^R(t) \\ \frac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^F(t)} = \lambda_C^F(t) + \theta_{Xi} \lambda_P^F(t) \end{cases} \tag{8.3b}$$

Second, sum all of the objective functions (8.2d) of each individual profit-maximization problem to obtain:

$$\begin{aligned} &\Pi[q_{Xi}^E(t), q_{Xi}^R(t), q_{Xi}^F(t)] \\ &= \alpha(t) \sum_{i=1}^{N} q_{Xi}^E(t) - \beta(t) \sum_{i=1}^{N} [q_{Xi}^E(t)]^2 \\ &\quad - \beta(t) \sum_{i=1}^{N} [q_{Xi}^E(t) \sum_{j \neq i}^{N} q_{Xj}^E(t)] - \sum_{i=1}^{N} C_{Xi}[q_{Xi}^E(t)] \\ &\quad + \lambda^R(t) \sum_{i=1}^{N} q_{Xi}^R(t) + \lambda_C^F(t) \sum_{i=1}^{N} q_{Xi}^F(t) + \lambda_P^F(t) \sum_{i=1}^{N} \theta_{Xi} q_{Xi}^F(t) \end{aligned} \tag{8.3c}$$

Third, consider (8.3b) and take the derivative of the above function (8.3c) with respect to $q_{Xi}^E(t)$, $q_{Xi}^R(t)$ and $q_{Xi}^F(t)$:

$$\begin{cases} \frac{\partial \Pi[q_{Xi}^E(t), q_{Xi}^R(t), q_{Xi}^F(t)]}{\partial q_{Xi}^E(t)} \\ \quad = \alpha(t) - 2\beta(t) q_{Xi}^E(t) - 2\beta(t) \sum_{j \neq i}^{N} q_{Xj}^E(t) - \frac{dC_{Xi}[q_{Xi}^E(t)]}{dq_{Xi}^E(t)} \\ \quad = \frac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^E(t)} - \beta(t) \sum_{j \neq i}^{N} q_{Xj}^E(t) \\ \frac{\partial \Pi[q_{Xi}^E(t), q_{Xi}^R(t), q_{Xi}^F(t)]}{\partial q_{Xi}^R(t)} = \lambda^R(t) = \frac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^R(t)} \\ \frac{\partial \Pi[q_{Xi}^E(t), q_{Xi}^R(t), q_{Xi}^F(t)]}{\partial q_{Xi}^F(t)} = \lambda_C^F(t) + \lambda_P^F(t) \theta_{Xi} = \frac{\partial \pi_{Xi}(t)}{\partial q_{Xi}^F(t)} \end{cases} \tag{8.3d}$$

Finally, considering (8.3a) and (8.3d), the potential function is:

$$\begin{aligned}&\sum_{t=1}^{T}\hat{P}[q_{Xi}^{E}(t),q_{Xi}^{R}(t),q_{Xi}^{F}(t)]\\&=\sum_{t=1}^{T}\left\{\begin{array}{l}\alpha(t)\sum\limits_{i=1}^{N}q_{Xi}^{E}(t)-\beta(t)\sum\limits_{i=1}^{N}[q_{Xi}^{E}(t)]^{2}-\beta(t)\sum\limits_{i=1}^{N}\sum\limits_{j<i}^{N}q_{Xi}^{E}(t)q_{Xj}^{E}(t)\\+\lambda^{R}(t)\sum\limits_{i=1}^{N}q_{Xi}^{R}(t)+\lambda_{C}^{F}(t)\sum\limits_{i=1}^{N}q_{Xi}^{F}(t)\\+\lambda_{P}^{F}(t)\sum\limits_{i=1}^{N}\theta_{Xi}q_{Xi}^{F}(t)-\sum\limits_{i=1}^{N}C_{Xi}[q_{Xi}^{E}(t)]\end{array}\right.\end{aligned}\tag{8.3e}$$

It can be seen from (8.3e) and (8.3c) that the potential function is not a simple summation of the objective functions of each individual profit-maximization problem but needs further derivation to satisfy (8.3a).

Thus, the conventional Nash-Cournot model can be transformed into an integrated single-level optimization problem by introducing the potential function as the new objective function and subjecting it to all of the technical constraints of various generators. The potential-function based Nash-Cournot problem is presented as follows, for $t=1,2,\ldots T$:

$$\begin{aligned}&\max\sum_{t=1}^{T}\hat{P}[q_{Xi}^{E}(t),q_{Xi}^{R}(t),q_{Xi}^{F}(t)]\\&\text{s.t.}\left\{\begin{array}{lll}(1\text{b})-(1\text{e}), & \text{Ei}=1,2,\ldots\text{ES}\\(1\text{g})-(1\text{i}), & \text{Ti}=1,2,\ldots\text{TP}\\(1\text{k})-(1\text{l}), & \text{Hi}=1,2,\ldots\text{HP}\\(1\text{n}), & \text{Ri}=1,2,\ldots\text{RE}\\(2\text{b})-(2\text{c})\end{array}\right.\end{aligned}\tag{8.3f}$$

Obviously, the above problem is a mixed integer quadratically constrained programming (MIQCP) problem, and it can be proved that the first-order KKT conditions of (8.3f) are completely equivalent to those of the multi-individual profit-maximization problems presented in subsection C of 2 (Kulkarni and Shanbhag 2014). Moreover, the CPLEX under GAMS is adopted to solve the MIQCP problem. For problems with integer variables, CPLEX uses a branch and cut algorithm. For the quadratically constrained program (QCP) model, only the Barrier method can be applied to solve the problem (GAMS Development Corporation 2009).

8.4 Numerical Examples

8.4.1 Basic Data

Numerical examples are implemented on a modified large scale power system with various types of generation technologies for theoretical analysis. The total installed capacity including the renewable units is 18270 MW and the peak load is 7500 MW, and the maximum demands of regulation and reserve services are 75 and 643.8 MW. The slope of the inverse demand function $\beta(t)$ which represents the demand elasticity is assumed to be as constant as 0.01 and the intercept $\alpha(t)$ which indicates the highest payment that consumers are willing to pay is shown as the following curve (Fig. 8.2).

Two basic scenario settings are included for comparison in this chapter. The Benchmark Scenario (BS) consists of 34 thermal units (26 % of the total generation mix), 9 hydro units (5 %), 90 wind turbines (47 %) and 52 photovoltaic stations (22 %), representing an optimistic case with high penetration of renewables. The Comparison Scenario (CS) introduces additional 12 ESSs (5 % of the total installed capacity of BS). Three typical wind power generation curves, i.e., inverse-peak, positive-peak and flat pattern, are studied in the two scenarios.

The generation costs, as well as the technical parameters of the thermal units, are referenced from (Vahidianasab and Jadid 2010). The technical parameters of hydro units are obtained from the real data of Guangdong Province in China. The wind farms have three kinds of rated capacities: 180, 100 and 60 MW and the number of each kind of wind units are 9, 54 and 27, respectively. For solar stations, they have four kinds of rated capacities: 100, 90, 80 and 60 MW and the number of each kind of solar units are 8, 4, 16 and 24. The rated capacity and energy limits of ESSs are 80 MW and 180 MWh. The overall efficiency of the ESSs is 81 %.

The capacity and performance prices of regulation service are referenced from the PJM regulation market on July 24, 2014 (PJM 2014). The capacity prices of reserve service are selected from the Texas reserve market on April 19, 2014 (Electric Reliability Council of Texas (ERCOT) 2014), as illustrated in Fig. 8.3.

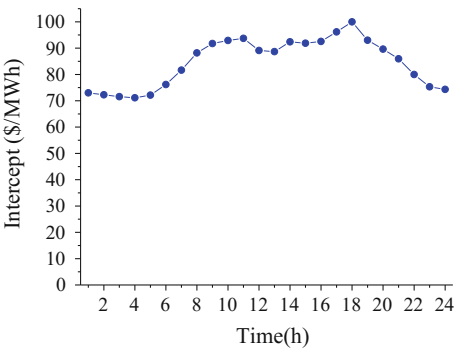

Fig. 8.2 The curve of the intercept of the inverse demand function

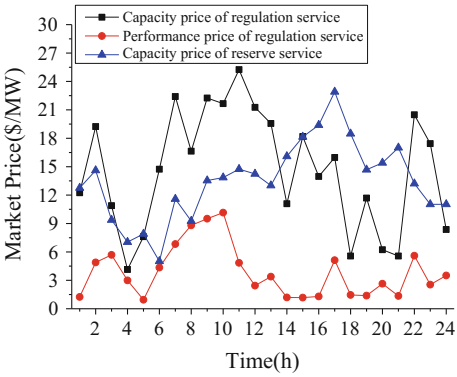

Fig. 8.3 Capacity and performance prices of the regulation service and capacity price of the reserve service

The stochastic nature of renewables is included in determining the demand of reserve services, according to the experience of the current operating markets. Apart from the emergency reserve (usually determined by the capacity of the largest generator), an additional 5 % of the forecasted generation of renewables is also included in the reserve demand to handle their intermittencies (Holttinen et al. 2012).

8.4.2 Comparison in the Energy Market

The competitive strategies for ESSs are intuitively to produce electricity when the energy price is high and to charge while the price is relatively low. When large-scale wind and solar power participate in the energy market, their cost-effective and stochastic properties may further widen the price difference between peak and valley and thus make the price curves more jagged, distorted and less predictable. These constantly varying prices will lead the ESSs to spontaneously respond and to strategically interact to maximize their own profits, as illustrated in Fig. 8.4. Fortunately, the profit-driven behaviors of the ESSs indirectly accommodate more renewable energy and achieve load shifting. The energy price will go down with large-scale renewable generation, for example, during the night with the inverse peak pattern, in the afternoon for the positive peak pattern, and in both the night and the afternoon for the flat pattern. During these periods the ESSs will choose to charge and consequently provide more flexibility for renewable generation.

When wind farms generate with the inverse-peak pattern, the most power produced by wind farms and solar stations can be accommodated because of the peak shifting of the renewables. By contrast, for the positive-peak pattern, too much wind power and solar power simultaneously produce electricity from 11:00 a.m. to 19:00 p.m. and part of the renewables would be curtailed due to the limited capacity of ESSs and the minimum output constraints of thermal units. In addition, the available integration of renewables in the flat pattern falls in between the inverse-peak pattern and the positive-peak pattern.

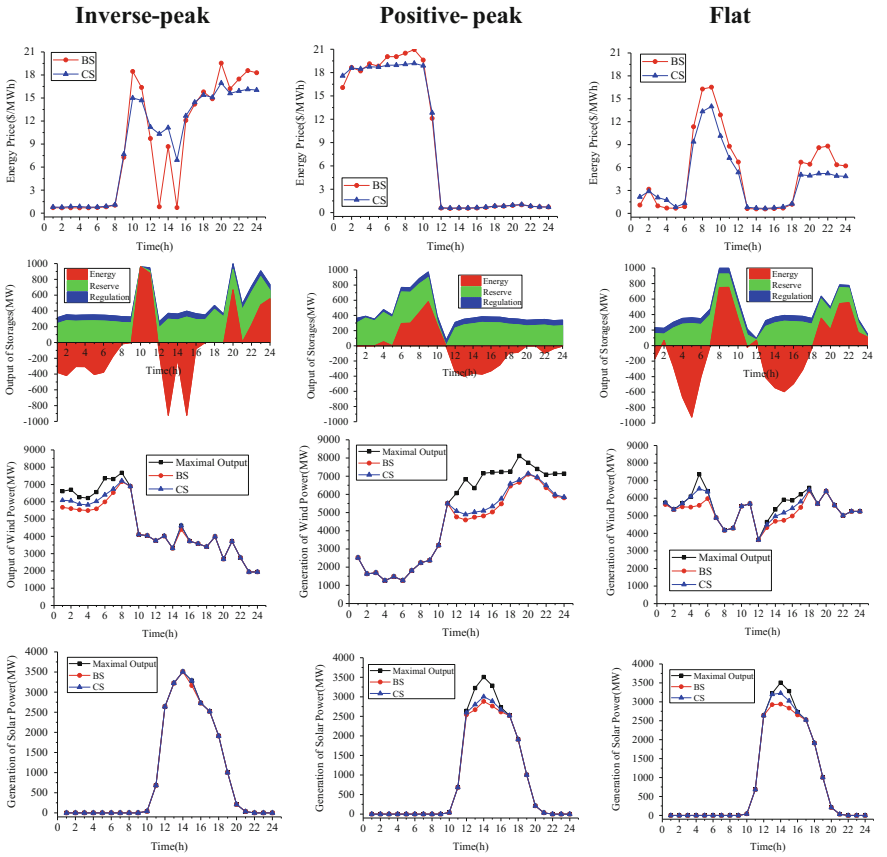

Fig. 8.4 Energy prices, capacity allocations of EESs, generation curves of wind farms as well as solar stations successively

8.4.3 Comparison in the AS Markets

When large-scale renewable generation is integrated into the power system, its stochastic and intermittent nature will require more regulation and reserve services provided by the conventional thermal units, hydro units and the emerging ESSs. The increasing demands of AS services will drive the regulation and reserve market prices higher; consequently, more fast response resources, such as hydropower and ESSs, will be attracted to provide the flexibility and ramping capability to handle the high penetration of renewables.

In addition, considering the "pay for performance" scheme and the operation characteristics of ESSs, more thermal plants can be replaced in the AS markets. Especially in the regulation market, one unit of storage capacity would be equivalent to several units of thermal capacity; thus, the ESSs will earn more payments

because of the different regulation mileage ratio for different response resources. In the numerical examples, the performance payments for ESSs and hydropower are three times those for the slow response thermal power (PJM 2015). Therefore, the ESSs and hydropower are more competitive, and thermal power is less willing to participate in the regulation market, as illustrated in Figs. 8.5 and 8.6. Compared to BS, almost all of the regulation demand and much of the reserve demand are satisfied by ESSs in the CS, leading the thermal units to be replaced by the ESSs in

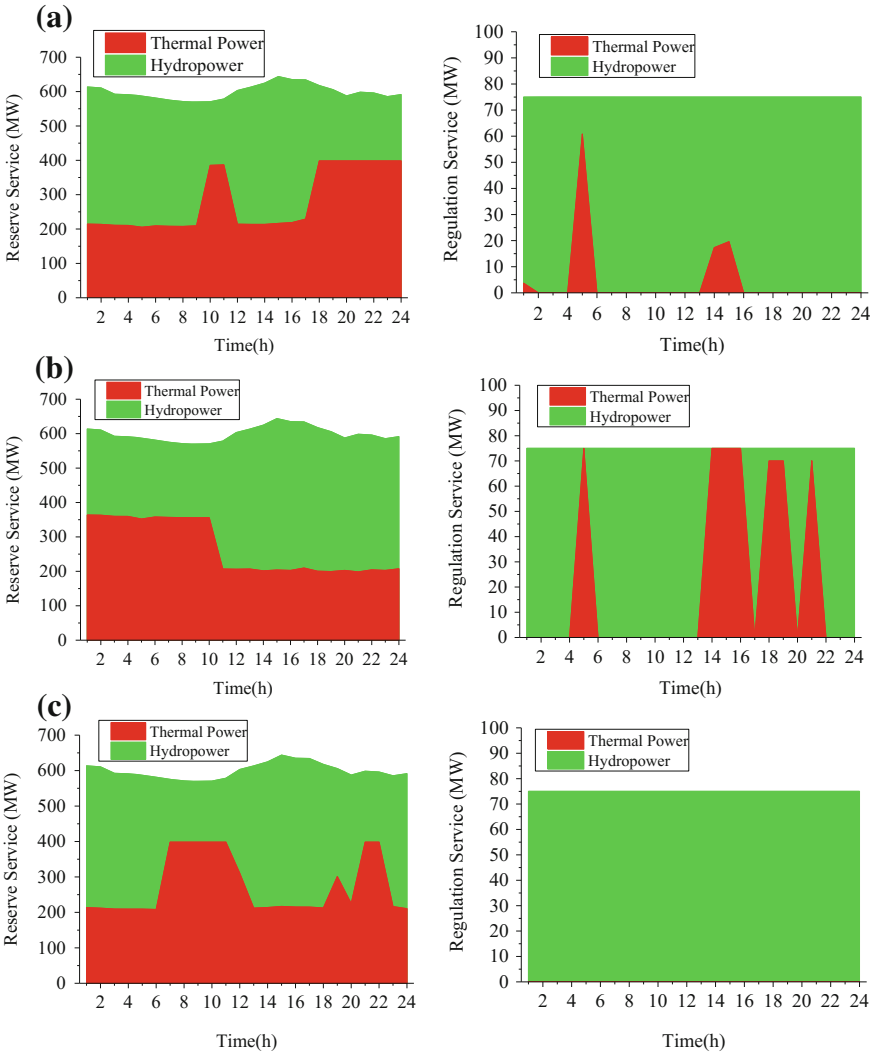

Fig. 8.5 Reserve and regulation market clearing results of the BS scenario. **a** Inverse-peak, **b** positive-peak, **c** Flat

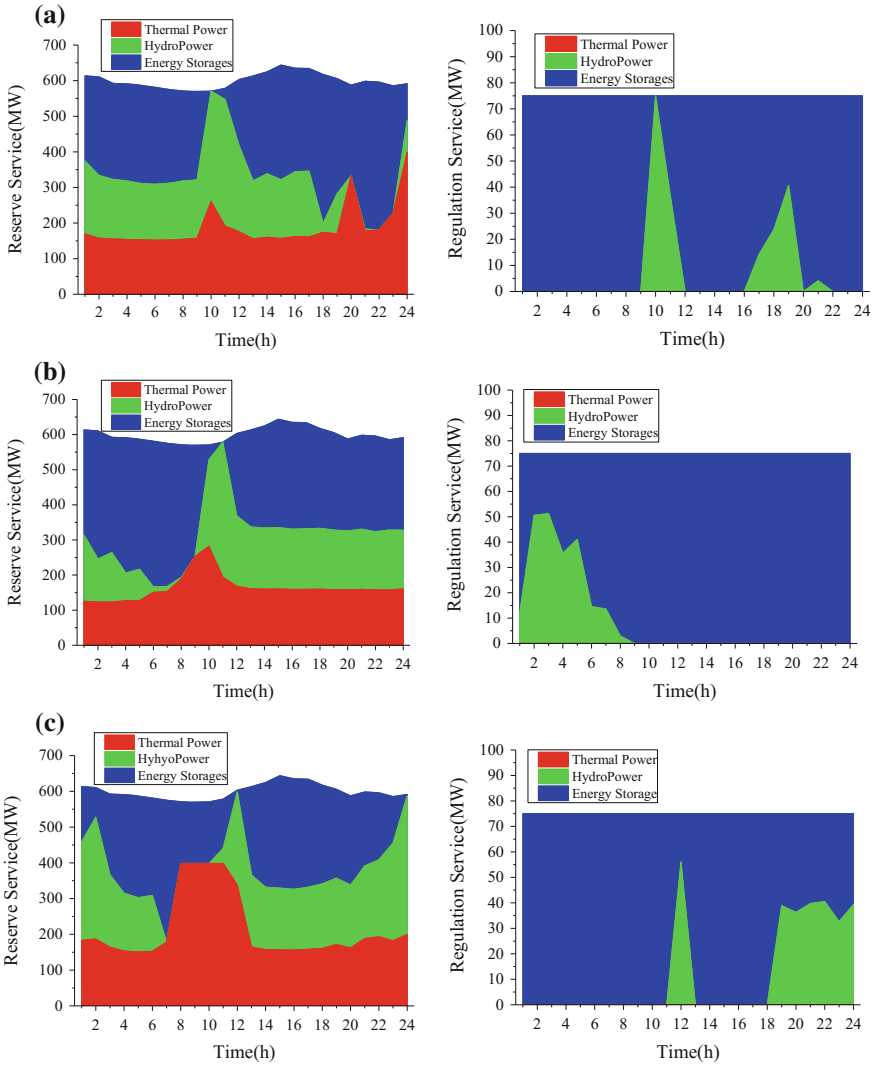

Fig. 8.6 Reserve and regulation market clearing results of the CS scenario. **a** Inverse-peak, **b** Positive-peak, **c** Flat

the AS markets. Then, the hydro units which are also substituted by the ESSs in the regulation and reserve markets will continue to replace the thermal units in the energy market. In this case, more renewables could be accommodated in the off-peak periods because the thermal units could further lower their output levels with providing less energy and reserve service and without providing regulation services. Meantime, more ramping capabilities of the thermal units can also be released. Moreover, the respective demand requirements of traditional regulation

Table 8.1 Renewable curtailment reductions with storages in the AS markets

Wind generation curve patterns	Wind curtailment reduction (MWh)	Solar curtailment reduction (MWh)
Inverse-peak pattern	271.9	0
Positive-peak pattern	411.2	42.8
Flat pattern	301.7	64.94

resources and fast regulation resources are not considered for the purpose of simplicity. This assumption does not impose significant effects on the results.

Moreover, in the CS scenario, the curtailment reduction of wind power could be as much as 411.2 MWh and that of solar power is 64.94 MWh, as presented in Table 8.1.

8.4.4 Algorithm Efficiency Comparisons

The potential-function (PF) based approach, GS and NCP methods are compared based on the CS scenario. The simulations are carried out on an Intel(R) Core(TM) i5-3210 M with 1 Quad-Core processors running at 2.5 GHz and 4 GB of RAM. The computation platform is GAMS 24.4 and the results are presented in Table 8.2.

The computation efficiency is significantly improved with the potential function, which is only approximately 7 s, whereas the GS method consumes more than one hour and the NCP method cannot obtain a feasible solution. The relative infeasibility of PF method decreases very fast while these of the GS and NCP methods go down slowly with many minor oscillations, as illustrated in Fig. 8.7. Moreover, the convergence criteria of the PF method is equal to 10–6, which is much stricter than the other two methods. The absolute feasible error of final computation results of the GS method is 0.088 and that of the NCP method is 2.775.

In short, the PF approach overwhelmingly outperforms the GS and NCP methods. The reformulated Nash-Cournot model is more efficient, robust and accurate. Moreover, the computation results of the GS method are similar to those of the PF approach, which not only reveals that the above example has a unique equilibrium point, but also indicates that the PF method has obtained the correct results. Moreover, if there exist multiple equilibria, the equilibrium that can achieve

Table 8.2 Efficiency comparisons of three solution methods

Methods	Computation time (s)	Iteration number	Solution status	Solver type	Convergence criteria
GS	5255.1	337	Feasible	CPLEX	0.1
NCP	126.82	629	Infeasible	DICOPT	0.1
PF	7.1500	35	Optimal	CPLEX	10^{-6}

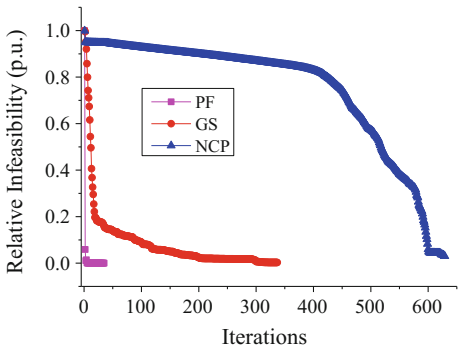

Fig. 8.7 The relative infeasibility curves of the three methods

the total generation cost minimization or the total profit maximization of all participants will be selected as the meaningful one (Carlos et al. 2012).

8.5 Conclusion

To explicitly and precisely evaluate the contribution of the independently-operated ESSs for supporting large-scale renewable generation in joint energy and AS markets, this chapter proposes a multi-period Nash-Cournot model to compute the market equilibrium with a new reformulation approach. The traditional bi-level equilibrium model can be transformed into an integrated single-level optimization problem by introducing a potential function as the new objective function and concatenating all the technical constraints of each generator. The numerical results validate the effectiveness of this novel model transformation technique and indicate that the individual profit-maximizing behaviors of ESSs in joint energy and AS markets indirectly provide better flexible services needed in the power system operation. In the energy market, the prices go down with large-scale wind or solar productions when the ESSs choose to charge and further improve the integration of renewables. In the AS markets, the ESSs with fast response properties are more competitive, and the conventional thermal units can be replaced to further lower the generation level and more ramping capabilities can be released to support more renewable generation. These are the contributions of the ESSs to improve the integration of large-scale renewables.

The work of this chapter is a new exploration and more effort will be performed in the future, including taking the transmission constraints and unit commitment into consideration. Hopefully, this chapter can provide some new ideas in the area of promoting the development of ESSs.

References

Aigner, T., Jaehnert, S., Doorman, G. L., & Gjengedal, T. (2012). the effect of large-scale wind power on system balancing in Northern Europe. *IEEE Transactions on Sustainable Energy, 3* (4), 751–759.

Akhavan-Hejazi, H., & Mohsenian-Rad, H. (2014). Optimal operation of independent storage systems in energy and reserve markets with high wind penetration. *IEEE Transactions on Smart Grid, 5*, 1088–1097.

Ali, Karimi V., Ali, D., & Hassan, M. (2011). A new self-scheduling strategy for integrated operation of wind and pumped-storage power plants in power markets. *Applied Energy, 88*, 5002–5012.

Carlos, R., Antonio, J., & Conejo, Yves S. (2012). Equilibria in an oligopolistic electricity pool with stepwise offer curves. *IEEE Transactions on Power Systems, 27*, 752–761.

Deb, C. (2004). Multicommodity spatial cournot model for generator bidding analysis. *IEEE Transactions on Power Systems, 19*(1), 267–275.

Díaz-González, F., Sumper, A., Gomis-Bellmunt, O., & Villafáfila-Robles, R. (2012). A review pf energy storage technologies for wind power applications. *Renewable and Sustainable Energy Reviews, 16*, 2154–2171.

Drew, F., & Jean, T. (2013). *Game Theory* (Vol. I, p. 14). Cambridge: MIT Press.

Electric Reliability Council of Texas (ERCOT). (2015). http://www.ercot.com/.

Energy Information Administration. (2013, April). *Updated capital cost estimates for utility scale electricity generating plants*. U.S. Department of Energy, Washington, DC, Technical report TR-20585 (6–7).

Gabriel, S. A., Conejo, A. J., Fuller, J. D., et al. (2013). C*omplementarity modeling in energy markets* (vol. I, p. 283). New York: Springer.

GAMS Development Corporation. (2009). *GAMS: The solver manuals* (vol. I,, p. 183). GDC.

He, G., Chen, Q., Kang, C., Pinson, P., & Xia, Q. (2015). Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life. *IEEE Transactions on Smart Grid*.

Holttinen, H., Milligan, M., Ela, E., Menemenlis, N., Dobschinski, J., Rawn, B., et al. (2012). Methodologies to determine operating reserves due to increased wind power. *IEEE Transactions on Sustainable Energy, 3*(4), 713–723.

Jian, Y., Shmuel, O. S., & Llan, A. (2007). Cournot equilibria in two-settlement electricity markets with system contingencies. *International Journal of Critical Infrastructures, 3*(1), 142–160.

Jiang, R., Wang, J., Zhang, M., & Guan, Y. (2013). Two-stage MInimax regret robust unit commitment. *IEEE Transactions on Power Systems, 28*(3), 2271–2282.

Klemperer, P. D., & Meyer, M. A. (1989). Supply function equilibria in oligopoly under uncertainty. *Econometrica, 57*(6), 1243–1277.

Kulkarni, A. A., & Shanbhag, U. V. (2014). *A shared-constraint approach to multi-leader multi-follower games*. Anal: Set-Valued Var.

Li, G., Shi, J., & Qu, X. (2011). Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market—A state-of-the-art review. *Energy, 36*(8), 4686–4700.

Maria, D., Giuseppe, F., Mariagiovanna, P., & Michele, T. (2012). Planning and operating combined wind-storage system in electricity market. *IEEE Transactions on Sustainable Energy, 3*(2), 209–217.

Masiello, R. D., Roberts, B., & Sloan, T. (2014). Business models for deploying and operating energy storage and risk mitigation aspects. *Proceedings of the IEEE, 201*(7), 1052–1064.

Molina, J. P., Zolezzi, J. M., & Contreras, J. (2011). Nash-Cournot equilibria in hydrothermal electricity markets. *IEEE Transactions on Power Systems, 26*(3), 1089–1101.

Monderer, D., & Shapley, L. S. (1996). Potential games. *Games and Economic Behavior, 14*, 124–143.

PJM. (2015). *Working to perfect the flow of energy*. http://www.pjm.com/.

Tamaschke, R., Docwra, G., & Stillman, R. (2005). Measuring market power in electricity generation: A long-term perspective using a programming model. *Energy Economics, 27*(2), 317–350.

Thatte, A. A., Xie, L., Viassolo, D. E., & Singh, S. (2013). risk measure based robust bidding strategy for arbitrage using a wind farm and energy storage. *IEEE Transactions on Smart Grid, 4*, 2191–2199.

The BP Energy Outlook. (2015). http://www.bp.com/.

Usaola, J. (2011). Operation of concentrating solar power plants with storage in spot electricity market. *IET Renewable Power Generation, 6*(1), 59–66.

Vahidianasab, V., & Jadid, S. (2010). Stochastic multi-objective self-scheduling of a power producer in joint energy and reserve markets. *Electric Power Systems Research, 80*, 760–769.

Voorneveld, M. (2000). Best-response potential games. *Economics Letters, 66*, 289–295.

Wang, C., Lu, Z., & Qiao, Y. (2013). A consideration of the wind power benefits in day-ahead scheduling of wind-coal intensive power systems. *IEEE Transactions on Power Systems, 28* (1), 236–245.

Willems, B., Rumiantseva, I., & Weigt, H. (2009). Cournot versus supply functions: What does the data tell us? *Energy Econ., 31*(1), 38–47.

Woo, C., Zarnikau, J., Kadish, J., Horowitz, I., Wang, J., & Olson, A. (2013). The impact of wind generation on wholesale electricity prices in the hydro-rich Pacific Northwest. *IEEE Transactions on Power Systems, 28*(4), 4245–4253.

Xiao, Y., Su, Q., Bresler, F. S., Carroll, R., & Schmitt, J. R. (2014). Performance-based regulation model in PJM wholesale markets. In *Proceedings of the IEEE Power Engineering Society on General Meeting*.

Yousefi, A., Iu, H. H., Fernando, T., & Trinh, H. (2013). An approach for wind power integration using demand side resources. *IEEE Transactions Sustainable Energy, 4*(4), 917–924.

Chapter 9
Employing Chemical Processes as Grid-Level Energy Storage Devices

Michael Baldea

Abstract The contribution of power generation from renewable sources to the U.S. energy markets has witnessed an explosive growth, leading to increased variability and uncertainty in grid operations. Demand response (DR) operation of chemical processes, based on overproducing and storing product(s) during off-peak times, when grid demand is low, and reducing production rates and (partially) meeting product demand using the stored products during peak price times, can be used to create a virtual grid-scale "battery" and help alleviate power generation variability. In this chapter, we discuss the interplay between process-centric and utility-centric optimization of the operating schedule of DR chemical processes. Given the frequent production rate changes required in such circumstances, the scheduling calculations are carried out under a set of dynamic constraints that represent the scheduling-relevant closed-loop behavior of the processes. We propose a reconciliation mechanism, based on two-way communication of a limited set of information between the utility and process operators. Utility operators use a simplified, battery-like representation of the energy storage capabilities of each DR process to compute the utility-optimal process energy consumption profile. Then, process operators determine the energy price profile that would allow them to implement this consumption profile without incurring economic losses, and communicate it to the utility company. This scheme reduces the dimension of the optimization problem that must be solved at the utility level and provides an equitable and transparent means for chemical process operators to engage in DR schemes.

Nomenclature

Sets

- $i = 1, \ldots, N_i$ scheduling time slot
- $j = 1, \ldots, N_j$ discretized time points within each scheduling time slot
- $k = 1, \ldots, K$ DR process
- $l = 1, \ldots, L$ generator

M. Baldea (✉)
McKetta Department of Chemical Engineering, The University of Texas at Austin, Austin, TX 78712, USA
e-mail: mbaldea@che.utexas.edu

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_9

Parameters

- S_k^{max} maximum capacity of storage system in DR process k (kg)
- t^p scheduling time horizon (divided in N_i time slots) (h)
- $\bar{M}_k$ nominal production rate of process k (kg/h)
- $H_{V,NG}$ natural gas heating value (kJ/m^3)
- P^o power demand of other consumers (kW)
- P_i^o power demand of other consumers (grid load) in time slot i (kW)
- ΔM_k flexibility of process k, defined in terms of turn-up/turn-down capacity relative to the nominal production rate $\bar{M}_k$ (kg/h)
- η_l efficiency of generator l (%)
- δ_k demand rate for process k (kg/h)
- ζ_k specific energy consumption of DR process k (kJ/kg)
- π^e real-time energy price ($/kWh)
- π_i^e real-time energy price at time corresponding to slot i ($/kWh)
- π^{NG} natural gas price ($/m^3)
- π_k^p product price of process k ($/kg)
- ρ_l maximum ramp rate of generator l (kW/h)
- τ_k time constant of process k (h)

Continuous variables

- F_l^{NG} natural gas flow rate to generator l (m^3/h)
- $F_{i,j,l}^{NG}$ natural gas flow rate to generator l at time point j in scheduling slot i (m^3/h)
- $M_{i,j,k}$ production rate of process k at time point j in scheduling slot i (kg/h)
- $\hat{M}_k$ target production rate of process k (kg/h)
- $\hat{M}_{i,k}$ target production rate of process k in scheduling slot i (kg/h)
- π_k^e tailored energy price/incentive for process k ($/kWh)
- P_k^p power demand of process k (kW)
- ΔP_k^p change in power demand of process k (kW)
- $P_{i,j,k}^p$ power demand of process k at time point j in scheduling slot i (kW)
- P_l^g power generation of generator l (kW)
- $P_l^{g,max}$ maximum power generation of generator l (kW)
- $P_l^{g,min}$ minimum power generation of generator l (kW)
- $P_{i,l}^g$ power generation of generator l in scheduling slot i (kW)
- $P_{i,j}^{mkt}$ market power purchase at time point j in scheduling slot i (kW)
- Q_k "state of charge" of battery equivalent of process k (kg)
- $\hat{\pi}_{i,k}$ incentive energy price for process k at time slot i ($/kWh)
- $S_{i,j,k}$ amount of stored product in process k at time point j in scheduling slot i (kg)
- $\alpha_{i,j,k}$ flow rate of product *into* storage in process k at time point j in scheduling slot i (kg/h)
- $\beta_{i,j,k}$ flow rate of product *from* storage in process k at time point j in scheduling slot i (kg/h)

9.1 Introduction

The U.S. energy markets, and in particular the power generation portfolio, have witnessed significant changes in the past decade. Developments in extraction technology have made natural gas a much more economical hydrocarbon feedstock, enabling at the same time developments in efficient and more environmentally benign technologies such as combined heat and power. The deployment of renewable-based generation technologies has also grown explosively, with recent reports showing a tripling of the contribution of renewable energy to U.S. electricity generation in the past decade (US Energy Information Administration 2013b).

While highly encouraging from the environmental and energy security perspectives, these developments also bring about several challenges related to meeting an increasingly time-variable demand. Specifically, the amount of solar and wind power generation undergoes daily variations, peaking at mid-day (solar) and midnight or later (wind); meanwhile, grid demand peaks in the late afternoon (Ondeck et al. 2015). Grid operators and utility companies must close this peak-time gap by increasing generation rates of conventional (fossil-based) power plants or by purchasing energy from the open market. The latter option can be very costly, since peak-time energy prices are often considerably higher than off-peak tariffs under the real-time pricing scenarios enabled by increasingly widespread market deregulation measures.

The need to balance generation and demand rates has spurred efforts aimed at developing demand response (DR) strategies, defined by the Federal Energy Regulatory Commission (FERC 2013) as

> Changes in electric usage by demand-side resources from their normal consumption patterns [in response to economic incentives and dynamic pricing structures] at times of high wholesale market prices or when system reliability is jeopardized.

In the medium term, the implementation of DR is expected to lower peak power demand in the U.S. by about 5 % (about 50GW) (US Energy Information Administration 2013a). Efforts are underway to implement such strategies in all major classes of electricity users, i.e., commercial buildings, residential (home) users, and the industrial sector. Buildings can use thermal energy storage (TES) to alter their grid load, especially those associated with the use of air conditioning. This entails pre-cooling the building structure and/or storing refrigerant (e.g., chilled water) during the off-peak hours, and using the stored refrigeration to satisfy cooling demand during peak price times (which typically coincide with peaks in ambient temperature. This strategy can shift a considerable portion of the energy demand of buildings from peak to off-peak times (see, e.g., Touretzky and Baldea (2016) and references therein for more information).

In the industrial category, chemical processes represent a very promising contributor owing to several DR-favorable features (Siirola and Edgar 2012; Soroush and Chmielewski 2013). First, many chemical processes have the ability to modulate their production rates and the associated power use over time scales relevant to the fluctuations in grid demand (i.e., minutes to hours). Moreover, excess capacity

and the ability to store excess product are often available. Intuitively, these features support a DR behavior that consists of, (i) overproducing and storing product(s) during off-peak times, when grid demand is low, and, (ii) reducing production rates and (partially) meeting product demand using the stored products during peak times. By following this operating pattern (which is analogous to the use of TES in buildings as described above), a chemical process acts as a grid-scale storage system, accumulating energy in the form of salable chemical compounds. This demand-side behavior must be supported on the supply side with time-varying energy price structures, such as time-of-use (TOU) pricing, whereby different price tiers are charged over well-defined time intervals during the day, or day-ahead or even real-time pricing, where prices vary over much shorter time periods and are correlated with the wholesale price paid by the utility when purchasing energy on the deregulated market.

Several literature studies have considered the DR applications in specific branches of the chemical industry. Paulus and Borggrefe (2011) analyzed the DR potential of the pulp and paper, aluminium, chlor-alkali, cement and steel sectors. DR of air separation plants was discussed, e.g., in Ierapetritou et al. (2002), Zhu et al. (2011), Pattison and Baldea (2014), Cao et al. (2015), Zhang et al. (2015). Increasing the involvement of the process utility system in DR was discussed by Feng et al. (2015). Other related DR applications include managing energy consumption related to the transport of liquids via pumping (Van Staden et al. 2011) and of solids via conveyor belts (Middelberg et al. 2009).

Most DR applications in the process industries focus on the tertiary/day-ahead market (Paulus and Borggrefe 2011), and daily chemical production schedules are typically determined by solving an optimization problem aimed at maximizing profit while accounting for time varying energy prices. Since most manufacturers are not willing to relinquish control of their facilities to the utility provider, this optimization calculation is performed by the customer, potentially following some interaction with the utility provider (Paulus and Borggrefe 2011; Feng et al. 2015).

While cost-optimal from a *demand* perspective, this approach is very likely to be suboptimal from the point of view of the *utility*: unilaterally optimizing production schedules for individual plants to take advantage of variable energy prices may pose the risk of (collectively) producing a new demand peak a different time of the day, thereby possibly exacerbating rather than solving the original peak demand issue.

Furthermore, participation in tertiary markets requires that the aforementioned production scheduling decisions be made over time horizons of the order of 1 day. As a consequence, the time scale of the scheduling problem overlaps with the response time of the plant itself (Pattison et al. 2016b), and the process dynamics should be accounted for explicitly in DR scheduling calculations in order to ensure that the resulting schedules are feasible from a *dynamic* point of view. While the role of dynamics in the interaction of a chemical process with the grid has been acknowledged in the literature (Garcia et al. 2013; Pattison et al. 2016b), the development of a comprehensive framework for incorporating dynamics information in scheduling the production of plants participating in DR schemes as well as in optimizing the operations of electric utility providers (and the grid), remains an open research question.

Motivated by the above, this chapter tackles three main objectives: first, we use recent results in integrating process dynamics and control into production scheduling to study the (dynamic) interaction of chemical plants with the power grid from the perspective of the *plant*. We refer to this problem as P1. Specifically, time-varying production targets are computed as the solution of a scheduling problem aimed at maximizing plant profit while explicitly accounting for the process dynamics and forecasts of energy prices. Subsequently, we consider the converse problem (which we refer to as P2), i.e., maximizing profit from the point of view of the *utility provider*. Owing to the facts that, (i) process operators may not be willing to share detailed process dynamic models with their utility suppliers, and, (ii) utility-level optimization may require accounting for a substantial number of DR entities (thereby considerably increasing the problem size), we introduce a new low-order modeling framework based on representing each chemical plant as a utility-level storage battery (together with the relevant dynamics). We reformulate the utility-level problem in terms of optimizing the "battery" charge-discharge cycles along with the timing and rate at which energy is purchased on the open market to meet a fluctuating outside (e.g., residential and commercial building) demand. Finally, we observe that the optimal plant operating schedules determined by P1 and P2 may differ considerably (owing to the need to avoid a secondary peak), and propose a reconciliation mechanism. We define a means for creating equitable customized price incentive structures for individual DR entities, whose effect is to shape individual energy use patterns such that the overall behavior of *the ensemble* of DR entities is at or near the optimum determined from P2. Throughout the chapter, we use a simple a case study to illustrate the theoretical results.

9.2 System Description

We consider an ensemble of K single-product process systems that act as DR entities, connected to a utility system which also comprises L natural gas-powered generators as shown in Fig. 9.1. The utility system operators have the option of meeting the power demand of the K process systems and other consumers (including, e.g., commercial and residential buildings) by a combination of, (i) modulating the power generation level, (ii) purchasing energy and, (iii) relying on the DR capabilities of the process systems. The latter entails scheduling chemical production such that the processes contribute to meeting the total demand. Our discussion focuses solely on *active* power as a measure of net energy transfer between the elements of the system considered. We assume that the presence of renewable energy-based generators is accounted for in the energy price structure, and do not model them explicitly in this system.

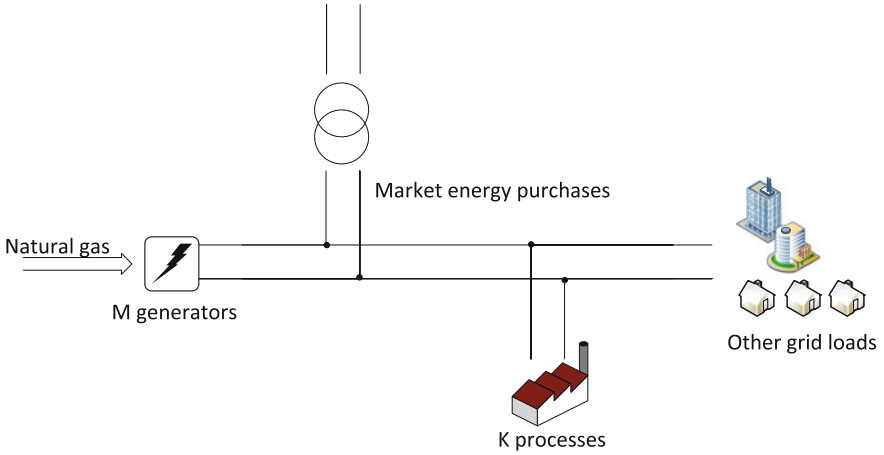

Fig. 9.1 Utility system with L generators and K demand-response (DR) chemical processes. The time-varying demand of, e.g., commercial and residential buildings can be met by a combination of, (i) increasing the production of generators, (ii) purchasing energy from the open market and, (iii) modulating the operation of the chemical plants

9.2.1 Process Systems

The dynamics of a process system are typically (see, e.g., Baldea and Daoutidis 2012) modeled by an input-affine system of differential equations of the form

$$\begin{aligned} \dot{\mathbf{x}}_k &= \mathbf{f}_k(\mathbf{x}_k) + \mathbf{G}_k(\mathbf{x}_k)\mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k) \end{aligned} \tag{9.1}$$

which relates the state and output vectors $\mathbf{x}_k$ and $\mathbf{y}_k$ of process $k \in \{1, \ldots, K\}$ to the input vector $\mathbf{u}_k$. Such models capture essential dynamic information regarding e.g., product quality and production rates, as well as the evolution of process states (temperatures, pressures, etc.) that must be known in order to abide by process safety constraints and equipment limitations, and are used for the design of *process control* systems.

The aforementioned dynamic information is typically discarded in *production scheduling* calculations, where it is replaced by static parameters approximating the transition times between process states (e.g., different products, different production rates) (Maravelias 2012). This representation is appropriate when the scheduling time horizon is extensive, and the time elapsed between two scheduled changes in process state is much longer than the time constant of the process (i.e., the process reaches steady state after each change) (Baldea et al. 2015; Du et al. 2015; Pattison et al. 2016b).

However, the DR operation of chemical processes in fast-changing electricity markets calls for changes in the production schedule to occur over much shorter

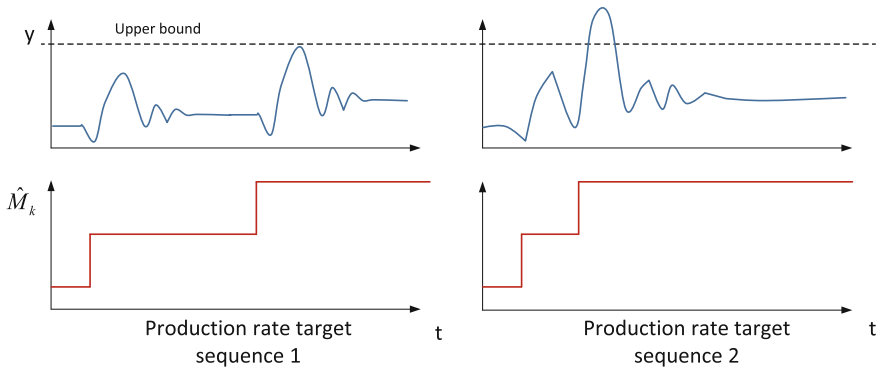

Fig. 9.2 (*Bottom*) Scheduled production rate target changes implemented in a hypothetical process. In the first case (*left*), the duration of the production time slots is sufficiently long to allow the process states and outputs (*top*) to reach steady state after each change. In the second case (*right*), the frequency of scheduling decisions increases, with the length of the production time slots being of the same order of magnitude as the time constant of the process. In this case, the process does not reach steady state after each production target change, and constraint violations may occur. In this latter case, scheduling calculations must account explicitly for the dynamics of the process (adapted from Touretzky (2016))

time horizons, which are comparable to (or even shorter than) the process response time. In this case, the dynamics of the process (and its control system) must be be explicitly accounted for in production scheduling decisions in order to ensure that the scheduled transitions remain feasible from a *process dynamics* point of view (Pattison et al. 2016b), as shown in Fig. 9.2.

At a first glance, this would entail embedding the process model (9.1) in the formulation of the production scheduling optimization problem. However, owing to the typically considerable dimensions and the multiple time scale nature of process models (Baldea and Daoutidis 2012), the result is a high-dimensional, multi-scale (mixed-integer) dynamic optimization (MIDO) problem that is difficult—if not impossible—to solve in a reasonable amount of time for systems of practical size. As a consequence, in our recent research (Baldea et al. 2015; Du et al. 2015; Pattison et al. 2016b), we have focused on characterizing the *scheduling-relevant* dynamics of chemical processes. In particular, we described the evolution of production rates, product compositions, and of the subset of state variables that correspond to process and equipment operating constraints. We have shown that these dynamics can be described using low-dimensional models (referred to as time scale-bridging models (SBMs)) that capture the closed-loop behavior of the process *and* its control system. SBMs relate the evolution of the aforementioned variables to the production rate and product purity/grade target (setpoint) values, which are time-varying and produced by the scheduling calculation.

For the purpose of the discussion in this chapter, we will rely on the developments in Du et al. (2015) and consider that the K single-product processes (together with their respective control systems) can be represented via linear first-order SBMs

capturing the dynamic response of the production rate, M_k, to changes in the production rate targets $\hat{M}_k$, which are defined by a DR scheduling calculation (which will be discussed later in the chapter):

$$\tau_k \dot{M}_k = (\hat{M}_k - M_k) \qquad \forall_k \tag{9.2}$$

Moreover, we assume that each process has facilities for materials storage, with the dynamics of the stored inventory levels, S_k, being described by:

$$\dot{S}_k = \alpha_k - \beta_k \qquad \forall_k \tag{9.3}$$

and the capacity constraint

$$S_k \leq S_k^{max} \qquad \forall_k \tag{9.4}$$

is imposed. The material flow rates to (α_k) and from (β_k) the storage systems are constrained to be within the turn-up/turn-down capability of the process:

$$\begin{aligned} 0 \leq \alpha_k \leq \Delta M_k \qquad \forall_k \\ 0 \leq \beta_k \leq \Delta M_k \qquad \forall_k \end{aligned} \tag{9.5}$$

and are chosen such that the product demand rate δ_k is met at all times:

$$\delta_k = M_k + \alpha_k - \beta_k \qquad \forall_k \tag{9.6}$$

The power consumption of process k is assumed to be proportional to the production rate:

$$P_k = \zeta_k M_k \tag{9.7}$$

9.2.2 Utility System Model

Each of the L natural-gas fueled generators is assumed to have efficiency η_l, which can be used to calculate power generation as a function of the flow rate of natural gas:

$$P_l^g = \frac{1}{3600} \eta_l H_{V,NG} F_l^{NG} \qquad \forall_l \tag{9.8}$$

Generators are subject to ramp rate constraints,

$$\frac{dP_l^g}{dt} \leq \rho_l \qquad \forall_l \tag{9.9}$$

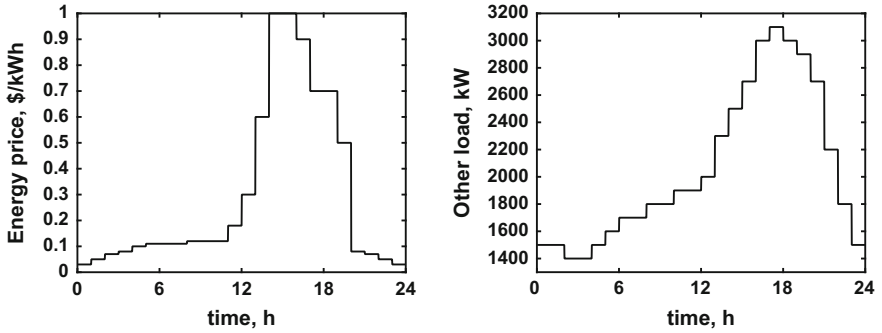

Fig. 9.3 *Left* Real-time energy prices, π^e; *Right* Time-dependent load profile, P^o

The operation of each generator and decisions concerning real-time market purchases, P^{mkt}, are based on meeting demand from the DR processes and other loads P^o (e.g., residential and commercial buildings—Fig. 9.1) at all times:

$$\Sigma_{l=1}^{L} P_l^g + P^{mkt} = \Sigma_{k=1}^{K} P_k^p + P^o \tag{9.10}$$

9.2.3 Load and Energy Price Profiles

We assume that residential consumption is the main cause of variability in the utility load (Wattles 2012) and consider a time-dependent load profile P^o with an early evening peak as shown in Fig. 9.3 (here we assume that the other large loads—such as buildings—do not participate in any demand response activities). Such demand profiles also lead to a mid-afternoon peak in real-time energy prices, and we consider the price profile π^e shown in Fig. 9.3.

9.3 Process-Centric Operation Scheduling

The net operating profit of process k is defined as the objective function

$$J_k^{process} = t^p \delta_k \pi_k^p - \int_0^{t^p} P_k^p \pi^e dt \tag{9.11}$$

which reflects the difference between the profit derived from product sales and the operating cost related to purchasing power. Here, the scheduling horizon t^p is assumed to be at least as long as the duration of the price and demand predictions (e.g., 24 h in Fig. 9.3).

Process-centric operation involves finding the production schedule (i.e., the time-varying values of $\hat{M}_k$, α_k, β_k over the horizon t^p) that maximizes the profit subject to accounting for the process dynamics and time varying energy prices, and is formulated as optimization problem P1, which is solved for each process k:

$$\begin{aligned} \underset{\hat{M}_k,\alpha_k,\beta_k}{max} \; & J_k^{process} \\ s.t. \; & \text{dynamic process model (SBM), (1.2)} \\ & \text{storage system model (1.3)} \\ & \text{capacity and flow bounds(1.4), (1.5)} \\ & \text{demand satisfaction(1.6)} \end{aligned} \tag{P1}$$

additionally, the constraint

$$S_k(t = t_p) \geq S_k(t = 0) \qquad \forall_k \tag{9.12}$$

is imposed, stipulating that the amount of product stored at the end of the scheduling horizon be greater or equal to the amount available at the beginning of the horizon, which avoids "falsifying" profits via the sale of previously stored material.

Thus, profit maximization is achieved via optimizing the interplay between modulating the production rate of the process (and hence its power consumption), and replenishing and depleting stored inventory.

Remark 1 The process-centric scheduling problem formulation (P1) above, as well as the utility-centric scheduling (P2), below, rely on several several assumptions, summarized here:

A1. The product demand rate δ_k is constant in time for all K processes.
A2. All K processes are continuous, single-product and production rate changes do not require turning equipment on or off (also, for now, we do not account for the DR potential of batch operations).
A3. The scheduling-relevant closed-loop dynamics of the process are described using a linear SBM (Eq. (9.2)).

Specifically, A2 eliminates the need for product and equipment selection/assignment decisions in the scheduling problem (and hence the need for binary variables in the problem formulation). Assumption A3 preserves the linear nature of the constraint system of the optimization problems; while we use first-order models for simplicity, the approach can easily be extended to higher-order systems.

However, it is important to note that while these assumptions simplify the formulation of the scheduling problems, they do not affect the fundamental concepts behind the proposed framework. In effect, multi-product facilities and equipment on/off cycling can be accommodated by incorporating the appropriate binary decision variables, while nonlinear SBMs can be constructed and used following, e.g., the approach proposed in Pattison et al. (2016a).

9.4 Utility-Centric Operation Scheduling

Similar to the case of process-centric scheduling, utility-centric operations scheduling entails optimizing the operating economics of the utility system. In this case, the objective function

$$J^{utility} = \int_0^{t^p} \left[\left(\Sigma_{k=1}^{K} P_k^p + P^o - P^{mkt} \right) \pi^e - \Sigma_{l=1}^{L} F_l^{NG} \pi^{NG} \right] dt \tag{9.13}$$

reflects the profit generated by selling energy (to the chemical processes and other consumers, the latter represented by the load P^o), and the expenditures associated with energy acquired from the market and natural gas purchases. We assume for simplicity that the price paid by the customers is the same as the real-time market price.

We will assume for now that the utility operators are omniscient and omnipotent, in the sense that they have complete knowledge of and control over the operation of all K processes, in addition to controlling the L generators. Thus, utility-centric operation involves finding the production schedule for each process, and the generating schedule for each generator, that maximizes the profit $J^{utility}$ subject to accounting for the process dynamics, ramp rate and capacity constraints, and time varying energy prices, and is formulated as problem P2 as follows:

$$\begin{aligned} &\max_{\hat{M}_k, \alpha_k, \beta_k, P^{mkt}, P_l^g} \; J^{utility} \\ &s.t. \quad \left.\begin{array}{l} \text{dynamic process model (SBM), (1.2)} \\ \text{storage system model (1.3),(1.12)} \\ \text{capacity and flow bounds(1.4), (1.5)} \\ \text{product demand satisfaction(1.6)} \\ \text{electricity demand satisfaction (1.10)} \\ \text{generator power calculation (1.8)} \\ \text{ramp rate constraints (1.9)} \end{array}\right\} \forall_k \end{aligned} \tag{P2}$$

9.5 Example

Let us consider a system comprising one power generation facility ($L = 1$) and one DR chemical process ($K = 1$), with the price and demand structures shown in Fig. 9.3. We assume that the generator and the process are modeled as described in Sect. 9.2, with the parameters in Table 9.1. The storage system is assumed to be full at the beginning of the day, $S(t = 0) = S^{max}$.

We solve both the process-centric and the utility-centric scheduling problems, as described in Sects. 9.3 and 9.4. We develop discrete-time versions of these problems, considering a scheduling horizon $t^p = 24h$, divided into 24 hourly time slots

Table 9.1 Parameters of example system

Process		Generator	
ΔM	200 kg/h	$P^{g,min}$	2300 kW
$\delta = \bar{M}$	1000 kg/h	$P^{g,max}$	4600 kW
τ	0.3 h	η	0.4
ζ	2 kWh/kg	$H_{V,NG}$	38376.71 kJ/m^3
π^p	1 \$/kg	π^{NG}	0.113 \$/m^3
S^{max}	1000 kg	ρ	1840.00 kW/h

($N_i = 24$), each in turn divided into 60 minute-long time intervals $N_j = 60$. Thus, the objective function for process-centric operation becomes:

$$J^{process} = t^p \delta \pi^p - \Sigma_{i=1}^{N_i} \Sigma_{j=1}^{N_j} P_{i,j}^p \pi_i^e \tag{9.14}$$

while the objective function for utility-centric operation is:

$$J^{utility} = \Sigma_{i=1}^{N_i} \Sigma_{j=1}^{N_j} \left[\left(P_{i,j}^p + P_i^o - P_i^{mkt} \right) \pi_i^e - F_{i,j}^{NG} \pi^{NG} \right] \tag{9.15}$$

We discretize the dynamic models of the process and its storage system, (9.2) and (9.3) using an implicit Euler scheme, obtaining:

$$N_j \tau (M_{i,j} - M_{i,j-1}) = \hat{M}_i - M_{i,j} \qquad \forall_{j>1}, \forall_i \tag{9.16}$$

and, respectively,

$$N_j (S_{i,j} - S_{i,j-1}) = \alpha_{i,j} - \beta_{i,j} \qquad \forall_{j>1}, \forall_i \tag{9.17}$$

The continuity conditions

$$M_{i,1} = M_{i-1,N_j} \qquad \forall_{i>1} \tag{9.18}$$

and

$$S_{i,1} = S_{i-1,N_j} \qquad \forall_{i>1} \tag{9.19}$$

are imposed at the boundaries of the N_i time slots, along with the generator ramp rate constraints

$$P_{i,j}^g - P_{i,j-1}^g \leq \frac{1}{N_j} \rho \qquad \forall_i, \forall_{j>1} \tag{9.20}$$

$$P^{g}_{i,j} - P^{g}_{i,j-1} \geq -\frac{1}{N_j}\rho \qquad \forall_i, \forall_{j>1} \tag{9.21}$$

The resulting optimization problems are linear and were implemented in GAMS (GAMS Development Corporation 2015) and solved using CPLEX (*CPLEX*, n.d.). We note here that P1 (and to a lesser extent P2) suffer from degeneracy due to the fact that the cost of adding or removing material from the storage system is not explicitly accounted for. To deal with this, we added a small regularizing term to the objective function (9.14), in the form:

$$R = -\Sigma_{i=1}^{N_i} \Sigma_{j=1}^{N_j} \chi(\alpha_{i,j} + \beta_{i,j}) \tag{9.22}$$

with the penalty χ set at 0.005 \$/(kg/h).

Figure 9.4 shows the optimal evolution of the variables relevant to the operation of the process for the two cases. Notice that the operating patterns are vastly different. On the one hand, the process-centric case produces an intuitive result, whereby the production rate is diminished to its lower bound during the peak energy price period, and stored material is used to meet demand. The inventory is replenished in the off-peak hours later in the day—a mode of operation that could potentially "shift" peak consumption to a later time in the day without achieving the desirable grid load "levelling" effect.

On the other hand, utility-centric operation maximizes utility profit by depleting the product inventory during the low demand period in the early hours of the day, where both the production rate and generation rate are set to their lower bounds (Fig. 9.5). During the peak demand and peak price in the afternoon, the generation rate is set to its maximum, while the process production rate is diminished gradually; energy demand is partially met by purchasing energy from outside the system. The production rate of the process is maintained at a high level during part of the peak energy price period to generate revenue for the utility. These observations are reflected in the results presented in Table 9.2. While the total daily energy consumption of the process for the two cases is the same (which is consistent with the requirement of meeting the same product demand and ensuring that the product inventory is replenished at the end of the day), the energy cost differs significantly, and this difference is reflected in the profit obtained by the process and the utility in the two scenarios.

Figure 9.5 also reveals a different mode of interaction between the utility system and the grid in the two scenarios: in the process-centric case, energy is purchased at a high level from the grid for a contiguous block of time during the peak demand period; on the other hand, in the utility-centric case, such purchases are limited to shorter bursts. The latter pattern is also indicative of the slight degeneracy of the problem formulation (as mentioned above). This can be attributed to the fact that changing the generation level and purchasing energy from the grid have an equivalent impact on the objective function at the operating point that the process is in between hours 12 and 18.

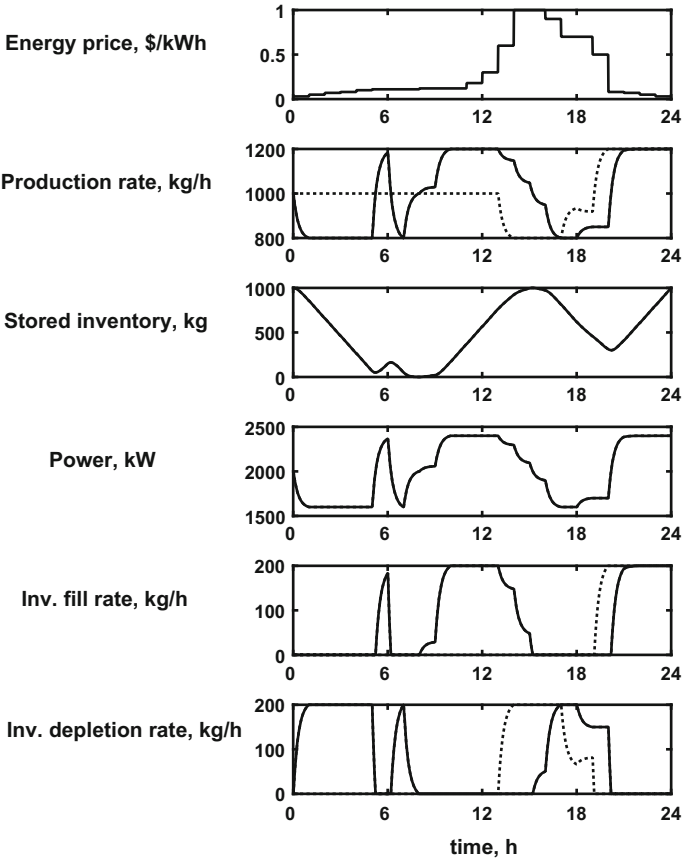

Fig. 9.4 Evolution of process variables over the scheduling horizon for the process-centric (*dotted line*) and utility-centric (*solid line*) optimization cases. The energy price profile is provided at the top as a reference

These results suggest that the operation of DR processes should be coordinated, and optimized in conjunction with the operation of the utility system, such that the operating costs of both entities are minimized. Clearly, process operators are unlikely to be willing to absorb the higher energy costs and profit loss associated with the utility-centric optimal solution obtained from P2. On the other hand, process-centric optimization of operations carried out unilaterally (the solution of P1) by each plant can lead to the advent of a second peak in grid demand ("peak shifting").

It is also noteworthy that problem P2 assumes that rather extensive information concerning the process dynamics and energy consumption/efficiency is available at the utility level; considering again the process operator perspective, it is unlikely that chemical and petrochemical companies would readily share such information with the utility supplier. Thus, defining the extent of (and need for) information sharing

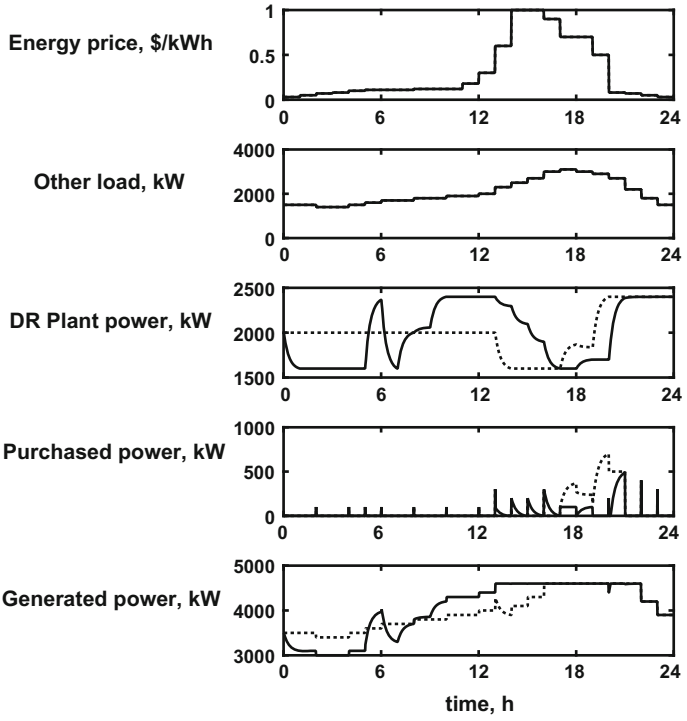

Fig. 9.5 Evolution of utility system variables over the scheduling horizon for the process-centric (*dotted line*) and utility-centric (*solid line*) optimization cases. The energy price profile and load P^o are are provided as a reference. Data for the process-centric case are obtained by solving problem P2 with the process power demand maintained fixed at the optimal profile obtained from solving P1

Table 9.2 Comparison of process profit for the process-centric and utility-centric operation optimization cases. Utility profit data for the process-centric case are obtained by solving problem P2 with the process power demand maintained fixed at the optimal profile obtained from solving P1.

	Process-centric (P1)	Utility-centric (P2)
Process profit ($)	11,112	10,171
Utility profit ($)	30,317	31,617
Energy consumed by process (kWh)	47,995	47,995
Energy cost for process ($)	12,888	13,829

(or, conversely, defining a utility-relevant model of the dynamics of process systems) is central to establishing such coordination mechanisms. Another key challenge consists of creating the means for equitably sharing the benefits (or costs) of leveling the load between the demand response entities and the utility operator.

In the following section, we address these challenges by, (i) using a battery analogy to represent the process dynamics that are of interest in scheduling the operation of the utility system and, (ii) introducing a decomposition-based coordination approach that relies on creating a new, tailored energy price/incentive structure for each demand response player, thereby driving the operation of the respective player to be utility-optimal while minimizing the impact of the change in operating paradigm on the operating cost.

9.6 A Coordination Approach for Utility-Optimal Operation of DR Process Systems

9.6.1 A Battery Representation of Process Systems

Mathematical models of batteries, irrespective of chemistry, typically aim to describe a set of variables that includes (Rao et al. 2003) the discharge behavior (state of charge, capacity fading) and the effect of temperature. We begin by noting that chemical processes acting as grid-level batteries present several advantages, including the fact that they do not exhibit capacity loss/fading (since the capacity is defined by the ability to store a material, which does not change in time), and that the "state of charge" can be measured easily as we will show below. On the other hand, the specific energy consumption ζ_k of the process (and therefore P_k) will likely be impacted by *ambient* temperature, a dependency that can be captured in the definition of this parameter.

We also note here that the battery analogy has proven successful in representing residential or commercial thermal loads in power system models (Göransson et al. 2014; Zerrahn and Schill 2015; Ali et al. 2015; O'connell et al. 2015) but we are currently unaware of any attempts at using this approach to model chemical processes engaged in DR activities.

In order to describe the behavior of a process k as a grid-level storage system, we model its operation in terms of a desired change $\Delta\hat{P}_k^p$ from the nominal power requirement of the process $\bar{P}_k^p$, which can be calculated from the nominal production rate as $\bar{P}_k^p = \zeta_k \bar{M}_k$. The rate at which ΔP_k^p can be altered in time is directly related to the process dynamics (9.2), having the same time constant τ_k as in (9.2), i.e.,

$$\tau_k \dot{\Delta P}_k^p = \Delta\hat{P}_k^p - \Delta P_k^p \qquad \forall_k \tag{9.23}$$

Then, the instantaneous power use of process k can be computed as:

$$P_k^p = \bar{P}_k^p - \Delta P_k^p \qquad \forall_k \tag{9.24}$$

These variables are then used to capture the inventory hold up of the process, which is representative of the "state of charge" of the corresponding battery:

$$\dot{Q}_k = -\frac{1}{\zeta_k}\Delta P_k^p \qquad \forall k \tag{9.25}$$

$$0 \le Q_k \le S_k^{max} \tag{9.26}$$

9.6.2 Reformulation of the Utility-Centric Operation Scheduling Problem

With the aid of the battery analogy provided above, we can reformulate the utility-centric optimal operation problem for a set of grid-connected process systems as:

$$\begin{aligned} \underset{\Delta\hat{P}_k^p, P^{mkt}, P_l^g}{max} \quad & J^{utility} \\ s.t. \quad & \text{dynamic model of process as a battery (1.23)-(1.24)} \\ & \text{state of charge model (1.25) and inventory constraint (1.12)} \\ & \text{electricity demand satisfaction (1.10)} \\ & \text{generator power calculation (1.8)} \\ & \text{ramp rate constraints (1.9)} \end{aligned} \tag{P3}$$

This formulation, which we refer to as "Problem P3", has the advantage of capturing the utility-relevant dynamics of the K processes with a lower-dimensional model. Moreover, it assumes that a minimum amount of process and production information is available at the utility level (namely, the nominal production rate, specific energy consumption and storage capacity). For now, product demand is implicitly assumed to be constant in time, an assumption that we will discuss in more detail later in the chapter.

9.7 Example (Continued)

We continue discussing the example introduced in Sect. 9.5, and solve the utility-centric operation problem while using the battery analogy to represent the dynamics and storage capabilities of the process, i.e., we solve Problem P3. The relevant results are shown in Fig. 9.6. The figure demonstrates that the state of charge/material inventory, and the pattern of electricity purchases are very similar to those obtained from solving P2.

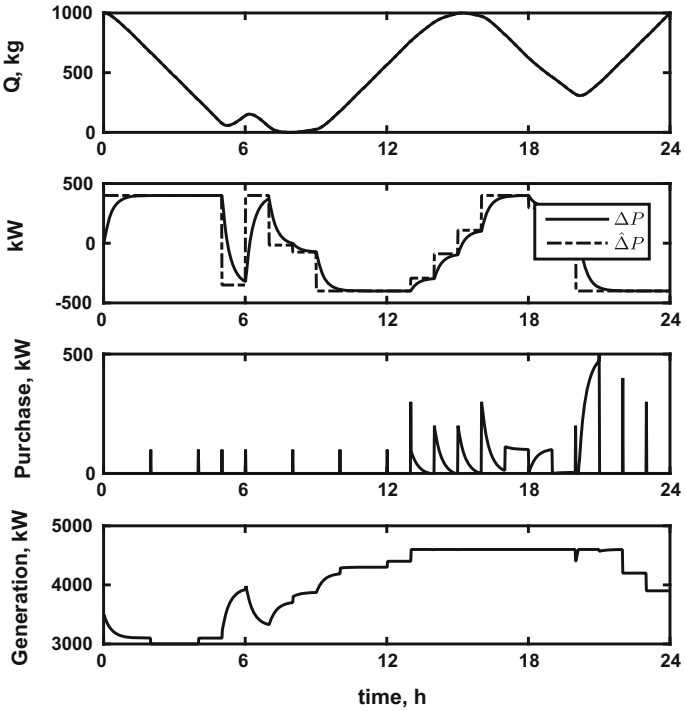

Fig. 9.6 Solutions of problem P3. *Top* evolution of state of charge Q. *Middle* change in target (*dashed*) and actual (*solid*) power consumption of the demand response process. *Bottom* electricity purchases made by the utility to satisfy demand. The corresponding results obtained from the solution of P2 are shown in dotted line where applicable

9.8 A Decomposition-Coordination Approach for Cooperation Between Demand-Response Chemical Processes and the Utility System

We propose a novel decomposition-based coordination approach, where the DR processes and the utility operators cooperate in establishing a tailored price structure (in the sense of a time-dependent daily energy price profile) for each process, such that, (i) the energy cost of each process does not differ from that corresponding to the minimum obtained when optimizing the production cost unilaterally (i.e., solving problem P1) and, (ii) the power demand profile of each process is as close as possible to the optimal profile obtained from solving the utility-centric problem P2, without disturbing product demand satisfaction by each process k.

Moreover, we restrict the exchange of information between the utility operator and process operators to the desired power demand profile (in the utility-to-process direction) and the feasible demand profile and associated price structure (in the

process-to-utility direction). These ideas are aligned with the electricity price negotiation mechanisms discussed, e.g., by Singh et al. (2015).

The proposed procedure consists of the following steps:

1. Unilaterally determine optimal values for operating costs (on the process side) and process demand profiles (on the utility side) given energy price profile π^e and outside power demand P^o, by solving solving the process-centric and utility-centric scheduling problems:

   - Solve P1 for each process k, and let C_k^{p*} be the optimal operating cost based on prices π^e, where operating cost is defined as

   $$C_k^p = \int_0^{t_p} P_k^p \pi^e dt \tag{9.27}$$

   - Solve P3, and let P_k^{p*} be the optimal power demand profile of process k resulting from the solution of this problem.

2. Identify a price incentive profile for each process k by solving the following optimization problem (which we refer to as Problem P4) at the process level:

$$\begin{aligned} \min_{\hat{\pi}_k^e, \hat{M}_k} \quad & \int_0^{t^p} (P_k^{p*} - P_k^p)^2 dt \\ s.t. \quad & \text{dynamic process model (SBM), (1.2)} \\ & \text{storage system model (1.3), (1.12)} \\ & \text{capacity and flow bounds(1.4), (1.5)} \\ & \text{demand satisfaction (1.6)} \\ & C_k^p = C_k^{p*} \end{aligned} \tag{P4}$$

   Solving P4 produces the tailored energy price structure $\hat{\pi}_k^e$ that would incentivize the operators of process k to diverge as little as possible from the power demand profile P_k^{p*} computed from P3. Note that the last constraint, $C_k^p = C_k^{p*}$, requires that process operators do so without any cost penalty to the process (relative to the scenario where energy is purchased at market rates). Also note that it is justified to assign the solution of this problem to each individual process since its formulation includes both process information (e.g., the dynamic process model) and business information (e.g., demand rates) which may not be readily shared by process operators with outside entities.

   In the ideal case, a price incentive structure can be constructed such that the time-dependent power demand of each process k computed by solving Problem P4 is the same as the utility-centric optimum obtained from P3. Equivalently, the value of the objective function of (P4) at the optimum is zero. However, there are cases when this minimum may not be reached due to, e.g., time-dependent

production rate constraints imposed on a process. Let $\tilde{P}_k^{p*}$ be the time-dependent power profile corresponding to the minimum of (P4) for process k.

3. Finalize utility operating schedule: the computed price incentive structures $\hat{\pi}_k^e$ and corresponding demand profiles, $\tilde{P}_k^{p*}$, $k = 1, \ldots, K$ are communicated to the utility. The final utility operating schedule is obtained by re-solving a modified version of Problem P3, whereby the power demand profile and price profile of all K DR processes is fixed to the values obtained from solving problems P4. We refer to this as Problem P3':

$$\begin{aligned} &\max_{P^{mkt}, P_l^g} \; J^{utility} \\ &s.t. \quad \left.\begin{array}{l} \text{dynamic model of process as a battery, (1.23)-(1.24)} \\ \text{state of charge model (1.25) and inventory constraint (1.12)} \\ \text{electricity demand satisfaction (1.10)} \\ \text{generator power calculation (1.8)} \\ \text{ramp rate constraints (1.9)} \\ \Delta\hat{P}_k^p = \bar{P}_k^p - \tilde{P}_k^{p*} \end{array}\right\} \forall_k \end{aligned} \tag{P3}$$

where the last constraint imposes that the change in power consumption required of processes k follow the trajectory $\tilde{P}_k^{p*}$.

The procedure described above is summarized in Table 9.3.

Table 9.3 Decomposition approach for coordination between DR chemical processes and the utility system

Step	Action	Inputs	Outputs
1a	Solve P1 for all K processes	Energy price π^e	Production rate target $\hat{M}^{k*}$
			Optimal operating cost C_k^{p*}
1b	Solve P3 at the utility level	Energy price π^e, Demand P^o	Generation schedule, P_l^{g*}
			Purchase schedule, P^{mkt*}
			Process demand profile P_k^{p*}
2	Solve P4 at the process level	Process demand profile P_k^{p*}, Optimal operating cost from P3 C_k^{p*}	Price incentive profile, $\hat{\pi}_k^e$
			Updated production rate target $\hat{M}_k^*$
			Updated process demand profile $\tilde{P}_k^{p*}$
3	Solve P3' to finalize utility operating schedule	Price incentive profile, $\hat{\pi}_k^e$, Process demand profile, $\tilde{P}_k^{p*}$	Updated generation schedule, P_l^{g*}
			Updated purchase schedule, P^{mkt*}

9.9 Example, Continued

We apply the coordination mechanism described above to the example system considered throughout this chapter. The quadratic program (P4) was solved using the CONOPT3 solver (*CONOPT*, n.d.). In this case, the optimal value of the objective function of problem (P4) is zero. The optimal price structure is shown in Fig. 9.7 (top). This tailored price profile is noticeably different form the real-time price profile, as is the corresponding optimal production rate profile required in order to maintain the same energy cost and profit levels throughout the day. The main discrepancies between the tailored profile and the real-time price structure are the higher off-peak prices, and the significantly lower peak price, resulting in an overall "flatter" price curve, whose maximum occurs much earlier in the day than the original price peak.

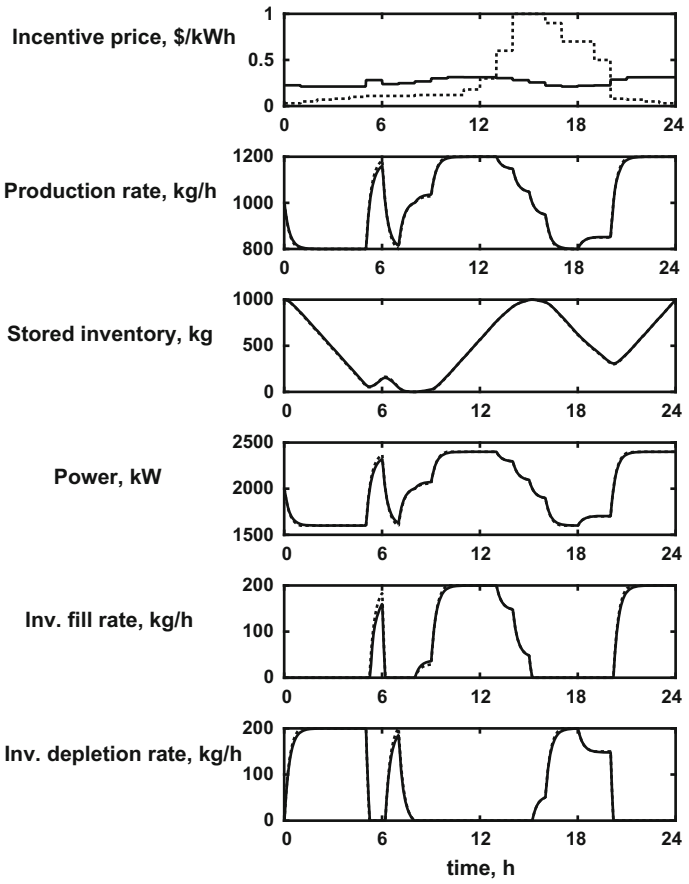

Fig. 9.7 Solution of P4, including the tailored energy price incentive structure (Subplot 1 at *top*, compared with the energy price profile π^e in *dotted line*). Subplots 2–6 show in *dotted line* the solutions of problem P2 for comparison

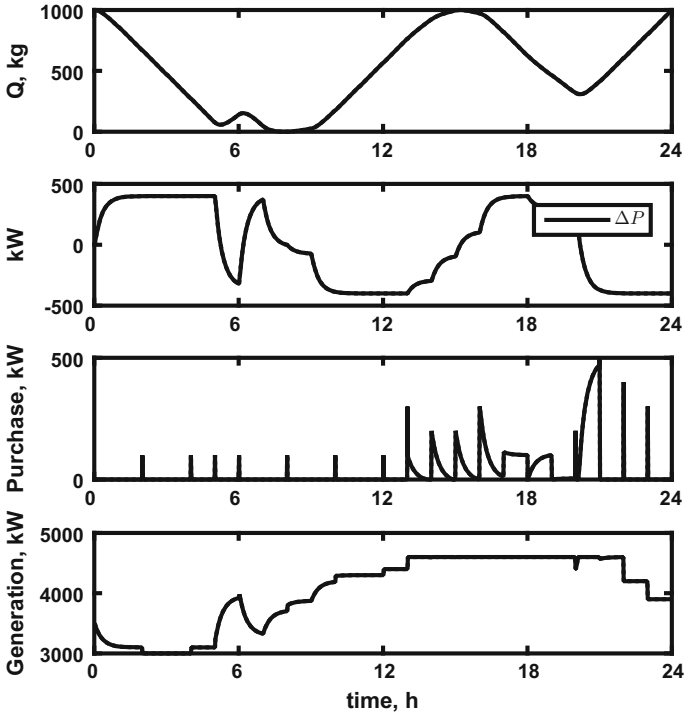

Fig. 9.8 Solution of P3'. The solution of problem P3 is shown in *dotted line* for comparison

Figure 9.8 shows that the operating schedule at the utility level, derived from solving P3' is the same as the solution of P3, which is to be expected given that the value of the objective function of P4 at the optimum is zero (i.e., the DR process can track exactly the demand profile prescribed from P3). Table 9.4 provides more insight into these results: while the process profit does not change from the value obtained from solving P1, the utility profit is lower than in the case of the utility-centric solution (P2), but higher than in the case of the process-centric solution (P1), showing the utility level benefits of properly coordinated demand-response action. The \$952 difference in utility profit between the (utopian) utility-centric value (\$31,617—see Table 9.2) and the value obtained following the coordination process (\$30,665), reflects—within the numerical tolerance of this problem– the price incentive of \$941 over the course of the day (computed as \$11,112-\$10,171 from the values of Table 9.2) offered to the DR process to follow the utility-optimal demand profile. The difference between the utility profit obtained after the coordination process and the value corresponding to the process-centric optimization, i.e., \$348, or about 1 % of profit, represents the utility-level benefit of implementing DR. Note that this profit could be partially redistributed to the chemical process operators as a further incentive to participate in DR schemes.

Table 9.4 Process and utility profit during coordinated operation

Process profit ($)	11,112
Utility profit ($)	30,665
Energy consumed by process (*kWh*)	47,995
Energy cost for process ($)	12,888

9.10 Conclusions and Perspective

In this chapter, we discussed the interplay between process-centric and utility-centric optimization of the operating schedule of demand-response chemical processes. We argued that due to the fast and frequent production rate changes required in such circumstances, the scheduling calculations must be carried out under a set of dynamic constraints that represent the scheduling-relevant closed-loop behavior of the processes. Since the solutions of the process-centric and utility-centric optimal scheduling problems may differ significantly, we proposed a reconciliation mechanism, based on two-way communication of a limited set of information between the grid and process operators. On the one hand, utility operators use a simplified, battery-like representation of the energy storage capabilities of each DR process to compute the grid-optimal process energy consumption profile, and communicate this profile to the process. Then, process operators determine the energy price profile that would allow them to implement this consumption profile without incurring economic losses, and communicate it to the utility (along with any discrepancy in the energy consumption profiles due, e.g., to changes in demand that impose production rate variations). This scheme presents several advantages, including, (i) reducing the dimension of the optimization problem that must be solved at the utility level (owing to the use of the battery analogy in modeling the DR processes) and hence reducing solution times, while (ii) offering an equitable and transparent means for chemical process operators to engage in DR schemes.

The discussion in this chapter is based on a linear approximation of the scheduling-relevant dynamics of a process, and on the assumption that the product demand of each process does not change during the day. However, these assumptions were used to preserve the clarity of the presentation and simplicity of the numerical solutions, rather than due to fundamental limitations of our approach. In effect, the proposed framework lends itself very naturally to the use of nonlinear process models and can accommodate more complex production scenarios involving multi-product processes and time-dependent production rates, and we expect that these benefits will contribute to its adoption in practice.

Acknowledgments Financial support from the National Science Foundation (NSF) through the CAREER Award 1454433 and Award CBET-1512379 is acknowledged with gratitude. The author is also grateful to Professor Ross Baldick for providing insightful comments on the problem formulation and results.

References

Ali, M., Alahäivälä, A., Malik, F., Humayun, M., Safdarian, A., & Lehtonen, M. (2015). A market-oriented hierarchical framework for residential demand response. *International Journal of Electrical Power & Energy Systems*, *69*, 257–263.

Baldea, M., & Daoutidis, P. (2012). *Dynamics and nonlinear control of integrated process systems*. Cambridge: Cambridge University Press.

Baldea, M., Park, J., Du, J., & Harjunkoski, I. (2015). Integrated production scheduling and model predictive control of continuous processes. *AIChE Journal, 61*(12), 4179–4190.

Cao, Y., Swartz, C. L. E., Baldea, M., & Blouin, S. (2015). Optimization-based assessment of design limitations to air separation plant agility in demand response scenarios. *Journal of Process Control, 33*, 37–48.

CONOPT. (n.d.). Retrieved August 3, 2014, from http://www.gams.com/dd/docs/solvers/conopt.pdf.

CPLEX. (n.d.). Retrieved August 3, 2014, from http://www.gams.com/dd/docs/solvers/cplex.pdf.

Du, J., Park, J., Harjunkoski, I., & Baldea, M. (2015). Time scale bridging approaches for integration of production scheduling and process control. *Computers & Chemical Engineering*, *79*, 59–69.

Feng, J., Brown, A., OBrien, D., & Chmielewski, D. J. (2015). Smart grid coordination of a chemical processing plant. *Chemical Engineering Science, 136*, 168–176.

FERC. (2013). https://www.ferc.gov/industries/electric/indus-act/demand-response/dem-res-adv-metering.asp.

GAMS Development Corporation. (2015). *General Algebraic Modeling System (GAMS) Release 24.5.6*. Washington, DC, USA. http://www.gams.com/.

Garcia, H. E., Mohanty, A., Lin, W. C., & Cherry, R. S. (2013). Dynamic analysis of hybrid energy systems under flexible operation and variable renewable generation-part I: Dynamic performance analysis. *Energy*, *52*, 1–16.

Göransson, L., Goop, J., Unger, T., Odenberger, M., & Johnsson, F. (2014). Linkages between demand-side management and congestion in the european electricity transmission system. *Energy*, *69*, 860–872.

Ierapetritou, M. G., Wu, D., Vin, J., Sweeney, P., & Chigirinskiy, M. (2002). Cost minimization in an energy-intensive plant using mathematical programming approaches. *Industrial & Engineering Chemistry Research*, *41*(21), 5262–5277.

Maravelias, C. (2012). General framework and modeling approach classification for chemical production scheduling. *AIChE Journal*, *58*(6), 1812–1828.

Middelberg, A., Zhang, J., & Xia, X. (2009). An optimal control model for load shifting-with application in the energy management of a colliery. *Applied Energy*, *86*(7), 1266–1273.

O'Connell, N., Hale, E., Doebber, I., & Jorgenson, J. (2015). *On the inclusion of energy-shifting demand response in production cost models: Methodology and a case study*, Technical report, National Renewable Energy Laboratory (NREL) report NREL/TP-6A20-64465.

Ondeck, A. D., Edgar, T. F., & Baldea, M. (2015). Optimal operation of a residential district-level combined photovoltaic/natural gas power and cooling system. *Applied Energy*, *156*, 593–606.

Pattison, R. C., & Baldea, M. (2014). Optimal design of air separation plants with variable electricity pricing. *Foundations of Computer Aided Process Design (FOCAPD)* (pp. 393–398). WA: Cle Elum.

Pattison, R. C., Touretzky, C. R., Johansson, T., Harjunkoski, I. & Baldea, M. (2016a). Moving horizon scheduling of an air separation unit under fast-changing energy prices. In *Proceedings of DYCOPS 2016*. Trondheim, Norway.

Pattison, R. C., Touretzky, C. R., Johansson, T., Harjunkoski, I., & Baldea, M. (2016b). Optimal process operations in fast-changing electricity markets: Framework for scheduling with low-order dynamic models and an air separation application. *Industrial & Engineering Chemistry Research*. http://dx.doi.org/10.1021/acs.iecr.5b03499.

Paulus, M., & Borggrefe, F. (2011). The potential of demand-side management in energy-intensive industries for electricity markets in Germany. *Applied Energy*, *88*(2), 432–441.

Rao, R., Vrudhula, S., & Rakhmatov, D. N. (2003). Battery modeling for energy aware system design. *Computer*, *36*(12), 77–87.

Siirola, J. J., & Edgar, T. F. (2012). Process energy systems: control, economic, and sustainability objectives. *Computers & Chemical Engineering, 47*, 134–144.

Singh, R., Ma, K., Thatte, A. A., Kumar, P. R., & Xie, L. (2015). A theory for the economic operation of a smart grid with stochastic renewables, demand response and storage. In *54th IEEE Conference on Decision and Control (CDC)* (pp. 3778–3785).

Soroush, M., & Chmielewski, D. J. (2013). Process systems opportunities in power generation, storage and distribution. *Computers & Chemical Engineering*, *51*, 86–95.

Touretzky, C. R. (2016). Embedding dynamics and control consideration in operational optimization of process and energy systems. Ph.D. Thesis, The University of Texas at Austin.

Touretzky, C. R., & Baldea, M. (2016). A hierarchical scheduling and control strategy for thermal energy storage systems. *Energy and Buildings*, *110*, 94–107.

US Energy Information Administration (2013a). *Form EIA 411, Coordinated Bulk Power Demand and Supply Report*.

US Energy Information Administration. (2013b). *The Electric Power Monthly*.

Van Staden, A. J., Zhang, J., & Xia, X. (2011). A model predictive control strategy for load shifting in a water pumping scheme with maximum demand charges. *Applied Energy*, *88*(12), 4785–4794.

Wattles, P. (2012). *ERCOT overview*. Retrieved January 18, 2016, from http://www.ercot.com/content/news/presentations/2012/ERCOT_SmartEnergySummit_2-28-12.pdf.

Zerrahn, A., & Schill, W. P. (2015). On the representation of demand-side management in power system models. *Energy*, *84*, 840–845.

Zhang, Q., Grossmann, I. E., Heuberger, C. F., Sundaramoorthy, A., & Pinto, J. M. (2015). Air separation with cryogenic energy storage: Optimal scheduling considering electric energy and reserve markets. *AIChE Journal*, *61*(5), 1547–1558.

Zhu, Y., Legg, S., & Laird, C. D. (2011). Optimal operation of cryogenic air separation systems with demand uncertainty and contractual obligations. *Chemical Engineering Science*, *66*(5), 953–963.

Chapter 10
Modelling and Analysis of China's Passenger Car Fuel Consumption up to 2030

Zheng Zhao, Pei Liu and Zheng Li

Abstract In this study, we present an analysis of future gasoline supply and consumption gap and a projection to future passenger car fuel consumption. In the first part, we study the gasoline and diesel supply from the refining sector in China. A virtual refinery model is established to analyze the productivity of gasoline and diesel. Based on the calculation of the model, we get minimum and maximum production of gasoline and diesel and the flexibility between them. In the second part, we study the ownership and sales of passenger car in China. A model for different types of passenger cars is established and many factors affecting the ownership are analyzed separately. Vehicle age distribution in passenger car is also taken into consideration. In the third part, alternative fuels and technologies for vehicles are studied. Fuel ethanol, natural gas, EV and PHEV are analyzed in detail. In the last part, we analyze the actual gasoline consumption rate and the average annual mileage of passenger car. Then we make an assumption as a basic scenario. Gasoline consumption in passenger car sector and other sectors in the future are calculated based on the assumption above. Gasoline consumption and supply are compared and a major finding is that the gasoline supply cannot meet the consumption since 2022 in basic scenario. To meet this gap, different measures in technology and policy which reducing gasoline consumption in passenger car sector are considered and analyzed.

10.1 Forecast of Gasoline and Diesel Yield from China's Refining Sector

10.1.1 Background

In China, most gasoline and diesel are obtained from the oil refinery industry. As shown in Figs. 10.1 and 10.2, annual gasoline and diesel production and consumption in China are roughly in balance. This also indicates that the capacity and

Z. Zhao · P. Liu (✉) · Z. Li
Department of Thermal Engineering, Tsinghua University, Beijing, China
e-mail: liu_pei@tsinghua.edu.cn

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_10

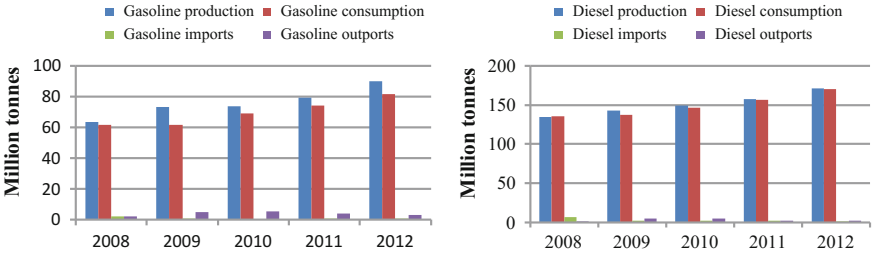

Fig. 10.1 Gasoline and diesel production and consumption in China, between 2008 and 2012

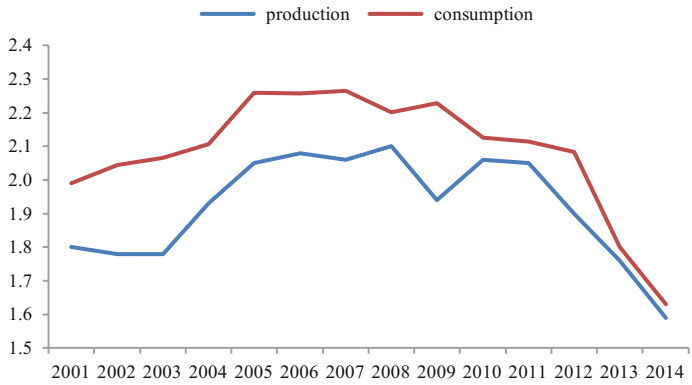

Fig. 10.2 Diesel to gasoline rate of production and consumption in China, between 2001 and 2014

operational flexibility of the refining sector are rather critical to the future balance (or imbalance) between supply and demand. Many studies have shown that the demand for gasoline and diesel will increase greatly with rapid increase of vehicle population (Ma et al. 2012). In this context, whether the refining industry in China can meet this increasing demand, or to what extent it can meet this demand, remains a question. In this part, we aim to tackle this problem via analyzing the development of the refining industry.

The production of gasoline and diesel in the refining sector depends on two factors: one is the amount and the property of the crude. The other is the productivity of gasoline and diesel in refinery. The crude processed in the refining sector comprises two parts: domestic and imports. In recent years, domestic produced crude oil production is relatively stable, around a level of 200 million tons per year. Imported crude oil amount increases with China's oil demand every year. In 2013, imported oil accounted for 57 % of China's total oil supply (NBSC). From the view point of energy security, oil import dependency should not be too high, and there should be an upper limit for oil import. In this work, this upper limit for oil import is set to be 70 % (Wang 2013), i.e., at maximum 70 % of China's oil can come from oil import.

The production of gasoline and diesel in the refining sector also depends on the structure of a refinery, its operational conditions, and crude oil property. The production rate can change greatly under different circumstances. For instance, a petrochemical refinery produces more chemical light oil and naphtha, but less gasoline or diesel. A fuel refinery is quite on the opposite side, where gasoline and diesel are its major products (Shen 2009).

At the national level, production rate of gasoline and diesel also changes year by year, mainly following their demand. Figure 10.2 indicates that the production range of the refining industry is not fixed. Instead, it exhibits an obvious wide range for operation. Therefore, quantifying this range is critical to calculating future supply capacity of gasoline and diesel.

10.1.2 Virtual Refinery Model

The refining process in China can be divided into three parts: primary processing, secondary processing and extended processing. Primary processing includes atmospheric and vacuum distillation, which is also the major part of refining. Secondary processing includes many types, including catalytic cracking, delay coking, hydro cracking, catalytic reforming and hydro treating. All or some of these technologies can be used for different production purposes and for processing crude with different properties. Extended processing is a deepening processing procedure, mainly to produce chemicals. As we mainly aim at the productivity of gasoline and diesel, only primary and secondary process are considered in this part.

Taking all the primary and secondary processing technologies into consideration, we establish a virtual refinery model at the national level, in which the capacity of each unit is the summation of corresponding unit capacity in the whole country. The connection between units includes all feasible connections based on actual refining process. Figure 10.3 shows the structure and the details of the model (Li 2013).

10.1.2.1 Model Structure and Formulation

A superstructure model is constructed to cover all possible physical flows, oil products, and processing routes. The model obeys the basic conservation law of mass. For each unit, the total mass of inputs and outputs are in balance. While the production rate of different products in different unit is different. The value can change in a certain range. The actual flow routes of products in each unit are also changeable in all the possible routes. Different value represents different operation condition. Considering so many variables above, when the crude volume, capacity of each unit and the demand for end products except gasoline and diesel are given, the minimum and maximum gasoline and diesel production and the flexibility can be calculated.

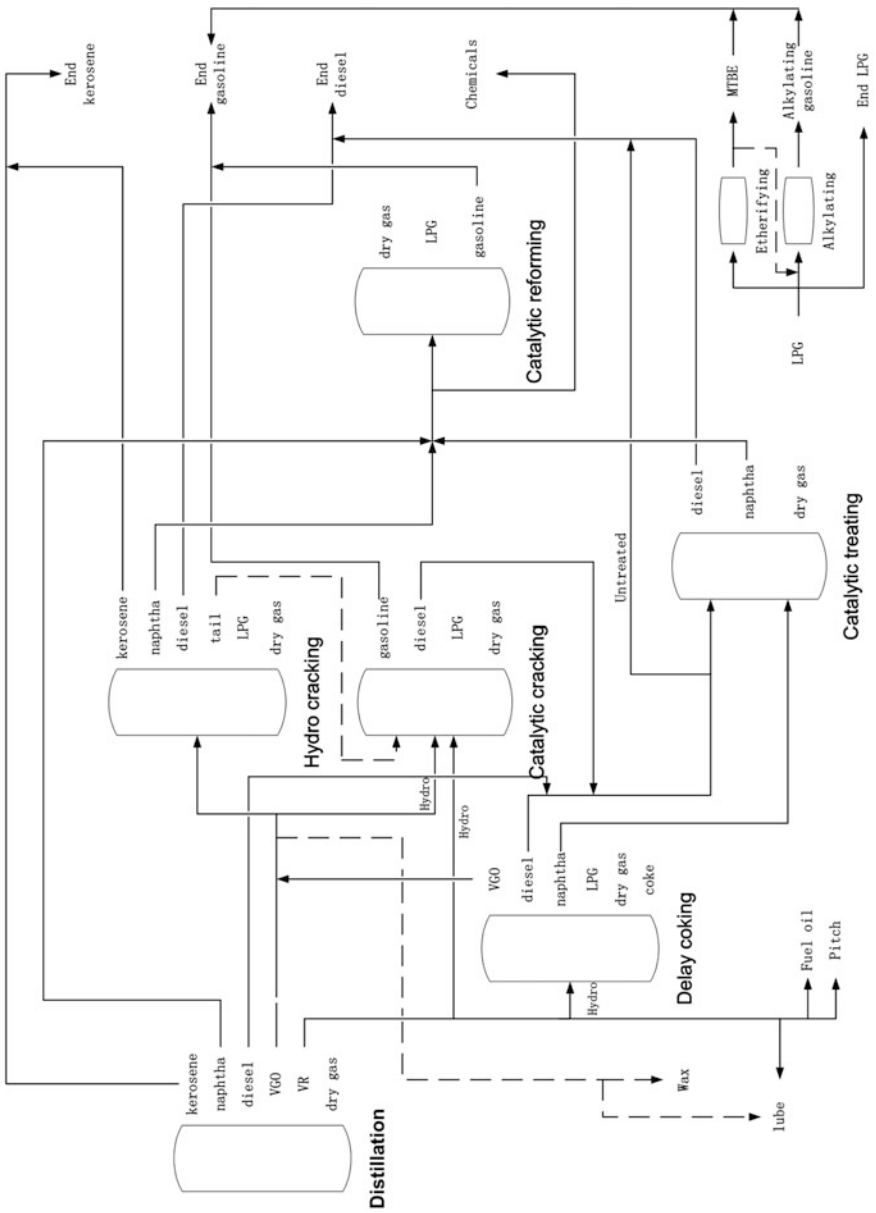

Fig. 10.3 The structure of virtual refinery model at the national level in China

10.1.2.2 Model Settings for Primary Processing

In primary processing, atmospheric and vacuum distillation is considered. The raw material of distillation is the crude oil. Its main liquid products are naphtha, diesel, kerosene, vacuum gas oil (VGO), vacuum residuum (VR) and liquefied petroleum gas (LPG). Naphtha from distillation has a low octane number, about 40−70. It is usually used as chemical or reforming raw material. Diesel and kerosene (D&K) contains impurities, so hydro treating is needed for D&K here. In D&K production, we assume that kerosene accounts for 10–30 % (Shen 2009; Hou 2011; Chen 2007; Li 2012; Li 2009; Hou 2005). VGO and VR need further processing.

The fraction of distillation depends on crude property. Crude oil in China's refining process comprises two parts: domestic crude and import crude. The property of each kind of crude is quite different. For domestic crude, which is produced in a limited number of giant oil fields, the property is relatively stable. According to the crude oil property data in domestic oil fields (Lu 2009), we get the fraction of distillation for domestic crude. For import crude, the source is relatively stable (Huang 2001). So we use source of import crude in 2014 as the criterion and get the fraction rate of import crude oil.

Based on above methods, as shown in Fig. 10.4, we get the fraction of distillation for both domestic and import crude. It can be seen that domestic crude is heavier than imports, more VGO and VR but less naphtha and D&K.

It is assumed that the production of domestic crude maintains at 210 million tons since 2015 (BP 2014 Qian and Jiang 2014). The upper limit of oil import dependency is 70 %. Then we get the proportion of domestic and import crude in the future. Based on the above data, the fraction of mixing crude from 2015 to 2030 can be calculated. As Table 10.1 shows, with oil import dependency higher, the share of naphtha and D&K rises, this is beneficial to produce more gasoline and diesel.

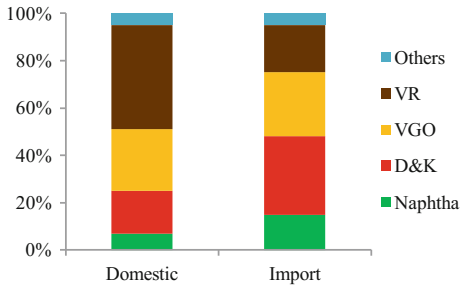

Fig. 10.4 Distillation fraction rate of domestic and import crude of China in 2014

Table 10.1 Fraction rate of mixing crude

	Naphtha (%)	D&K (%)	VGO (%)	VR (%)
2015	11.7	26.9	26.6	32.4
2020	12.1	27.6	26.6	31.3
2025	12.3	27.9	26.7	30.8
2030	12.6	28.5	26.7	29.9

10.1.2.3 Model Settings for Secondary Processing

Delay coking is a key process to transform VR, the share of diesel in products is high, but because of much impurity, diesel here should be hydro treated. Catalytic cracking is the main process in secondary processing. The share of cracking gasoline is very high, and with high octane number, it can be directly used as gasoline components. Catalytic cracking is an important process to produce gasoline. Hydro cracking is a key process to produce high quality oil products. One advantage of this process is the high quality of the products, with which hydro treating is not needed. Naphtha here can be directly used as high quality reforming raw materials. D&K can be directly used as components of end products. The other advantage of this process is the high flexibility. The range of its productivity can be widely adjusted. Its disadvantage is the high cost. Catalytic reforming is a key process to produce high octane number gasoline components and aromatics. Hydro treating is a key process to improve quality of oil products. Its processing object is rather wide. High octane Gasoline processing is a key process to produce high octane gasoline components. The capacity of this technology increases rapidly but still small. Its raw materials are C_4, mainly from LPG. Raw materials shortage is the biggest problem for this technology. Since its products are components of end gasoline product, the process is considered in the model (Shen 2009; Hou 2011; Chen 2007; Li 2012; Li 2009; Hou 2005). Table 10.2 shows the parameters of main processes in secondary processing.

10.1.2.4 Model Settings for Capacity of Different Process

For primary processing capacity, it will reach 940 million tons in 2030 (Qian and Jiang 2014). For secondary processing capacity, we do extrapolation based on average annual growth rate of primary and secondary processing capacity from 2010 to 2014. Figure 10.5 shows the calculated forecast of capacity for different process from 2015 to 2030 by extrapolation.

10.1.2.5 Model Settings for Products

Oil refining products are mainly divided into the following categories: light fuels (gasoline, kerosene and diesel), chemicals (naphtha and solvent), heavy oil products (lubricant, coke, wax, fuel oil and pitch) and C_1–C_4 products (LPG and chemical gas). Since other oil refining products, besides gasoline and diesel, can also affect the production range of the refining industry, it is also necessary to forecast their production rates. Table 10.3 shows the demand of other oil products based on existing research (Shi and Bai 2014; Cai and Liu 2011; Kong et al. 2014; Wang and Zhu 2013; Huang 2013; Fan 2012; An et al. 2015; Shi et al. 2012; Chi 2014). For ease of calculation, we assume that the production of other oil products is equal to the demand of corresponding products in the model.

Table 10.2 Parameters of refinery model in secondary processing

Delay coking	Raw materials:	VR
	Main products:	Naphtha: 10–16 % Diesel: 24–34 % VGO: 20–30 % Naphtha + Diesel + VGO = 65 % Coke: 25 % LPG: 5 %
Catalytic cracking	Raw materials:	VGO, VR
	Main products:	Gasoline: 45–55 % Diesel: 20–30 % Gasoline + Diesel = 75 % LPG: 20 %
Hydro cracking	Raw materials:	VGO
	Main products:	Naphtha: 10–65 % Kerosene: 10–60 % Diesel: 20–70 % Tail: 5–20 % Naphtha + Kerosene + Diesel + Tail = 95 % LPG: 5 %
Catalytic reforming	Raw materials:	Naphtha
	Main products:	Gasoline: 70 % Aromatic hydrocarbon: 30 %
Hydro treating	Raw materials:	Distilled naphtha, Distilled kerosene, Distilled diesel Coker naphtha, Coker diesel Catalytic gasoline, Catalytic diesel Delay coking raw materials: VR Catalytic cracking raw materials: VGO + VR

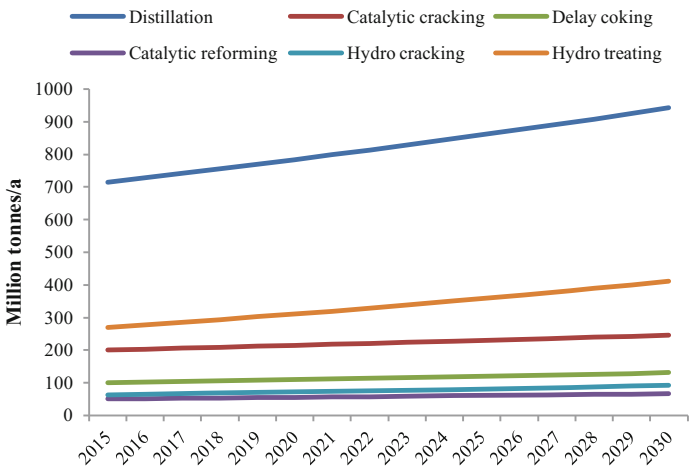

Fig. 10.5 Forecast of different refinery unit capacity, between 2015 and 2030

Table 10.3 Minimum demand of products apart from gasoline and diesel (million tons)

	2015	2020	2025	2030
Kerosene	32	50	75	100
Heavy	61	62	63	65
Chemicals	33	36	40	44
LPG	29	31	33	35

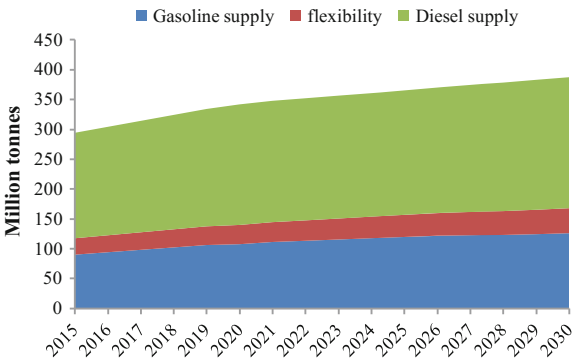

Fig. 10.6 Production of gasoline and diesel in China, between 2015 and 2030

10.1.3 Results of Modeling Calculation

Based on the above virtual refining model and modelling settings, we get minimum and maximum production of gasoline and diesel and the flexibility between the two products. As Fig. 10.6 shows, the green bar denotes the minimum production of gasoline, the blue bar denotes the minimum production of diesel and the red bar denotes the flexibility between gasoline and diesel. Summation of the three is total production of gasoline and diesel. Figure 10.6 indicates that when the production of other oil products is fixed, the total production of gasoline and diesel is fixed. But there is still flexibility between gasoline and diesel. The flexibility of productivity maintains at about 5 %. That means when refining structure is fixed and the production of other products is also fixed, flexibility of the production range is rather small only by change operational modes.

10.2 Analysis of Ownership and Sales of Passenger Car

10.2.1 Vehicle Classification

In this section, we present a model for studying ownership and sales of passenger car. This model and result are the basis of studying fuel consumption of passenger car sector. According to different functions of civil vehicles and based on the classification of Chinese Bureau of Statistics, civil vehicles are divided into

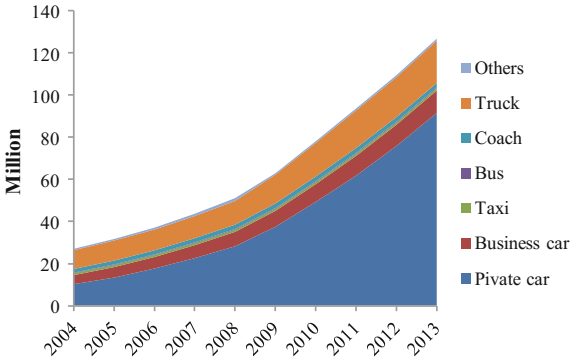

Fig. 10.7 Vehicle ownership in China, between 2004 and 2013

different categories. For passenger car, we further divide the three categories: private car, business car and taxi. The growth of private passenger car is much faster than other types of vehicles (Fig. 10.7). Since each type of passenger car has different ownership and annual mileage, such classification is necessary (Kobos et al. 2003; Ou et al. 2010; Han et al. 2011).

10.2.2 Modeling for Passenger Car Ownership

10.2.2.1 Private Car Ownership

Whether a person has a private passenger car or not mainly depends on his or her purchasing ability and purchasing desire. In this section, we use annual disposable income to represent the purchasing ability. Figure 10.8 shows the relationship

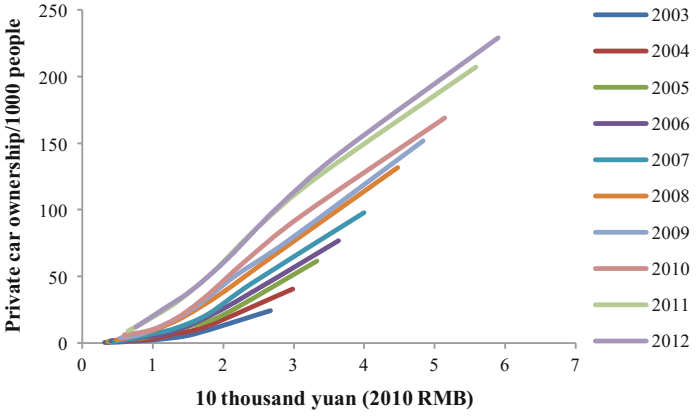

Fig. 10.8 Relationship between private car ownership per thousand people and annual disposable income, between 2003 and 2012

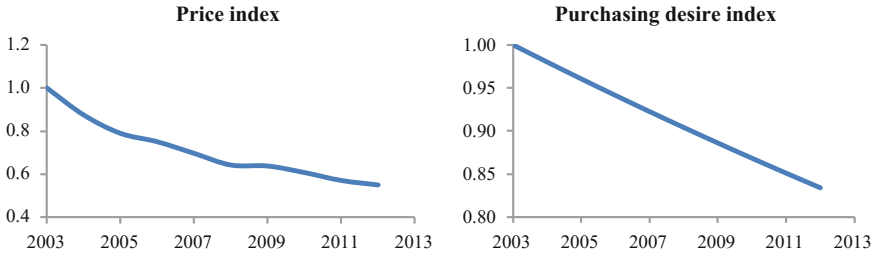

Fig. 10.9 Price index and purchasing desire index of private car in China based on 2003 case

between private passenger car ownership per thousand people and annual disposable income. It indicates that private passenger car ownership per thousand people increases with annual disposable income growing. But in different year, private passenger car ownership per thousand people is different under the same annual disposable income. It increases every year. That is because with the price of private passenger car declining, purchasing ability for cars improves, and because of demonstration effect of buying cars, purchasing desire of people enhances.

In order to reflect the influence of car price and purchase desire in the model, two indices are introduced, namely price index (*PI*) and purchase desire index (*PDI*). Figure 10.9 shows the index with the change of year.

Considering that car price declining is more sensitive to medium income group, we use normal distribution function $D(x)$ to describe this effect:

$$D(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}, \mu = 2.5, \sigma = 3.5, x = income \tag{10.1}$$

Taking effect of car price and purchasing desire into consideration, we use price index (*PI*), purchasing desire index (*PDI*) and $D(x)$ to modify annual disposable income:

$$Modified\_income(x,t) = \frac{x}{(1-(1-PI(t)) \times D(x)) \times PDI(t)}, x = income, t = year \tag{10.2}$$

Figure 10.10 shows the relationship between private passenger car ownership per thousand people and modified income (MI). We can see that the correlation of the data is good. Private passenger car ownership per thousand people increase as MI grows. When MI increases to a certain degree, the growth of private passenger car ownership per thousand people decreases and tend to a fixed value. According to existing researches (Han et al. 2011; Shen 2006), we can use Gompertz function to fit this variation:

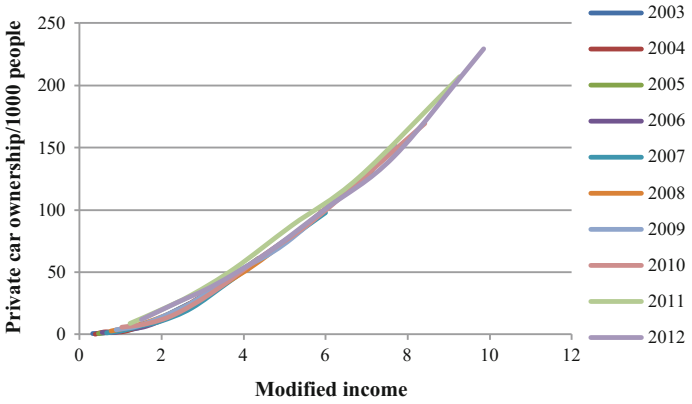

Fig. 10.10 Relationship between private car ownership per thousand people and modified income, between 2003 and 2012

$$S(y) = ae^{(-be^{-cy})} \tag{10.3}$$

a means the saturation value, b and c affect the shape of the curve.

To calculate private passenger car ownership per thousand people, we need to get the distribution of annual disposable income. The distribution of annual disposable income in China obeys the law of lognormal distribution (Han et al. 2011; Shen 2006):

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{\left(-\frac{(\ln(x)-u)^2}{2\sigma^2}\right)} \tag{10.4}$$

σ and μ depend on per capita disposable income α and Gini index G:

$$\alpha = e^{\left(u+\frac{1}{2}\sigma^2\right)}, G = 2F\left(\frac{\sigma}{\sqrt{2}}|0,1\right) - 1 \tag{10.5}$$

Based on the analysis above, the model for private passenger car ownership is described as follows:

$$\begin{aligned} &Onwership(t)_{Private} = \frac{TP(t)}{1000}\int_{x=0}^{\infty} F(x)S(MI(x))dx \\ &x = income(yuan), TP = population, t = year \end{aligned} \tag{10.6}$$

According to the data from National Bureau of Statistics, we can get urban and rural income distribution in China between 2004 and 2013. Considering to the cases in Japan and Europe, we assume the saturation value of Gompertz function is 400 in China. Then we can calculate private car ownership with our model. Figure 10.11

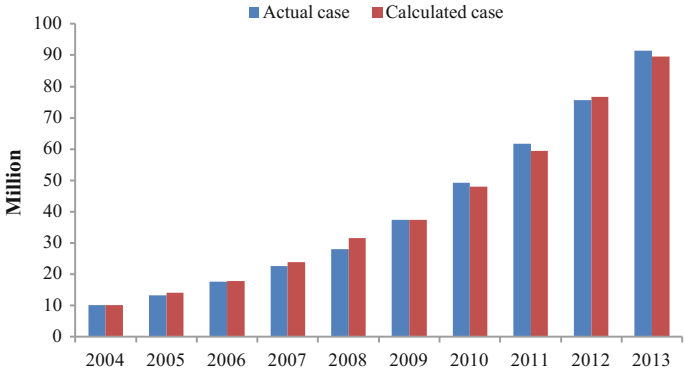

Fig. 10.11 Actual and calculated private car ownership, between 2004 and 2013

shows the comparison between results of calculation from the model and actual data from 2004 to 2013. We can see that the two sets of data are basically identical, which indicates the model is reliable.

10.2.2.2 Business Car Ownership

Business passenger cars are owned by government and public institutions, the main use of business passenger car is for official business. The growth of business car ownership is different from that of private car. As shown in Fig. 10.11, it can be seen that there is good linear correlation between business car ownership and GDP. So the model for business car ownership can be established by linearly fitting (Fig. 10.12).

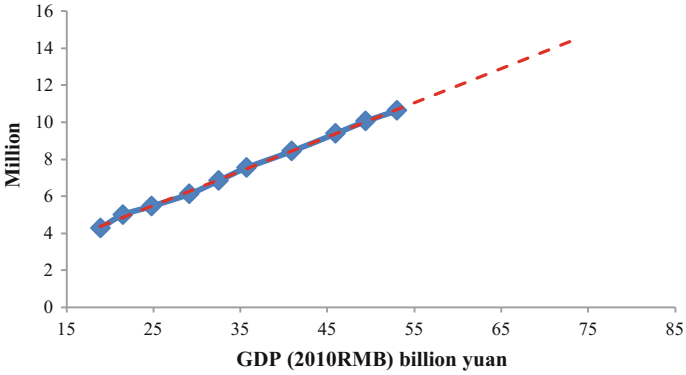

Fig. 10.12 Relationship between business car ownership and GDP in China

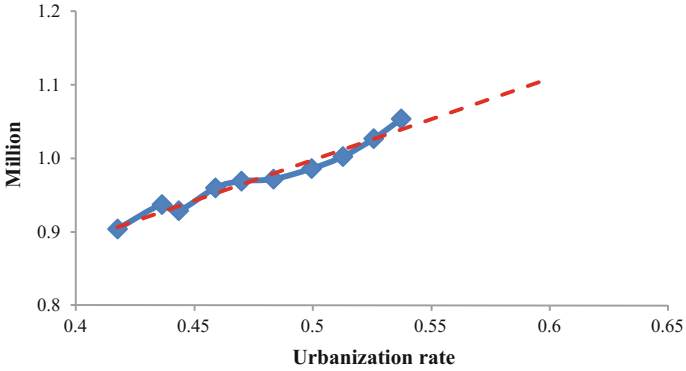

Fig. 10.13 Relationship between taxi ownership and urbanization rate in China

10.2.2.3 Taxi Ownership

Taxi ownership is controlled based on local policies. Figure 10.13 shows the relationship between taxi ownership and urbanization rate. Overall, taxi ownership and urbanization rate is positively correlated. Since taxi ownership change is relatively small, we assume that taxi ownership and urbanization rate meet the linear relationship. So the model for taxi ownership can be established by linearly fitting.

10.2.2.4 Modeling Summary

In the model for passenger car ownership, the input parameters are: population, per capita disposable income, Gini index, GDP and urbanization rate. Provided forecast of these parameters, ownership of passenger car in the future can be calculated with the model established above.

10.2.3 Modeling for Vehicle Age Distribution

Considering technology progress and policy adjustment, fuel economy of passenger cars produced in different years may be different. Improvements of fuel economy and the use of alternative fuels can only completed by cars replacement. So the survival rule of passenger car should be studied. Based on existing researches (Huo and Wang 2012), we can use Weibull distribution to describe the survival rule of passenger car:

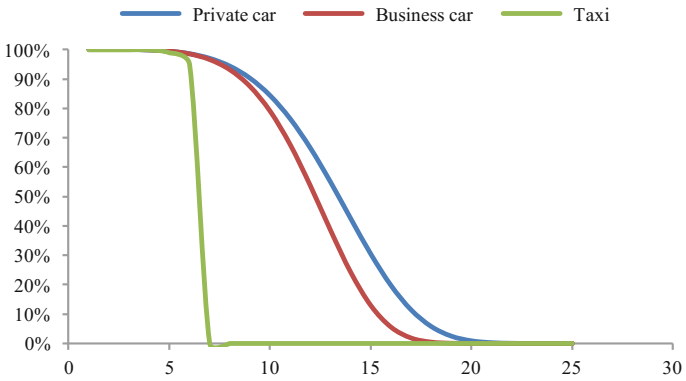

Fig. 10.14 Passenger car survival rate in China

$$SR(x) = e^{-\left(\frac{x}{T}\right)^k} \tag{10.7}$$

It indicates the car proportion of survival after serving x years. k and T are parameters depended on the type of cars. Figure 10.14 shows the survival curve of passenger car in China, we can see that private car has the longest survival time, after 7 years serving, it gradually begins to scrap. Business car is the second and similar to the private. Taxi scraps sharply when serving fixed years, since China implements a force scrap rule for taxi. In this section, we assume that after 6 years serving, taxi scraps directly (Huo and Wang 2012).

In this section, we assume that car survival rate doesn't change by year in the future. Based on the survival curve and ownership, we can calculate sales of passenger car in each year.

10.2.4 Forecast of Ownership and Sales of Passenger Car

According to existing researches and linear extrapolation, we get the input data of the model: population, GDP, per capital income and Gini index in China in the next 15 years. Based on our model and the input data, we can calculate passenger car ownership between 2015 and 2030. The result shows in Fig. 10.15.

We can see that there will be a great increase in passenger car ownership in China in the next 15 years. Most of the increase is contributed by the fast growth in private car. Business car and taxi account for a little percentage of the total ownership in passenger car sector. This result of passenger car ownership represents the total demand of passenger car in the future. According to the analysis of the vehicle age distribution in this part, if it is assumed that there are only gasoline cars in the future, we can get passenger car sales in each year in the future. But this assumption is unreasonable since vehicle alternative fuels are developing so fast that they might

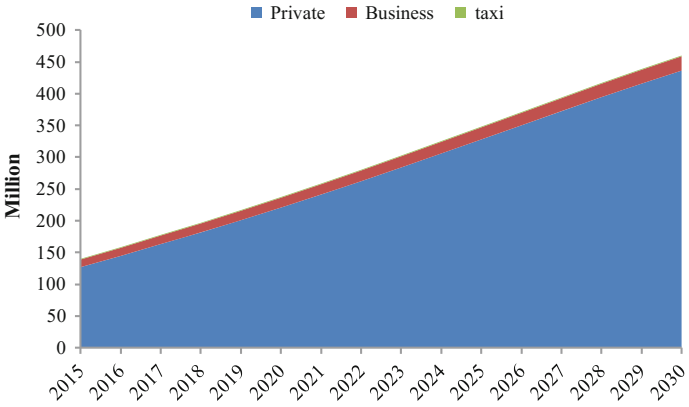

Fig. 10.15 Passenger car ownership forecast in China between 2015 and 2030

account for a certain percentage of the total passenger car ownership in the future which means sales of passenger car should be calculated separately by different fuels. So it is necessary to analyze the vehicle alternative fuels development in the next 15 years in China.

10.3 Alternative Fuel Analysis

10.3.1 Overall Analysis of Vehicle-Use Alternative Fuels

Using the model in part 3, sales and ownership of passenger car in the coming 15 years can be calculated. But this forecast can only tell us total demand of passenger cars in the future. We don't know how many of these passenger cars in each year are driven by gasoline or by other alternative fuels. So in order to calculate accurate gasoline consumption of passenger car sector in China, the utilization of vehicle-use alternative fuels in the future should be analyzed and forecasted quantitatively.

So far there are many kinds of vehicle-use alternative fuels have been commercialized or in demonstration in the world. According to the type of fuels, we can divide vehicle-use alternative fuels into 3 classes: biomass fuel which mainly includes ethanol and methanol, natural gas and power. In the power driven class, from the aspect of vehicle technology, there are electric vehicle (EV), plug-in hybrid electric vehicle (PHEV) and fuel cell electric vehicle (FCV). The developments of vehicle-use alternative fuels are quite different in different countries. In China, fuel ethanol was demonstrated in 2001 and the production has reach to a certain degree, though still not large. In 2014 the production of fuel ethanol is about 2.3 million tonnes. Fuel methanol has a negative prospect after 10 years demonstrating in Shanxi province. There is no potential of development for it in China in

the future. Natural gas has been widely used in taxies and buses in China in recent years. In some provinces many private passenger cars have choose to use natural gas for its low price. In 2014 ownership of natural gas cars has reach to about 2.7 million. For EV and PHEV, there are many encouraging policies and financial support. Clear plans of development are also made for them. For example, the government put forward that the ownership of EV and PHEV should reach to 5 million by 2020 in *Energy saving and new energy vehicles development plan*. For FCV, there is only support for the research but no clear policy for commercial promotion in China. So there is little potential of development for it in China in the next 15 years. In this chapter, we will only consider the development of ethanol, natural gas, EV and PHEV in China and get a result of forecast for each of them.

10.3.2 Analysis of Fuel Ethanol

According to different raw materials, there are 3 types of fuel ethanol. The ethanol produced from grain plants is the 1 generation (1G) ethanol. The main raw materials of the 1G ethanol are wheat and maize. In China the 1G ethanol producing technology is mature and has been commercialized. Most fuel ethanol which produced in these years is the 1G ethanol. The ethanol produced from non-grain plants which mainly include cassava, sweet potato and sweet sorghum is the 1.5 generation (1.5G) ethanol. In China cassava ethanol has just started to be commercialized, but the capacities of sweet potato ethanol and sweet sorghum are very small, the production technology is still in demonstration. The ethanol which produced from agricultural and forestry residues is the 2 generation (2G) ethanol which just begin demonstration in China (Huo 2012).

In order to forecast total production of ethanol in China, we need to analyze the 3 types of ethanol separately. Main companies for ethanol in China and their capacities are listed in Table 10.4. We can see that most ethanol produced in China is the 1G ethanol. Capacities of the 1.5 and 2G ethanol are very small.

Table 10.4 The Capacity of main companies for ethanol in China (Lei 2015)

Year	Main companies for ethanol in China	Feedstock	Capacity (thousand tonnes)
2011	Henan Tianguan	Maize/Cassava	500
	Jilin ethanol	Maize	550
	Anhui Fengyuan	Wheat/Maize	450
	COFCO biochemical in Zhaodong	Maize	250
	COFCO biomass energy in Guangxi	Cassava	200
2012	Henan Tianguan	Cellulose	10
2013	Longli	Cellulose	50
2014	Zhongxing energy	Sweet sorghum	100

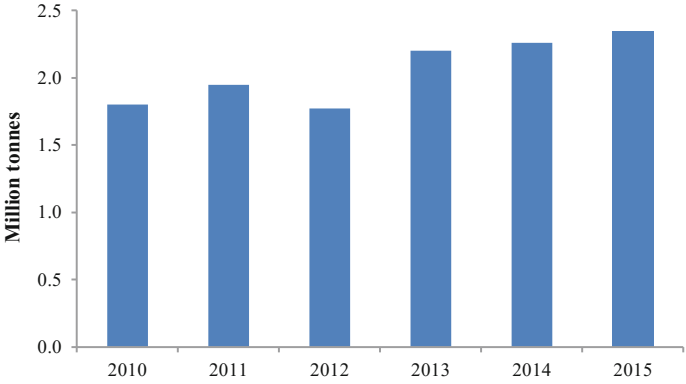

Fig. 10.16 Fuel ethanol production in China since 2010

Figure 10.16 shows the fuel ethanol production in China between 2010 and 2015. We can see that there is only small increase in the ethanol production. The government put forward that production capacity of grain plants ethanol will not increase in the future in *Medium and Long Term Renewable Energy Development Plan* published in 2007. So there is no increase in the production of 1G ethanol since then. The capacity just maintain at the original level. We assume that in the next 15 years the capacity of 1G ethanol will maintain at the present level which is 190 tonnes per year.

The 1.5G ethanol begins to demonstrate since 2007 and has a capacity of about 40 tonnes per year at present in China. The production of non-grain plants is the key factor that affects the production of the 1.5G ethanol. To preserve "food security", China has set a minimum area for the country's farmland—120 million hectares in *National long term plan for food security*, just below the present level. That means existing farmland cannot be used for non-grain plants. Only unused land will be considered in the production forecast.

The potential capacity of the 1.5G ethanol is small because of resource limits. Table 10.5 shows the potential production of non-grain plants. We can see that the suitable area in unused land is relatively small for cassava and sweet potato. There are more lands for sweet sorghum, but its mass yield of ethanol is small. Because of the distribution of suitable lands, it's difficult to utilize all the suitable areas in

Table 10.5 Potential production of non-grain plants

Plants	Plant area (10 thousand hectares)	Suitable area in unused land	Yield per hectare(tonnes)	Mass yield of ethanol (%)
Cassava	40	2	20.5	17
Sweet sorghum	Small	440	60	6.25
Sweet potato	350	30	21.5	12.5

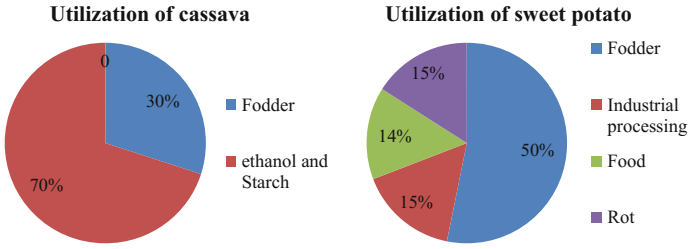

Fig. 10.17 Utilization of cassava and sweet potato in China in 2011

Table 10.6 Potential production of the 1.5G ethanol in China

Plants	Fuel ethanol potential production (10 thousand tonnes)	
Cassava	100	70 % for ethanol
Sweet sorghum	1650	100 % for ethanol
Sweet potato	150	15 % for ethanol

unused lands. Besides, most non-grain plants have other uses, not all of them can be used to produce fuel ethanol.

Figure 10.17 shows the utilization of cassava and sweet potato in China in 2011, and it didn't change much since then. The production of sweet sorghum is very small in China. Producing fuel ethanol is its only utilization (Xu et al. 2010; Zhang 2009; Zhang et al. 2010a; Zhang et al. 2011; Zhang et al. 2010b).

Based on the analysis above, we can calculate the potential production of the 1.5G ethanol in China which shows in Table 10.6. It can be seen that the potential production of sweet sorghum is much larger than cassava and sweet potato. As sorghum utilization is still on demonstration, financial support and encouraging policies are needed for its promotion.

The raw material for the 2G ethanol is cellulose which is mainly comes from agricultural and forestry residues. These residues are very useful in many aspects. As about 15 % of agricultural residues and 27 % of forestry residues can be utilized for energy use (Ouyang 2006). Only part of these energy-use residues are for liquid fuels. Forecasts of the residues for liquid fuels in China are different in many references. So we choose two groups of data to represent the low case forecast and the high case forecast which shows in Fig. 10.31. The mass yield of ethanol out of cellulose is assumed to be 23 %. So we can get the forecast of potential production for the 2G ethanol in China which shows in Fig. 10.18 (China Automotive Energy Outlook 2012).

It can be seen that the potential production of the 2G ethanol is about 20–40 million tonnes, a little larger than that of the 1.5G ethanol. Since the 2G ethanol production is still under research and it just start to demonstrate in China, the cost is

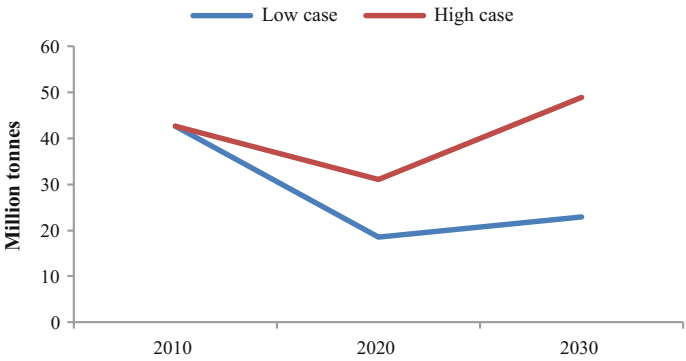

Fig. 10.18 Potential production for the 2G ethanol in China

very high. More encouraging policies and financial supports are needed for the development of the 2G ethanol production.

Generally speaking, from the point of view of policies in China, there were great policy supports between 2001 and 2007, when vehicle-use ethanol is mentioned in the 10th and 11th five-year plan. Between 2001 and 2005, *Vehicle use ethanol pilot program* and *detailed rules of vehicle use ethanol pilot work* were published to promote the application of fuel ethanol in China. In 2007, the government put forward some requirements to the development of fuel ethanol in *Medium and long term renewable energy development plan.* It said that the capacity of grain plants ethanol would not increase in the future, the 1.5G ethanol was the key project in the short term and the 2G ethanol was developed in the medium and long term. It also required that vehicle-use ethanol consumption would reach to 5 million tonnes in 2015 and 10 million tonnes in 2020. But actually the fuel ethanol production in 2015 was only about 2.3 million tonnes. Since 2008, the policy support for fuel ethanol has been smaller than before. In the 12^{th} five-year plan, vehicle-use ethanol development is totally not mentioned. This is also the same case in the 13^{th} five-year plan.

Besides, the financial aids for fuel ethanol production have been reducing since 2010. For the 1G ethanol the financial aids fall from 2000 yuan per tonne before 2010 to 500 yuan per tonne in 2012, and in 2016 there will not be financial aids anymore. For the 1.5G ethanol the financial aids are 750 yuan per tonne since 2012. For 2G ethanol there are no clear financial aids. So we assume that there will not be clear encouraging policies or large amount of financial aids for fuel ethanol in the foreseeable future.

Based on the analysis above, we assume that the 1G ethanol production capacity will maintains at the same level as in 2015, 1.9 million tonnes per year. For the 1.5G ethanol, we assume that cassava ethanol and sweet potato ethanol total production capacity will reach to 1 million tonnes per year in 2030 due to their small production potential in China. For sweet sorghum ethanol, it is assumed that there will be 50 thousand tonnes increase per year due to no clear policy support. For the

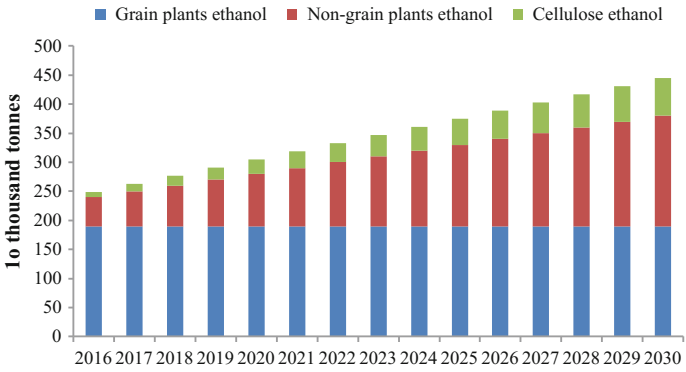

Fig. 10.19 Forecast of fuel ethanol production in China between 2016 and 2030

2G ethanol, the case is similar to that of sweet sorghum. Assuming that there is no clear policy support and no great technical progress in production, we just think that the production capacity of the 2G ethanol will be 40 thousand tonnes increase per year. Thus we can get the forecast of fuel ethanol production in China between 2016 and 2030, just as shown in Fig. 10.19.

10.3.3 Analysis of Natural Gas Vehicle

As the development of exploration techniques and theories, more and more natural gas resources in China are surveyed and mined in recent years. In 2013 the geological resources of natural gas in China has been surveyed to reach to 68,000 billion cubic meters, 40,000 billion cubic meters of them are recoverable and the proven reserves are about 12,000 billion cubic meters. National production and imports of natural gas increase every year to meet the growing demand in China. In recent years, natural gas is mainly utilized in the urban use, industry, chemical industry and power generation. The demand of natural gas in transportation grows fast and accounts for a certain proportion of the total. Based on existing researches, natural gas resource in China is huge and has large potential to be surveyed in the future, and with great amount of imports, the supply of natural gas can easily satisfy the demand (BP 2014) (Fig. 10.20).

As China is a big country, the distribution of natural gas is uneven in the whole area. But a great numbers of pipe lines have been built meet the natural gas transportation demands. Now the natural gas pipe line network has basically covered the whole area of the country, and More than 10 coastal LNG stations have been built or under construction to support the natural gas supply (www.cafv.com.cn www.chiagasmap.com) (Fig. 10.21).

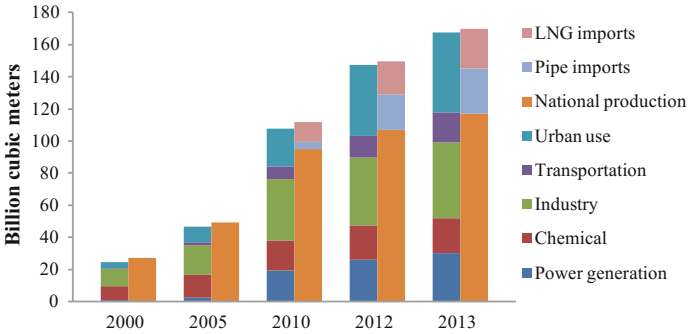

Fig. 10.20 Natural gas supply and demand in China between 2000 and 2013

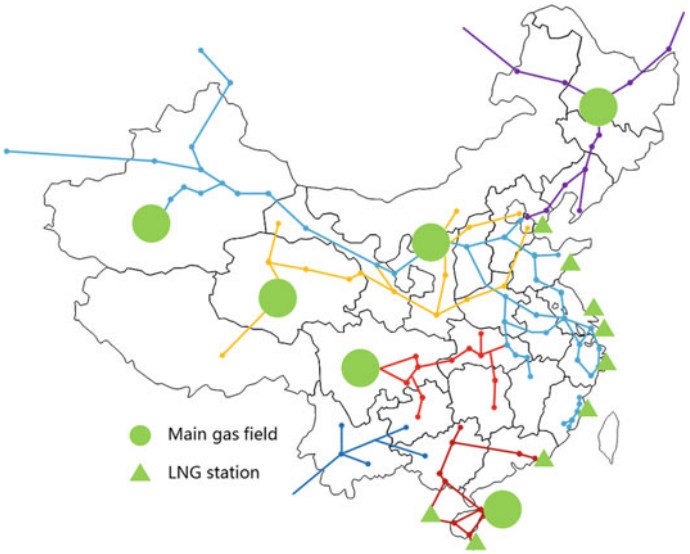

Fig. 10.21 Supply of natural gas in China

Vehicle-use natural gas utilization has two forms: compressed natural gas (CNG) vehicle and Liquefied natural gas (LNG) vehicle. Compared to LNG vehicles, CNG vehicles have lower pressure in the tank and the technology of which is more mature. So nearly all natural gas passenger cars are CNG ones in China. LNG vehicles are mainly trunks and coaches which are not considered in this article.

CNG taxies began to replace gasoline ones in 2008 and the ownership of CNG taxies increase fast since then. In 2014, about 53 % of taxies in China are CNG ones. In 12th five year plan for natural gas development, Vehicle use of natural gas is in the first priority category. So there is great nationwide gasification promotion

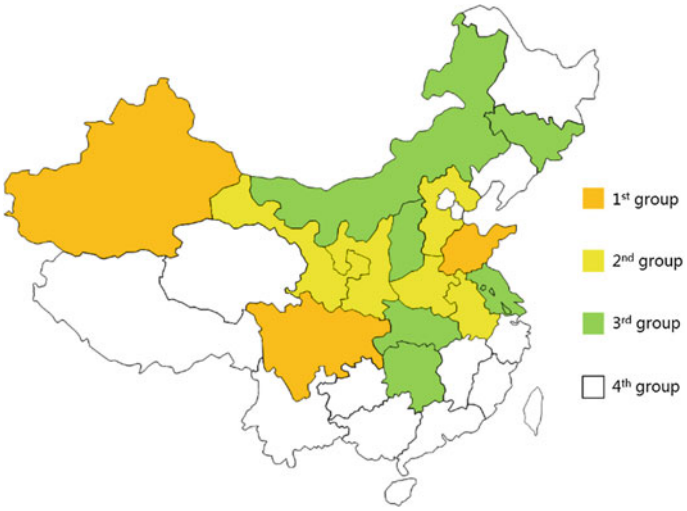

Fig. 10.22 Four groups in taxies gasification in China

in taxies and buses. Many provinces make their own targets of promoting CNG taxies. For example, Shandong put forward that the gasification of taxies should reach to 95 % in 2020 and Shijiazhuang says that the gasification of taxies will reach to 100 % in 2020.

According to the speed of gasification of taxies, we can divide the whole country to 4 groups which shows in Fig. 10.22. In the 1st group which near natural gas resources, the price of natural gas is low. There are more mature technology and production chain in the 1st group because of encouraging policies. The average rate of taxies gasification is 90 % and higher. For other groups, the rates are relatively lower: average 70 % in the 2nd group and 50 % in the 3rd group. In the 4th group, the gasification of taxies is still in promotion. But in Super Cities in China, such as Beijing, Shanghai, Guangzhou and Shenzhen, the gasification of taxies is extremely low: 3 % in Beijing, 6 % in Shanghai, 3 % in Guangzhou and no CNG taxi in Shenzhen. The taxi ownership in the 4 Super Cities accounts for about 10 % of total numbers of taxies (NBSC).

Based on the analysis above, we assume that the share of CNG taxies will reach to 90 % in 2020 and maintains until gasoline taxi reach to 0. According to policies of EV and PHEV, we consider that EV taxies will emerge in 2016 and will increase by 1 % per year. Thus we get the shares of taxies ownership in China between 2015 and 2030 which shows in Fig. 10.23. Combined with total numbers of taxies, we can forecast sales and ownerships of each type of taxi in the future.

CNG private and business cars are mainly promoted in the 1st group since 2010 due to mature technology and production chain, low price of natural gas and local policy supporting. As there is no clear policy for CNG private and business cars, its promotion will be affected by local gasoline and natural gas price. We assume that

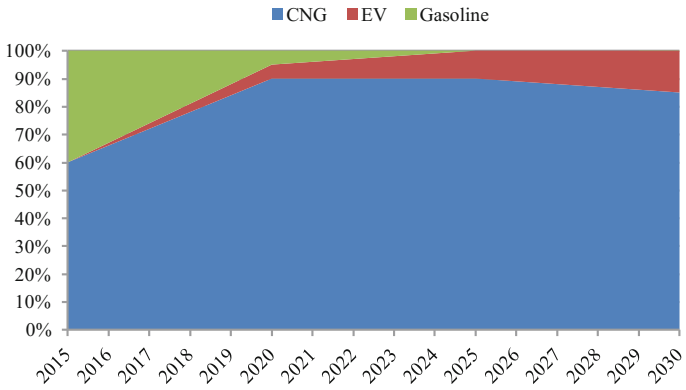

Fig. 10.23 Shares of taxies ownership in China between 2015 and 2030

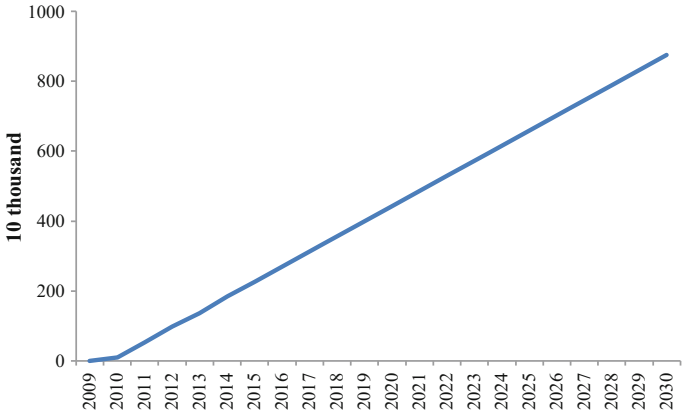

Fig. 10.24 CNG private and business cars ownership in China

CNG private and business cars will still be promoted in the 1st group, and the growth of ownership will maintain at the present speed. Thus we get the forecast of CNG private and business cars ownership in China which shows in Fig. 10.24.

10.3.4 Analysis of EV and PHEV

There are many encouraging policies and financial supports for EV and PHEV in China since 2010. The target of the promotion is very clear. The government put forward that the ownership of EV and PHEV should reach to 5 million by 2020 *in Energy saving and new energy vehicles development plan*. It is also required that charging ports should reach to 4.5 million by 2020 to support new energy vehicles

Fig. 10.25 Three groups in EV and PHEV promotion in China

in *Electric vehicle charging infrastructure construction program*. Due to *13th five year plan on new energy automobile charging infrastructure reward policy and strengthening the popularization and application of the new energy vehicles (exposure draft)* the whole country has been divided into 3 groups in EV and PHEV promotion based on the economy and the environment in different regions.

As shown in Fig. 10.25, the 1st group covers atmospheric pollution control key areas including Beijing, Shanghai, Tianjin, Hebei, Shanxi, Jiangsu, Zhejiang, Shandong, Guangdong and Hainan. The 2nd group covers Midland provinces including Anhui, Jiangxi, Henan, Hubei, Hunan and Fujian. Other provinces are in the 3rd group.

There are different requirements on new passenger car sales in different groups. Table 10.7 shows the percentage of new passenger car sales in each year that EV and PHEV must account for between 2016 and 2020 in the 3 groups in China. The data in Table 10.7 all comes from the *exposure draft.* We assume that the actual sales of new passenger car between 2016 and 2020 will match the requirements in Table 10.7. Data between 2020 and 2030 will be obtained by extrapolation. The forecast of EV and PHEV shares of new passenger car sales in the 3 groups shows in Fig. 10.26.

Table 10.7 Requirements on new passenger car sales in China between 2016 and 2020

Echelon	10 thousand	2016	2017	2018	2019	2020
1st	Sales	3	3.5	4.3	5.5	7
	Percentage	3 %	4 %	5 %	8 %	10 %
2nd	Sales	1.8	2.2	2.8	3.8	5
	Percentage	2 %	3 %	4 %	5 %	6 %
3rd	Sales	1	1.2	1.5	2	3
	Percentage	0.50 %	1 %	1.50 %	2 %	3 %

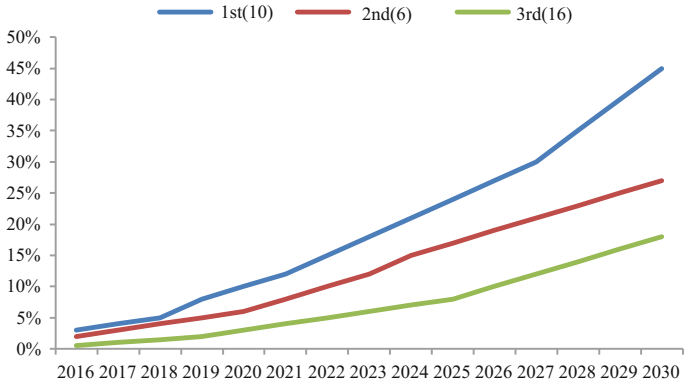

Fig. 10.26 EV and PHEV Shares of new passenger car sales in the 3 groups

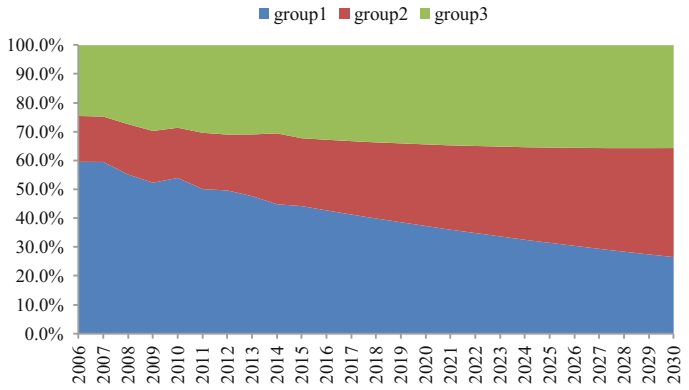

Fig. 10.27 Shares of private and business car sales in the 3 groups

As Fig. 10.27 shows, based on historical data, we can get shares of private and business car sales in the 3 groups between 2006 and 2015. The forecast data between 2016 and 2030 can be obtained by extrapolation. We can see that shares of private and business car sales decrease in the 1st group but increase in the 2nd and 3rd groups. That is because regions in the 1st group are more developed than those in the 2nd and 3rd groups. In the 1st group, private and business passenger car ownership is higher and closer to the saturation level, so the rate of growth decreases every year. But in the 2nd and 3rd group, the situation is quite the opposite. So the 2nd and 3rd groups will contribute more in the growth of passenger and business car ownership in China in future.

Based on the data above and the model in Sect. 3, we can calculate EV and PHEV sales between 2016 and 2030. According to the EV and PHEV sales in recent 5 years, EV accounts for more percentage than PHEV. In Beijing, only EV can get the local financial aid. *National financial aid for EV and PHEV in 2015* put

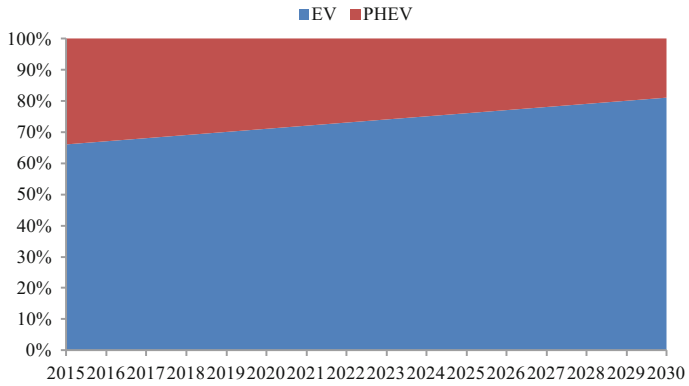

Fig. 10.28 Shares of EV and PHEV sales in China between 2015 and 2030

forward that EV (80–150 km) can obtain 31.5 thousand yuan per car, EV (150–250 km) can obtain 45 thousand yuan per car, EV (>250 km) can obtain 54 thousand yuan per car, while PHEV can only obtain 31.5 thousand yuan per car. So it can been that EV is more recommended by the government. PHEV is more acceptable at present for the absence of charging pipes at present.

We think that shares of EV will increase in the future with more supports of the government and more charging pipes built. So we just assume that there will be 1 % increase of EV shares per year. Figure 10.28 shows the forecast for shares of EV and PHEV sales between 2015 and 2030. Thus EV and PHEV sales in the next 15 years can be calculated separately.

10.3.5 Forecast for Passenger Car Sales and Ownership in Different Fuels

In this part, we combine the alternative fuel analysis with the model in Sect. 2 and get the result of passenger car sales and ownership in different fuels. According to the forecast in Sect. 2, we know that annul mileage and the vehicle age of business car is similar to that of private car, share of business car is much smaller than that of private car. So for ease of calculation, we merge the two types of passenger car into one category. This is helpful for the combination between the analysis in Sects. 2 and 3, for alternative fuel private passenger car and business one are difficult to forecast separately. It is more sensible to put them together. So the results in this part are divided into two categories, one is private and business sector and the other is taxi. Figures 10.29 and 10.30 show the sales cases and Figs. 10.31 and 10.32 show the ownership cases.

We can see that in private and business car sector gasoline car still accounts for most percentage of the total in sales and ownership. In taxi sector gasoline car

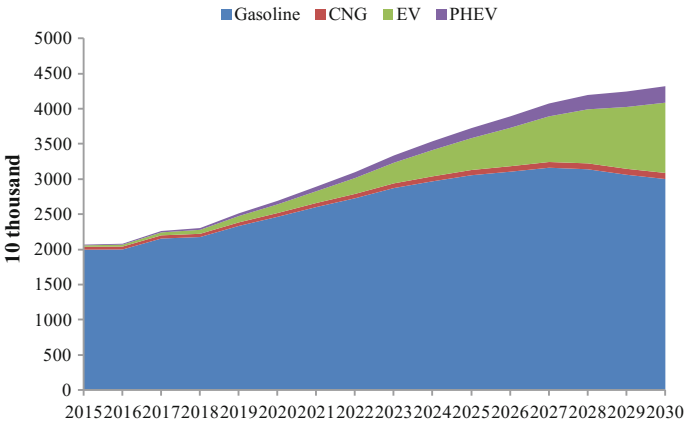

Fig. 10.29 Private and business passenger car sales in China based on fuels

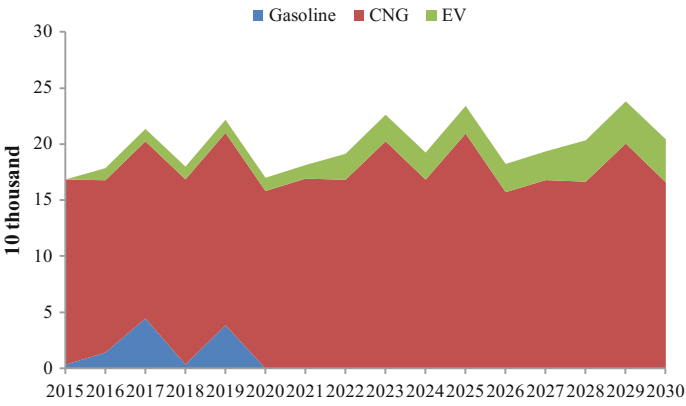

Fig. 10.30 Taxi sales in China based on fuels

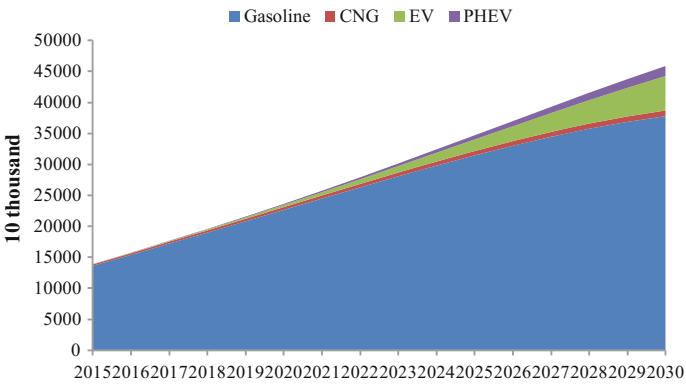

Fig. 10.31 Private and business car ownership in China based on fuels

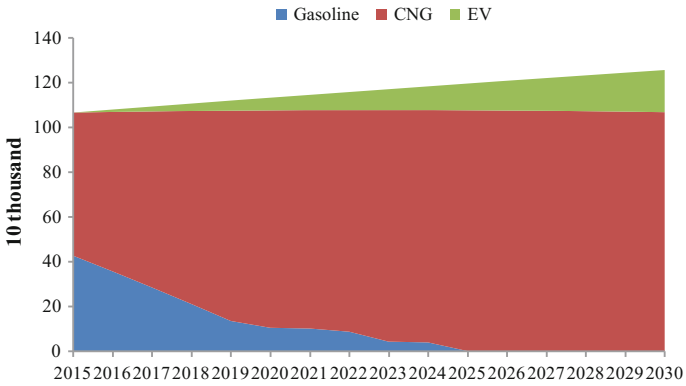

Fig. 10.32 Taxi ownership in China based on fuels

decrease by year and will disappear in 2025. CNG car will be in dominant position in sales and ownership of taxi, EV accounts for the rest percentage of the total in the future. Since fuel ethanol is mixed in gasoline in China, it doesn't affect the type of passenger car, it will be considered when calculating fuel consumption of passenger car. The results in this part are the basis for further calculation in fuel consumption.

10.4 Analysis of Gasoline Consumption and Improving Target of Gasoline Car

10.4.1 Actual Gasoline Consumption Rate of Passenger Car

For passenger car in China, there is a value of nominal gasoline consumption per hundred kilometers. Figure 10.33 shows the average nominal gasoline consumption per hundred kilometers in different years. Since 2003, under the influence of national policy, the average nominal gasoline consumption per hundred kilometers of passenger car decreases gradually, from the 9.11 L/100 km before 2003 down to 7.31 L/100 km in 2013. The data after 2013 refers to the policy in *Energy saves and new energy vehicles industry development planning (2012–2020)* that in 2020, the average nominal gasoline consumption per hundred kilometers should fall down to 5.0 L/100 km [33]. Data after 2020 can be forecasted based on different scenario settings. As a reference scenario, we just assume that the data maintains at 5.0 L/100 km from 2020 to 2030.

The actual gasoline consumption is generally higher than nominal gasoline consumption. The actual gasoline consumption depends on driving habits of drivers and traffic conditions. Since driving habits of drivers have a strong randomness, we only consider the effect of traffic conditions in this report. Based on existing researches (Beijing traffic development report 2013), actual to nominal ratio of

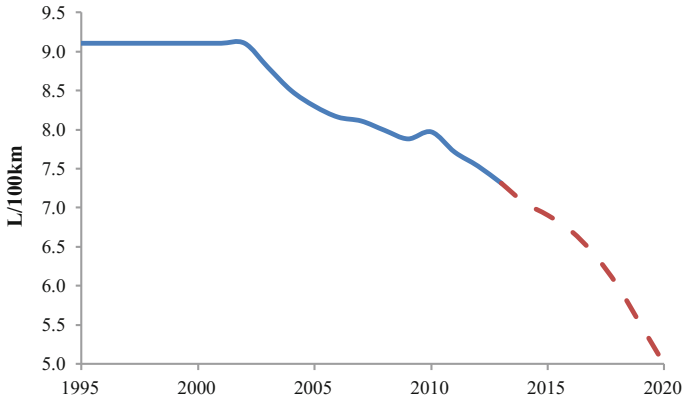

Fig. 10.33 Average nominal gasoline consumption rate of new produced passenger car in China

gasoline consumption increases by year in China. That is because with passenger car ownership in China growing, traffic congestion is more and more serious. But considering cases in developed countries and based on common sense, the actual gasoline consumption will not increase infinitely, but turn to a relatively stable value.

In this section, we assume the nominal gasoline consumption per hundred kilometers is the lower limit of the actual consumption. Based on cases that gasoline consumption in urban conditions is 30 % higher than comprehensive conditions, assuming there is 10 % increment for the bad driving habits. We assume that the maximum of gasoline consumption ratio of actual to nominal is 1.4.

Since traffic condition is related to ownership of passenger car. We establish the relationship between fuel consumption ratio of actual to nominal and ownership of passenger car per thousand people as shown in Fig. 10.34. According to the changing law, Gompertz is used to fitting the data. It indicates that gasoline

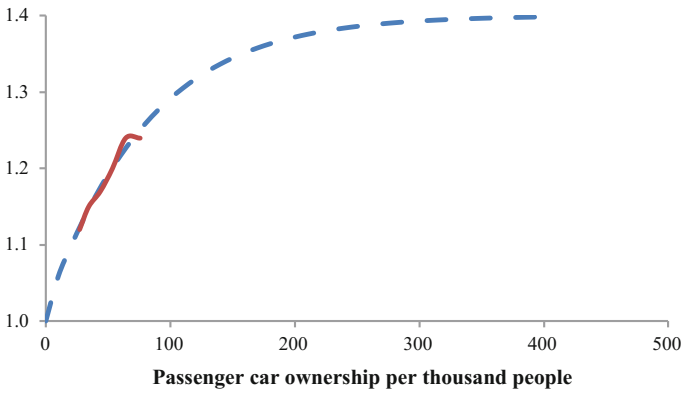

Fig. 10.34 Relationship between average actual to nominal ratio of passenger car gasoline consumption and passenger car ownership per thousand people

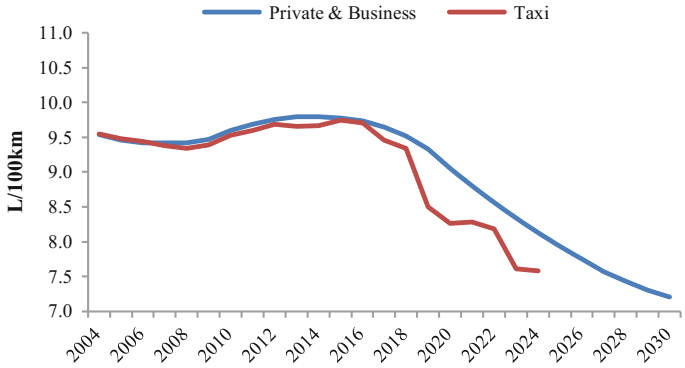

Fig. 10.35 Average actual gasoline consumption rate of passenger car, between 2004 and 2030

consumption ratio of actual to nominal increases with ownership of passenger car per thousand people growing.

Based on the nominal data, actual to nominal ratio and vehicle age distribution, we can get the actual gasoline consumption rate of passenger car in the next 15 years. Figure 10.35 shows the average actual gasoline consumption per hundred kilometers of the two types of passenger car from 2004 to 2013. It can be seen that the nominal gasoline consumption per hundred kilometers decreases by year since 2003, but the ratio of actual to nominal increase sharply. So the actual data didn't change much between 2004 and 2015. Since 2016 the ratio of actual to nominal tends to a fixed level, the actual data has really decreased by year. The decrease of taxi is faster than that of private and business passenger car because the vehicle age of taxi is much shorter.

10.4.2 Analysis of Average Annual Mileage of Passenger Car

The average mileage per year is absence of nationwide statistics. The taxi average annual mileage data from 2002 to 2009 in other researches (Huo 2012) is selected in our model. We assume the average annual mileage of taxi maintains at 100 thousand kilometers since 2010. For private and business passenger car, we can get the results by calculation based on the data we get in previous chapters. The data of passenger car gasoline consumption between 2004 and 2015 can be obtained in other researches, combined with the data of gasoline passenger car ownership and passenger car actual fuel consumption per hundred kilometers, we can calculate the average annual mileage of private and business passenger car between 2004 and 2015, as shown in Fig. 10.36.

It indicates that the average annual mileage of private and business passenger car declines by year. The average annual mileage is 29 thousand kilometers in 2004 and

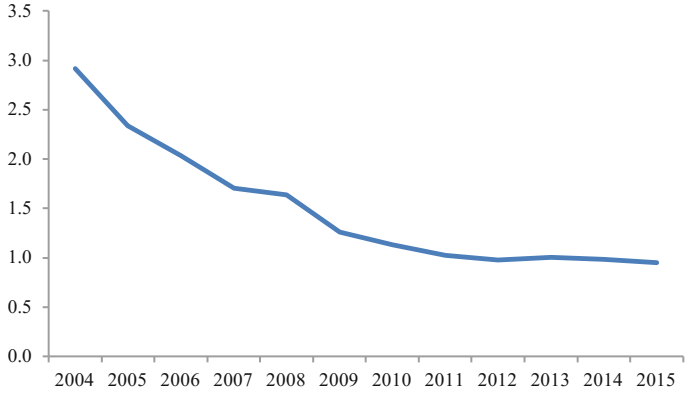

Fig. 10.36 Average annual mileage of private and business passenger car, between 2004 and 2015

drops to about 10 thousand kilometers in 2011 and maintains at the level since then. We assume that the average annual mileage of private and business passenger car will maintain at 10 thousand kilometers in the next 15 years.

10.4.3 Comparison of Gasoline Supply and Consumption in China

Based on the assumption in previous parts, we can calculate the gasoline consumption of passenger car sector. It can be seen that though the speed of growth slows down by year, the amount of increase is still very high. Gasoline is mainly consumed in passenger car sector in China, while other sectors also accounts for small percent of the total, which cannot be neglected. According to existing researches (Yang 2013), we get the data of gasoline consumption in other sectors. Combined with gasoline consumption in passenger car sector we calculate above, we can get the total consumption of gasoline in China in the future, which shows in Fig. 10.37.

Based on the analysis in part 1, if refining sector doesn't change its structure of production, the flexibility of productivity between gasoline and diesel is relatively small. The maximal output rate of gasoline from crude oil in China can be assumed as 24 %. The maximal oil dependency ratio in China is set to 70 %. The maximal supply of gasoline from refining sector in China is 168 million tonnes, which shows as a red line in Fig. 10.37. If we compare the gasoline consumption in this chapter with the maximal gasoline supply we get in part 1, we can conclude that the gasoline supply from refining sector cannot meet the gasoline consumption since 2022. In China gasoline consumption mainly comes from passenger car sector. So some effective measures should be taken to reduce gasoline consumption in passenger car sector.

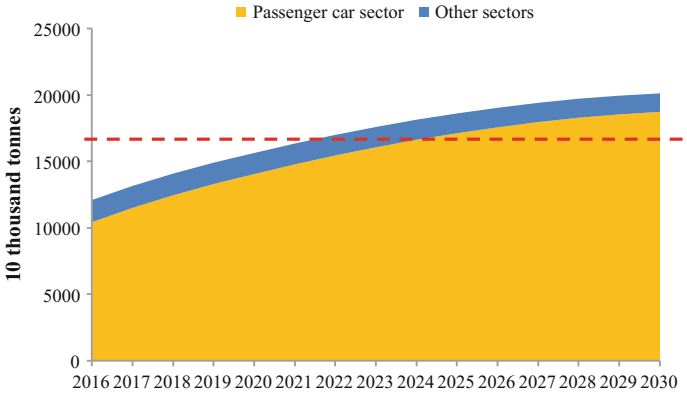

Fig. 10.37 Gasoline consumption in China between 2016 and 2030

10.4.4 Improving Targets of Reducing Gasoline Consumption

Ownership of gasoline passenger car, average annual mileage and actual gasoline consumption rate are the three key factors to affect gasoline consumption in passenger car sector. Since the ownership of gasoline passenger car has been analyzed in previous parts, we do not take measures to affect it here. Only average annual mileage and actual gasoline consumption rate are considered in this section.

Considering technology improvement in passenger car, reducing nominal gasoline consumption rate is a feasible way to improve fuel economy. In the previous analysis, the nominal gasoline consumption rate of new cars between 2016 and 2020 refers to *Energy saves and new energy vehicles industry development planning (2012–2020)*, from 6.7L/100 km to 5.0L/100 km. We assume that the data maintains at 5.0L/100 km between 2020 and 2030. In this section, for private and business passenger car, we consider 3 types of fuel economy cases. One is the reference case that nominal gasoline consumption rate of new cars maintains at 5.0L/100 km since 2020, the second case is that the fuel economy data falls to 4.0L/100 km in 2025 and maintains at the level since then, the third case is the fuel economy improvement case that the data falls to 3.0L/100 km in 2030. The average actual fuel economy of private and business passenger car shows in Fig. 10.38.

We can see that the actual data doesn't decline too much as the nominal data of new cars. That is because ownership of "old cars" still accounts for certain percentage of the total. The effect of fuel economy improvement is delayed. As high ratio between actual and nominal (A/N ratio) is mainly caused by traffic congestion. If we focus on developing specific techniques to reduce gasoline consumption rate in traffic jam, the A/N ratio may decline. Taking this into consideration, we assume that new cars of A/N ratio in 2020 are only 90 % of the ratio in the same year, and

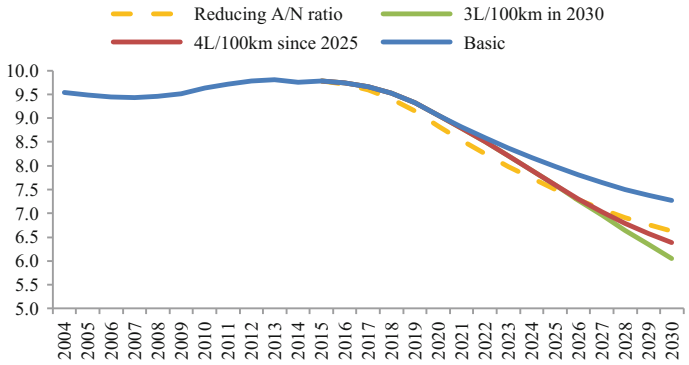

Fig. 10.38 Average actual fuel economy of private and business passenger car

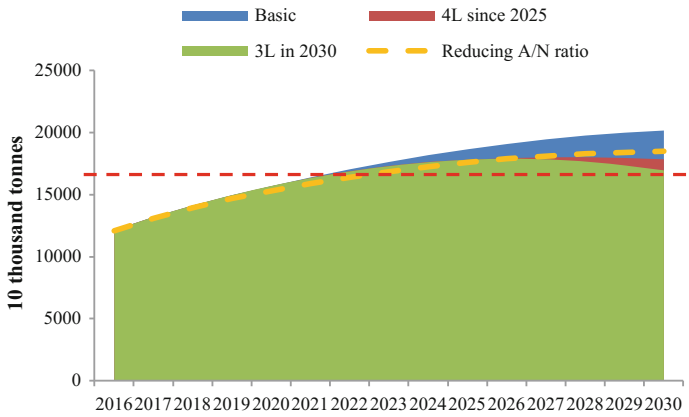

Fig. 10.39 Gasoline consumption in China in technology improvement case

the data maintains to 2030. In this case, as the yellow line shows in Fig. 10.38, actual consumption rate will have a great decline in the future.

Based on these four cases, we get the result of gasoline consumption which shows in Fig. 10.39. We can see that in each case there is a great decline of gasoline consumption between 2025 and 2030. Total gasoline consumption begins to decrease since 2025 in the fuel economy scenario. But due to the delayed effect, the gasoline consumption between 2020 and 2025 doesn't change much compared to the reference scenario. The supply still can't meet the consumption since 2022. So we can't get the supply and consumption of gasoline in balance only by improving fuel economy of new cars or by reducing A/N ratio. Besides gasoline consumption in 3L case appear to be little different from that in 4L case. So there is no necessity to develop much lower nominal rate in the near future.

If we combine 4L case and reducing A/N ratio case which shows in Fig. 10.40, we can find that the decline in gasoline consumption is considerable, especially

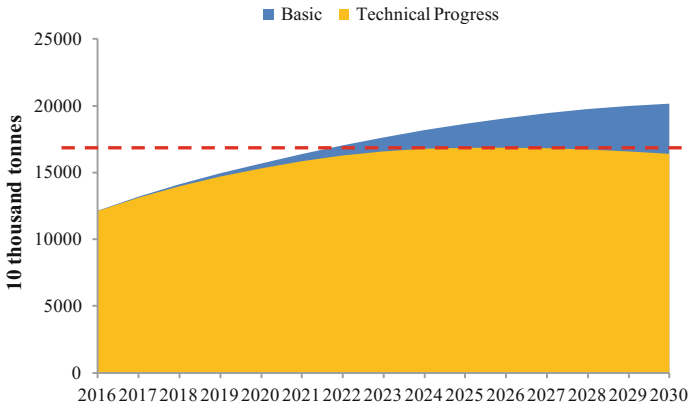

Fig. 10.40 Gasoline consumption in combined case

between 2025 and 2030. The gasoline consumption can just reach to the highest line of supply. So technology improvement in reducing actual rate is an effective way to reduce gasoline consumption in the long run. The results indicate that gasoline consumption between 2016 and 2025 changes little due to the delayed effect of the technology improvement case. Some other measure should be taken to complement this.

The average annual mileage in previous analysis is assumed to maintain at 1.0 km since 2016. That is because the data changes little between 2011 and 2015 duo to our results of calculation. Here we assume that the average annual mileage would fall from 1.0 km to 0.9 km between 2016 and 2020, and maintain at 0.9 km since then. This is the case of reducing mileage scenario. The results show in Fig. 10.41.

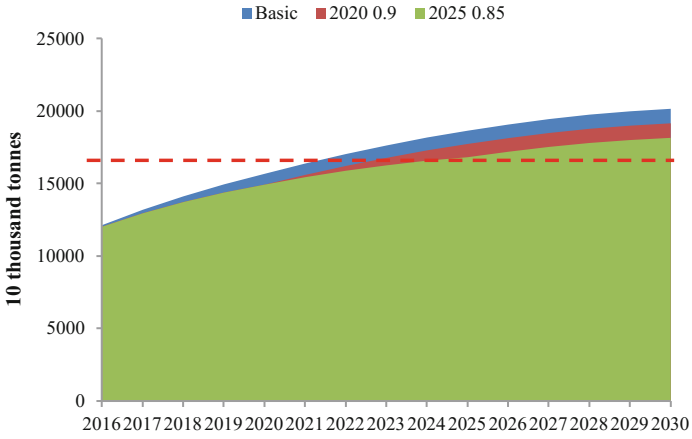

Fig. 10.41 Gasoline consumption in China in reducing mileage scenario

It can be seen that reducing average annual mileage is an effective way to reducing gasoline consumption, especially between 2016 and 2025. In the reducing mileage scenario, though there is no improvement in fuel economy, the gasoline supply can meet the consumption until 2025. If we just improve the fuel economy of new cars to 4.0L/100 km in 2025 and maintain at the level since then, the supply and consumption of gasoline can be in balance. This just compensates for the lack of technology improvement case.

So we can conclude that reducing average annual mileage of private and business passenger car is an essential way to reduce gasoline consumption between 2016 and 2025. Fuel economy improvement of new passenger cars is also necessary, but the target is not so high. The basis requirement is that the nominal gasoline consumption rate of new cars reaches to 4.0L/100 km in 2025 and maintains since then. Measures to reduce A/N ratio is also important. New sold car in 2020 should have at least 10 % lower in the A/N ratio. In this way, the gasoline supply from the refining sector in China can meet the gasoline consumption in the future.

10.5 Conclusions

In this chapter, we firstly present calculation of the gasoline and diesel supply from the refining sector in China. A virtual refinery model is established to analyze the productivity of gasoline and diesel. Based on the calculation of the model, we get minimum and maximum production of gasoline and diesel and the flexibility between them. We conclude that when refining structure is stable and the production of other products is fixed, the flexibility of the production range is rather small only by change operational modes. The capacity structure of refining sector in China in the future tends to produce more gasoline. Considering the flexibility, the maximal output rate of gasoline in refining sector in China is about 24 %.

Future estimation of ownership and sales of passenger car in China is provided afterwards. A model for different types of passenger cars is established and many factors which affect the ownership are analyzed separately. Vehicle age distribution in passenger car is also taken into consideration. Sales in each year are developed based on the ownership and vehicle age analysis. The forecast of sales and ownership of passenger car between 2016 and 2030 is obtained as the basis of further study.

In the third part, alternative fuels and technologies are studied. Fuel ethanol, natural gas, EV and PHEV are analyzed in detail. Combined with the results in Sect. 2, we get the forecast for sales and ownership of passenger car in different fuels and find that in private and business car sector gasoline car still accounts for most percentage of the total in sales and ownership. In taxi sector gasoline car decrease by year and will disappear in 2025. CNG car will be in dominant position in sales and ownership of taxi, EV accounts for the rest percentage of the total in the future.

In the last part, we first analyze the actual gasoline consumption rate and the average annual mileage of passenger car. Then we make an assumption as a reference scenario. The gasoline consumption in passenger car sector and other sectors in the future are calculated based on the assumption above. Result indicate that the gasoline supply in the refining sector cannot meet the gasoline consumption in 2022 and afterwards. Effective measures should be taken to reduce gasoline consumption in passenger car sector. Key measures include: (a) the average annual mileage falls to 0.9 km in 2020 and maintains in the next few years, (b) the nominal gasoline consumption rate of new cars drops to 4.0L/100 km in 2025 and maintains afterwards, and (c) newly sold cars in 2020 should have at least 10 % lower in the actual-to-nominal fuel consumption ratio. In this way, the gasoline supply from the refining sector in China can meet the gasoline consumption in the future.

References

An, J., et al. (2015). Supply and demand trend of lube base oil at home and abroad. *Petroleum Economics, 1*, 11–19.

Beijing traffic development report (2013).

BP Statistical Review of World Energy (June 2014). www.chinagasmap.com.

Cai, Q., & Liu, J. (2011). Supply and demand trend of liquefied petroleum gas in China [J]. *Petroleum Planning & Engineering, 22*(3), 6–9.

Chen, S. (2007). *Comparison of high sulfur crude oil processing route*, Research Institute of Petroleum Processing.

Chi, H. (2014). Petroleum coke market supply and demand pattern analysis in China, *China petroleum and chemical industry of economic analysis*, (7).

China automotive energy outlook 2012. (2012). Science Press. www.cafv.com.cn.

Fan, J. (2012). Solvent oil market review and outlook. *Liaoning Chemical Industry*, *41*(8).

Financial support policy for new energy vehicles between 2016 and 2020.

Han, H., Hewu, W., & Minggao, O. (2011). Fuel conservation and GHG (Greenhouse gas) emissions mitigation scenarios for China's passenger vehicle fleet. *Energy, 36*(11), 6520–6528.

Hou, F. (2005). *The development of China's oil refining industry*, China Petrochemical Corporation.

Hou, F. (2011). China's oil refining technology, China Petrochemical Press.

Huang, J. (2001) *Imported crude oil evaluation data set*, China Petrochemical Press.

Huang, J. (2013). China's fuel oil market present situation and prospect, *Sino-Global Energy*, *18* (9).

Huo, H., & Wang, M. (2012). Modeling future vehicle sales and stock in China. *Energy Policy, 43* (3), 17–29.

Huo, H., Zhang, Q., He, K., et al. (2012). Vehicle-use intensity in China: Current status and future trend[J]. *Energy Policy, 43*(3), 6–16.

Kobos, P. H., Erickson, J. D., & Drennen, T. E. (2003). Scenario analysis of Chinese passenger vehicle growth. *Contemporary Economic Policy, 21*(2), 200–217.

Kong, J., et al. (2014). China's asphalt market for medium and long-term prospects. *International Petroleum Economy, 22*(7), 93–97.

Lei, Q. (2015). Fuel ethanol technology research status and development trend analysis. *Guangzhou Chemical Industry, 5*, 42–43.

Li D. (2009). *High sulphur crude oil processing process*. China Petrochemical Corporation.

Li, J. (2012). *Evaluation and feasibility of four kinds crude oil processing scheme*, Petro china dalian petrochemical company.
Li, W., Fu, F., Ma, L., et al. (2013). A process-based model for estimating the well-to-tank cost of gasoline and diesel in China [J]. *Applied Energy, 102*(2), 718–725.
Lu, W. (2009). The characteristics of crude oil in China.
Ma, L., Fu, F., & Li, Z., et al. (2012).Oil development in China: Current status and future trends *Energy Policy*, *45*(2), 43–53.
National Bureau of Statistics of China (2014).
Ou, X. M., Zhang, X. L., & Chang, S. Y. (2010). Scenario analysis on alternative fuel/vehicle for China's future road transport: Life-cycle energy demand and GHG emissions. *Energy Policy, 38*(8), 3943–3956.
Ouyang, M. (2006). China's energy saving and new energy vehicle development strategy. *Automotive Engineering, 04*, 317–321.
Qian, X., & Jiang, X. (2014). Domestic oil and gas industry development report.
Shen, Z. (2006). Forecast of car ownership based on the income distribution curve in China. *China Energy, 08*, 11–15. doi:10.3969/j.issn.1003-2355.2006.08.003.
Shen, B. (2009). Petroleum refining technology, China Petrochemical Press.
Shi, B., & Bai, X. (2014). China's refined oil market analysis and prospect in 2014, *Chemical Industry*, *32*(9).
Shi, B., et al. (2012). Global naphtha market supply and demand, trade flow analysis and Suggestions of China's purchasing. *International Petroleum Economy, 20*(10), 78–83.
The 12th five year plan on natural gas development.
The13th five year plan on new energy automobile charging infrastructure reward policy and strengthening the popularization and application of the new energy vehicles (exposure draft).
Wang, X. (2013). Surveying the development of China's oil reserves and the countermeasure analysis. *Technology, Economics & Management Research, 02*, 102–110.
Wang, C., & Zhu, Y. (2013). Forecast of kerosene in China. *China Petrochemical Industry, 10*, 27–28.
Xu, Z., et al. (2010). Sweet sorghum potential research. *Renewable Energy Sources, 28*(4), 118–122.
Yang, X. (2013). China's refined oil market demand forecasting research, China Petrochemical Press.
Zhang, L. (2009). Maize development and the byproducts usage. *Food Engineering, 4*, 14–15.
Zhang, C., et al. (2010a). Potential and spatial distribution of the sweet potato ethanol in China. *Resource Science, 32*(3), 505–511.
Zhang, C., et al. (2010b). Potential and spatial distribution of the cassava ethanol in China. *Resource Science, 30*(17), 4765–4770.
Zhang, C., et al. (2011). Potential and spatial distribution of the cassava ethanol in China. *Resource Science, 30*(8), 1726–1731.

Chapter 11
Comparison of Decomposed Regional Energy Consumption in China

Lingying Pan

Abstract Currently, most of studies on China's energy demand and consumption took the huge country as a whole. As China is a huge country composed of more than 30 administrative regions with different scales, structures and intensities of energy consumption, this manuscript revealed the regional disparities in energy consumption of China's 30 provinces. Based on a hybrid energy input-output model, the total energy consumption of different regions was decomposed and compared using three measurements of embodied energy in inter-regional trade: (1) only inter-regional energy trade was considered; (2) embodied energy in flow-out of final goods and services was considered; (3) embodied energy in flow-in of final goods and services was considered. According to the results of the second and third measurements, the 30 regions were categorized into four groups by their energy intensity and per capita GDP. This manuscript discussed the common characteristics of decomposed regional energy intensity, and provided policy implication for regional energy conservation. The results implicated that for developed regions with low energy intensities, such as Shanghai, energy conservation should focus on promoting low energy-consuming life style. For under-developed regions with low energy intensities, such as Guangxi, economic development is more urgent than energy conservation. For developing and energy absorbing regions, improving energy efficiency in industries is significant. For developing and energy exporting regions, transforming primary energy into high value-added products would be beneficial for economic development and energy conservation.

Keywords Energy consumption · Input-output model · Regional disparity · Energy intensity

L. Pan (✉)
Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
e-mail: panly05@usst.edu.cn

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_11

Abbreviation:

GDP: gross domestic production
IO: input-output
CO2: carbon dioxide
RMB: currency measurement unit of China
PPP: purchasing power parity

11.1 Introduction

In current days, energy conservation is a fundamental and prioritized national energy policy of China to ensure its sustainable development. China is promoting energy conservation by aiming at reducing energy intensity and controlling total energy consumption (Liao et al. 2013; SCIO 2012; Ma et al. 2011). In August 2012, the State Council of China enacted the 12th Five Year Plan of Energy Conservation & Emission Mitigation (SCC 2012). As part of this plan, China determined that by 2015, the total energy consumption would be under 4 billion tce (tonne coal equivalent) and the energy consumption per 10000 RMB GDP (gross domestic production) would be reduced by 16 % compared to that of 2010. Moreover, in November 2012, the 18th National Congress Report of Chinese Communist Party declared to promote a revolution of energy production and consumption to control the total energy consumption and enhance energy conservation for national energy security (XNA 2012).

For the implementation of these energy conservation policies, considering that China is a huge country composed of more than 30 administrative regions with different scales, structures and intensities of energy consumption, the corresponding regional targets and measures should be developed according to the specific characteristics of each region (SCC 2012). Therefore, it is important to study China's regional disparities with respect to energy consumption.

Regional disparity in energy consumption has become a popular and important research topic in recent years. Accordingly, while it has already been noted that regional disparities exist in some areas of China, it has further been found that energy consumption has a strong relationship with economic growth (Herrerias et al. 2013) and with the development of secondary industry (Zhang and Xu 2012a) and trade (Liu et al. 2010) at the country level (Herrerias et al. 2013; Meng et al. 2011; Liang et al. 2007; Liang et al. 2010). For an integrated analysis of regional disparities in energy consumption that considers a number of interrelated regions, the input-output (IO) method is widely applied because it deals well with the problems of embodied energy analysis (Duchin 1992; Miller and Blair 1985; Costanza 1980; Treloar 1997), which observes the sum of all the energy required to produce any goods or services used for regional final consumption expenditures, gross capital formation, and inter-regional trades. For example, the IO method analyses embodied energy and CO_2

in international or inter-regional trade (Su 2014, 2010, 2011; Liu et al. 2010; Guo et al. 2012; Zhang et al. 2013) and direct and indirect energy consumption in typical sectors and industries (Liu et al. 2009; Park and Heo 2007; Cellura et al. 2011; Zhu et al. 2012), and it decomposes national energy consumption as embodied energy in goods and services (Fu et al. 2013). Embodied energy is defined as the total (direct and indirect) energy required for the production of economic or environmental goods and services (Costanza 1980) or the energy consumed in all activities necessary to support a process, including upstream processes (Treloar 1997). Input-output analysis is well suited to calculate indirect effects in a systematic and all-inclusive accounting framework (Costanza 1980). Though some previous studies have discussed China's regional disparities in energy consumption and the causality between energy consumption and economic growth (Wang et al. 2012; Ali Akkemik et al. 2012; Wang et al. 2013; Zhang and Lin 2012b; Herrerias et al. 2013; Zhang and Xu 2012a; Zhang et al. 2013; Li et al. 2013; Yu 2012), this problem deserves further discussion because regional disparities in energy consumption can only be comprehensively understood by a more integrated analysis of energy embodied in regional final consumption expenditures, gross capital formation and inter-regional trades, together with an analysis of regional economic levels and industrial structures.

This manuscript provides a comprehensive understanding of China's regional disparities in energy consumption from the view of embodied energy. By establishing a hybrid energy input-output model, regional energy consumption is decomposed into embodied energy in goods and services for final consumption expenditures, gross capital formation, and inter-regional trades. The decomposed embodied energy is compared with statistical data of energy consumption, in order to comprehensively present and explain the deviations in accounted regional energy consumption by different measurements. Referring to the similarities of some regions with respect to decomposed energy intensity of final consumption expenditures, gross capital formation, flow-in and flow-out and regional per capita GDP, 30 regions of China are classified into different groups as a reference for creating differentiated regional policies of energy conservation for the regions which have similar characteristics of energy consumption and at similar level economic development in China.

In Sect. 11.2, the methodology and data input for establishing the regional hybrid energy input-output model and the principal of regional energy consumption decomposition are introduced. In Sect. 11.3, based on the results of decomposition, an overview of energy consumption (decomposed energy amount and intensity) of 30 regions (Tibet, Hong Kong, Macao and Taiwan are not included due to lack of data) in China is provided, and energy consumption based on the measurements in the national statistical yearbook and the input-output analysis is compared. According to decomposed regional energy consumption intensities and regional economic levels, the 30 regions are categorized into 4 groups and the common energy consumption characteristics of the regions in the same group are discussed. In Sect. 11.4, the manuscript provides conclusions concerning the effects of different measurements of energy consumption on the understanding of regional energy consumption, and suggestions are provided for promoting regional energy conservation by summarizing the disparities of energy consumption of the four groups.

11.2 Methodology and Data Input

In this section, a hybrid energy input-output model is established to study the direct and indirect energy consumption of 30 regions in China in 2007.

11.2.1 The Hybrid Energy Input-Output Model

The hybrid energy input-output model was established using data from regional input-output tables and energy balance tables (PCESY 2008b; PCESY 2008a; PCSY 2010b; PCSY 2010a).

The original regional input-output tables consist of monetary flows of intermediate use, final use, flow-in and gross output, as shown in Table 11.1.

There are 42 intermediate sectors in each regional input-output table, including agriculture, industry and tertiary sectors. In Table 11.1, intermediate use Z_{ij}^0 refers to the monetary valued flow from sector i to sector j, which represents the number of products of sector i used for sector j's production. Final use FU_i^0 refers to the monetary valued flow from sector i to final consumption expenditures, gross capital formation and flow-outs. Flow-in FI_i^0 refers to the monetary valued flow of sector i's products, which are imported from other regions. Gross output GO_i^0 refers to the gross output of sector i in this region.

$$GO_i^0 = \sum_{j=1}^{42} Z_{ij}^0 + FU_i^0 - FI_i^0 \tag{11.1}$$

Value added VA_i^0 refers to the labor compensation, net production taxes, depreciation of fixed assets, and operating surplus of sector i. Total input TI_i^0 refers to the total amount of monetary valued input from the other sectors to sector i.

The original regional energy balance tables consist of mass flows of varied types of energy for primary energy supply, energy transformation, and final energy consumption.

Table 11.1 Original composition of regional input-output table

		Intermediate use				Final use	Flow-in	Gross output
		Sector 1	Sector 2	…	Sector 42			
Intermediate sectors input	Sector 1	Z_{11}^0	Z_{12}^0	…	$Z_{1,42}^0$	FU_1^0	FI_1^0	GO_1^0
	Sector 2	Z_{21}^0	Z_{22}^0	…	$Z_{2,42}^0$	FU_2^0	FI_2^0	GO_2^0
	…	…	…	…	…	…	…	…
	Sector 42	$Z_{42,1}^0$	$Z_{42,2}^0$	…	$Z_{42,42}^0$	FU_{42}^0	FI_{42}^0	GO_{42}^0
Value added		VA_1^0	VA_2^0	…	VA_{42}^0			
Total input		TI_1^0	TI_2^0	…	TI_{42}^0			

To establish a hybrid energy input-output model, part of the monetary flows in original input-output tables need to be replaced with energy flows. As the sector classification in energy balance tables is different from that in input-output table, first, the 42 sectors in input-output tables are combined into 10 "combined" sectors to adapt the sector classification in energy balance tables. After combination, the monetary flows of combined sectors are summed, and the combined sectors are listed in Table 11.2.

In order to analyze the regional energy consumption, energy providing and consuming sectors need to be distinguished. These 10 "combined" sectors can be categorized into 4 energy sectors (energy providing sectors) and 6 non-energy sectors (energy consuming sectors). Energy sectors include mining and washing of coal; extracting of petroleum and natural gas; processing of petroleum, coking, and nuclear fuel; and production and supply of electric power and heat power. The others are non-energy sectors such as agriculture.

The monetary valued flows of energy sectors in the input-output table were replaced by energy valued flows, which were calculated using the data from regional energy balance tables. Regional energy balance tables provide the data of indigenous production, imports and exports, inputs and outputs of transformation, and final consumption of coal, oil, natural gas, heat and electricity. First, the physical quantity data in the energy balance tables are converted into energy values measured by energy unit, tce (ton coal equivalent). The categories in the original energy balance tables are then combined to match the sector categories in the regional input-output tables.

Finally, the monetary flows of the energy sectors in the re-organized regional input-output tables are replaced with the energy flows in the combined energy balance tables.

Following the aforementioned steps, we organize a hybrid energy input-output table for each of the 30 regions with the regional input-output tables and energy balance tables, which were ready for further analysis. The contents of a hybrid energy input-output table are presented in Table 11.3, and the symbols in the table are explained as follows:

- E_{ij}: the energy flow from sector i to sector j.
- Z_{ij}: the monetary flow from sector i to sector j.
- ER_i: the energy flow from sector i to the final consumption on expenditure, i.e., the energy consumed for residential and governmental use.
- R_i: the monetary flow from sector i to the final consumption on expenditure.
- EC_i: the energy flow from sector i to the gross capital formation.
- C_i: the monetary flow from sector i to the gross capital formation.
- EO_i: the energy flow exported from sector i to other regions.
- O_i: the monetary flow exported from sector i to other regions.
- EI_i: the energy flow imported from other regions to sector i.
- I_i: the monetary flow imported from other regions to sector i.
- EX_i: the total production of energy sector i.
- X_i: the total production of non-energy sector i.

Table 11.2 Sector combination of 42 sectors

Original sectors	Combined sectors
Mining and washing of coal	Mining and washing of coal
Extraction of petroleum and natural gas sector Production and supply of gas sector	Extraction and supply of petroleum and natural gas
Processing of petroleum, coking, and nuclear fuel	Processing of petroleum, coking, and nuclear fuel
Production and supply of electric power and heat power	Production and supply of electric power and heat power
Agriculture, forestry, animal husbandry and fishery sector Management of water conservancy, environment and public facilities sector	Agriculture, forestry, animal husbandry, fishery and water conservancy
Mining of metal ores Mining of non-metal ores and other ores Manufacturing of food, beverage and tobacco Manufacturing of textiles Manufacturing of textile wearing apparel, footwear and caps Processing of timber Manufacturing of wood and furniture, manufacturing of paper and paper products Printing Reproduction of recording media Chemical industry Manufacturing of non-metallic mineral products, smelting and pressing of metals Manufacturing of metal products Manufacturing of general purpose machinery and special purpose machinery Manufacturing of transport equipment, manufacturing of electrical machinery and equipment manufacturing of communication equipment, computers and other electronic equipment Manufacturing of measuring instruments and machinery for cultural activity and office work Manufacturing of artwork Other manufacturing Recycling and disposal of waste Production and supply of water	Non-energy industry
Construction	Construction
Transport and storage sector Post sector	Transport, storage and post
Wholesale and retail trades sector Hotels and catering services sector	Wholesale, retail trades, hotels and catering services

(continued)

Table 11.2 (continued)

Original sectors	Combined sectors
Information transmission	Other sectors
Computer services and software	
Financial intermediation	
Real estate leasing	
Business services	
Scientific research and technical services	
General technical services	
Services to households and other services	
Education	
Health, social security and social welfare	
Culture, sports and entertainment	
Public management and social organizations	

The values of the first four rows in Table 11.3 are measured using tonne coal equivalent (tce), and the values of the last six rows are measured originally in units of 10000 RMB and transformed into $10000. Thus, a regional hybrid energy input-output table is formed.

Next, by a series of matrix manipulation, we decompose the regional energy consumption into energy embodied in products and services for final consumption expenditures, gross capital formation and flow-outs. The derivational process is not mentioned here for the reason of space limit, but can be referred to in the book named Input-Output Analysis: Foundations and Extensions (Miller and Blair 1985), for the readers who are interested.

11.2.2 The Decomposition of Regional Energy Consumption

In the Provincial Energy Statistical Yearbook, the energy supply comes from flow-in and self-produced energy, and regional energy consumption is categorized as final use, processing input, stock change and flow-out (PCESY 2008b). Final use includes the energy consumption for urban and rural residential use. Processing input includes the energy used in farming, forestry, animal husbandry, fishery and water conservancy, industry, construction, transport, storage and post, wholesale, retail trade and hotel, restaurants, and other sectors. Stock change refers to the change of regional energy storage and to the net increase of energy product storage. Flow-out refers to the energy that is exported to other regions or used for the refueling of foreign airplanes and ships. The energy consumption category in the China Energy Statistical Yearbook is illustrated in Fig. 11.1.

Table 11.3 Contents of a hybrid energy input-output table

	Sector 1, 2, …,10	Final use			Flow-in	Total output
		Final consumption on expenditure	Gross capital formation	Flow-out		
Mining and washing of coal	E_{11}, E_{12}, …, $E_{1,10}$	ER_1	EC_1	EO_1	EI_1	EX_1
Extraction and supply of petroleum and natural gas	E_{21}, E_{22}, …, $E_{2,20}$	ER_2	EC_2	EO_2	EI_2	EX_2
Processing of petroleum, coking, nuclear fuel	E_{31}, E_{32}, …, $E_{3,10}$	ER_3	EC_3	EO_3	EI_3	EX_3
Production and supply of electric power and heat power	E_{41}, E_{42}, …, $E_{4,10}$	ER_4	EC_4	EO_4	EI_4	EX_4
Agriculture, forestry, animal husbandry, fishery and water conservancy	Z_{51}, Z_{52}, …, $Z_{5,10}$	R_5	C_5	O_5	I_5	X_5
Non-energy industry	Z_{61}, Z_{62}, …, $Z_{6,10}$	R_6	C_6	O_6	I_6	X_6
Construction	Z_{71}, Z_{72}, …, $Z_{7,10}$	R_7	C_7	O_7	I_7	X_7
Transport, storage and post	Z_{81}, Z_{82}, …, $Z_{8,10}$	R_8	C_8	O_8	I_8	X_8
Wholesale, retail trades, hotels and catering services	Z_{91}, Z_{92}, …, $Z_{9,10}$	R_9	C_9	O_9	I_9	X_9
Other	$Z_{10,1}$, $Z_{10,2}$, …, $Z_{10,10}$	R_{10}	C_{10}	O_{10}	I_{10}	X_{10}

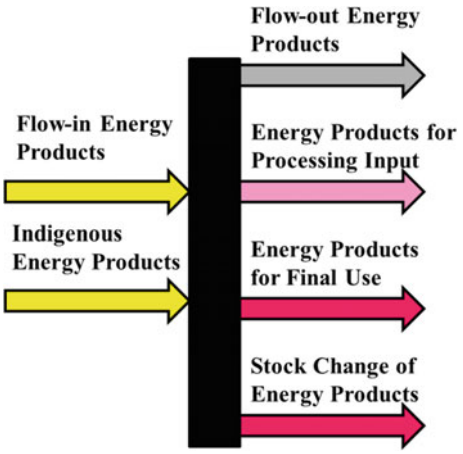

Fig. 11.1 Energy supply and consumption in energy balance statistical Yearbook

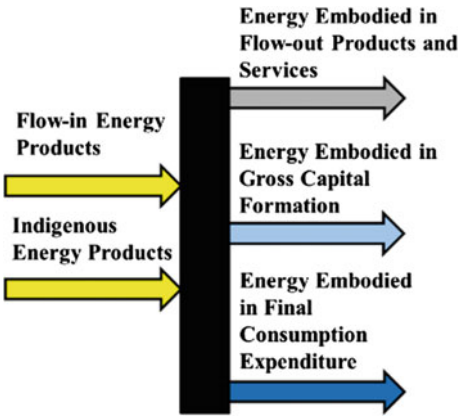

Fig. 11.2 Decomposed embodied energy with input-output model: Measurement A

However, the statistical data cannot provide information about how energy consumption contributes to the regional residential life and economic development, as a large amount of primary energy is not consumed directly but converted into secondary energy carriers or embodied in various non-energy products and services. Therefore, we decompose the regional energy consumption with an input-output method and consider it as the energy embodied in goods and services for final consumption expenditure, gross capital formation and flow-out. Accordingly, we determine how energy is consumed to support residential life, progress of urbanization and industrialization, and inter-regional trades. This measurement for regional energy consumption is defined as Measurement A in this study, as illustrated in Fig. 11.2.

In Measurement A, regional energy consumption is defined as the energy embodied in regionally produced goods and services. On the other hand, we define the regional energy consumption as the energy embodied in regionally consumed

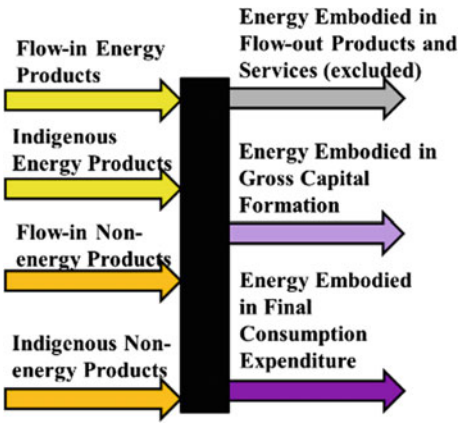

Fig. 11.3 Decomposed embodied energy with input-output model: Measurement B

goods and services for final consumption expenditure and gross capital formation. Accordingly, regional energy consumption refers to the energy embodied in indigenous and flow-in goods and services (including energy and non-energy products and services) that are ultimately consumed in the region for final use expenditure and gross capital formation. Meanwhile, information about where the flow-in goods and services are produced is not provided in the regional input-output tables, thus the energy embodied in flow-in goods and services is unknown. Herein, we assume that the energy embodied in unit flow-in goods and services (monetary valued) of one region is the same as the national average of the same type of product or service. In our consideration, this assumption is acceptable because in the energy-intensive industries in China, the levels of the energy utilization rates of a sector are similar, thus the deviation from region to region is not relevant. For a typical region, part of the energy embodied in flow-out products and services is sent to other regions in China, while the remaining energy is exported to foreign countries. This measurement for regional energy consumption is defined as Measurement B in this study and is illustrated in Fig. 11.3. As Measurement B only considers the regionally consumed embodied energy, the sum of 30 regions' embodied energy consumption would be less than that of Measurement A. The difference is the embodied energy in goods and services that are exported abroad. In Measurement B, energy embodied in flow-out products and services is not considered as regional energy consumption, as the embodied energy is exported and consumed in other regions or countries.

11.3 Results and Discussion

Based on the decomposition results of the energy consumption of 30 regions in China, we conducted several comparisons to determine China's regional disparities in energy consumption.

11.3.1 Overview of Per Capita Primary Energy Consumption in 30 Regions

The energy balance statistics provide information of direct primary energy consumption for final use, processing input and flow out. The regional energy consumption in 2007 according to Provincial Energy Statistical Yearbook 2008 is illustrated in Fig. 11.4.

In Fig. 11.4, the per capita energy consumption of 30 regions (column) is ranged from left to right by their per capita GDPs (PPP, current international $) of 2007. The solid bars represent per capita primary energy consumption for final use and stock change. The shaded bars represent per capita primary energy consumption for processing input. The blank bars represent per capita primary energy products that flowed out of the region. Generally, the result presents the following implications:

1. The per capita energy consumption in most economically developed regions (Shanghai, Beijing and Tianjin, whose per capita GDPs are higher than $10000) is relatively high.
2. The per capita energy consumption in most undeveloped regions (Chongqing, Ningxia, Shaanxi, Hainan, Hunan, Qinghai, Sichuan, Jiangxi, Guangxi, Anhui, Yunnan, Gansu and Guizhou, whose per capita GDPs are lower than $5000) is relatively low.
3. The per capita energy consumption of regions with abundant primary energy resources (e.g., Inner Mongolia and Shanxi) is the highest, with a large share of the primary energy consumption for flow-out.
4. For most regions in China, a large share of primary energy is directly consumed for processing input, indicating that industry was the largest energy consumer in China.

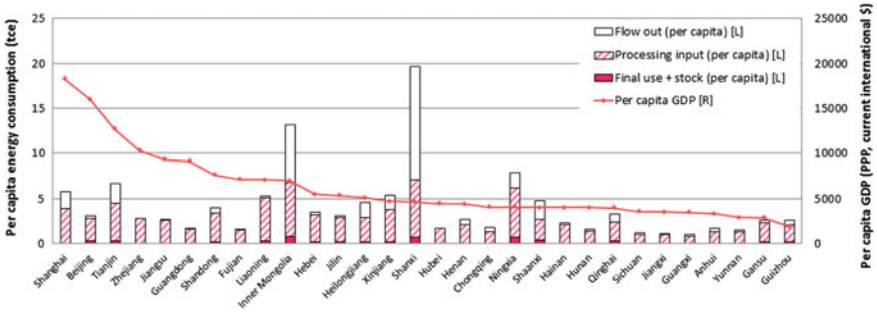

Fig. 11.4 Energy consumption statistics of 30 regions in China (2007) (All GDP numbers with units in US$ are based on purchasing power parity (PPP) and measured with current international $ according to the database of the World Bank.) ([L]: *left vertical* ordinate; [R]: *right vertical* ordinate)

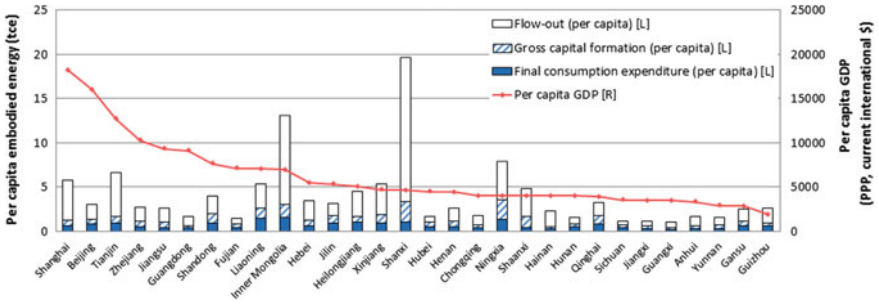

Fig. 11.5 Embodied energy consumption of 30 regions in China (Measurement A, 2007) ([L]: *left vertical* ordinate; [R]: *right vertical* ordinate)

However, the energy balance data alone cannot reflect the characteristics of the per capita energy consumption of various regions for final consumption expenditure, gross capital formation and exports of products and services.

To further understand how energy consumption contributes to the regional economy, we decompose the regional energy consumption with a hybrid input-output model and measure the decomposed embodied energy using two measurements. With Measurement A, the regional energy consumption is decomposed into energy embodied in goods and services for final consumption expenditure, gross capital formation and flow-out, as illustrated in Fig. 11.5.

In Fig. 11.5, the solid bars represent per capita embodied energy in products and services for final consumption expenditure, the shaded bars represent per capita embodied energy in products and services for gross capital formation, and the blank bars represent per capita embodied energy in flow-out products and services. In Fig. 11.5, the sum of per capita energy embodied in final consumption expenditure, gross capital formation and flow-out is the same as that in Fig. 11.4. However, the decomposed embodied energy further reveals how energy consumption contributes to residential life and economic development of the region.

1. Although the direct primary energy consumption for final use is small in most regions (see Fig. 11.4), the energy embodied in goods and services for final consumption expenditure is rather large (see Fig. 11.5), indicating that a large amount of energy is embodied in non-energy products that contribute to residential life.
2. Energy embodied in gross capital formation refers to the energy consumed for infrastructure construction and facility building. The developing regions whose per capita GDPs are between $5000 to $10000, have a substantial amount of per capita energy embodied in gross capital formation, indicating that these regions are in the process of industrialization and urbanization. Meanwhile, developed regions (the left 6 regions in Fig. 11.5) and under-developed regions (the right 10 regions in Fig. 11.5) has small embodied energy of gross capital formation. For developed regions, this indicates that the industrialization and urbanization

is close to complete. For under-developed regions, this indicates that they are still at the beginning of industrialization and urbanization.

3. Energy embodied in flow-out products and services varies region by region because different regions export different types of products and services. For Shanghai and Tianjin, the large per capita embodied energy in flow-out indicates that large quantities of energy-intensive goods (e.g., energy products, steel, chemicals, etc.) are exported from Shanghai Port and Tianjin Port. For the regions with abundant energy resources, e.g., Inner Mongolia and Shanxi, the large per capita embodied energy in flow-out indicates that these regions export large quantities of energy products, mostly coal and coal-derived products for Inner Mongolia and Shanxi, together with some non-energy products.

Using Measurement A, the decomposed embodied energy implies how primary energy consumption contributes to the regional economy. However, Measurement A does not present the actual energy consumed for residential life and ongoing urbanization and industrialization of the region, as flow-in non-energy products and services are excluded. Therefore, we conduct Measurement B, which considers the flow-in non-energy products and services as part of the regional energy consumption. The results are illustrated in Fig. 11.6.

Using Measurement B, regional energy consumption is decomposed into energy embodied in products and services for final consumption expenditure and gross capital formation. The products and services include regionally produced and imported products and services. With this measurement, energy embodied in flow-out products and services are not included in regional energy consumption. In Fig. 11.6, the solid bars and shaded bars, respectively, represent the primary energy embodied in final consumption expenditure and gross capital formation of the region. Through such decomposition, we determine how energy consumption contributes to the regional residential life and the ongoing urbanization and industrialization progress by being embodied in goods and services for final consumption expenditure and gross capital formation.

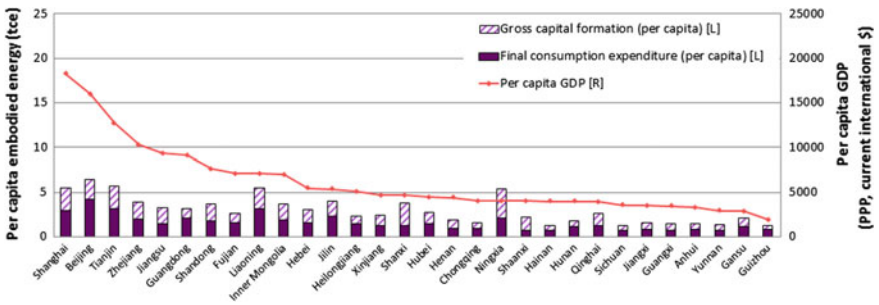

Fig. 11.6 Embodied energy consumption of 30 regions in China (Measurement B, 2007) ([L]: *left vertical* ordinate; [R]: *right vertical* ordinate)

1. Generally, economically developed regions (Shanghai, Beijing, Tianjin) have high per capita embodied energy in products and services for both final consumption expenditure and gross capital formation.
2. For most of the developing regions whose per capita GDPs are between $5000 to $10000, the sum of per capita embodied energy for final consumption and gross capital formation is lower than that of developed regions.
3. With respect to some of the developing regions, the per capita embodied energy for final consumption is higher than that for gross capital formation (e.g., Guangdong, Fujian, Liaoning, Jilin, Heilongjiang), thus indicating that these regions are developing tertiary industry, for example, the information transmission service.
4. For some other developing regions, the per capita embodied energy for gross capital formation is higher than that for final consumption (e.g., Zhejiang, Jiangsu, Shandong, Inner Mongolia, Shanxi), indicating that these regions are developing a significant industrial base.
5. For under-developed regions with per capita GDPs that are lower than $10000, both the per capita energy embodied in products and services for final consumption expenditure and gross capital formation is low, indicating low energy consumption levels in these regions.

The comparison of embodied energy consumption calculated using Measurements A and B is illustrated in Fig. 11.7.

As mentioned in Sect. 11.2.2, the total embodied energy of 30 regions measured with Measurement B would be less than that measured with Measurement A because energy embodied in exported products and services (including those products exported to foreign countries) are not included in Measurement B. Thus, in Fig. 11.7, per capita embodied energy with Measurement B refers to the embodied energy in regionally consumed products and services for final consumption expenditure and gross capital formation, while per capita embodied energy with Measurement A refers to the embodied energy in regionally produced products and services for final consumption expenditure, gross capital formation and flow-out. With Measurement A,

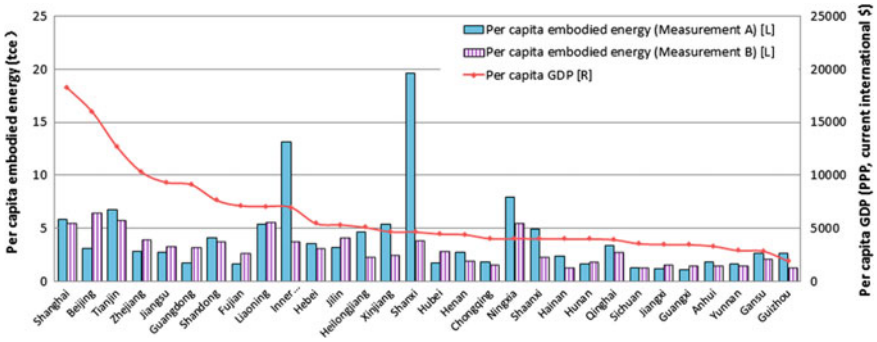

Fig. 11.7 Comparison of per capita embodied energy calculated with Measurement A and B ([L]: *left vertical* ordinate; [R]: *right vertical* ordinate)

per capita embodied energy implicates the energy consumption contributed to regional economy, while with Measurement B, per capita embodied energy implicates the energy consumption contributed to regional residential life and infrastructure construction. The implications of Fig. 11.7 are as follows:

1. The greatest deviation between the two measurements occurs in Shanxi and Inner Mongolia because Shanxi and Inner Mongolia produce and export large quantities of coal, while the actual energy consumption for regional final consumption expenditure and gross capital formation is rather low. This explanation can also be applied to the deviations of Ningxia, Shaanxi, Heilongjiang and Xinjiang, which are the major energy providing regions in China.
2. For Beijing, per capita embodied energy consumption with Measurement B is almost twice that with Measurement A. This indicates that as an economically developed region, Beijing is consuming much more energy than it provided, absorbing large quantities of energy from other regions. With Measurement B, Shanghai and Tianjin have per capita embodied energy consumption similar to that of Beijing, indicating their regional energy consumption is high. However, Shanghai and Tianjin also export large amounts of energy-intensive products as they are important port cities in China, thus resulting in high per capita embodied energy consumption with Measurement A.
3. For some of the developing regions whose per capita GDPs are between $5000 to $10000, e.g., Jiangsu, Zhejiang and Guangdong, the per capita embodied energy consumption with Measurement B is higher than that with Measurement A, indicating that these regions are consuming much energy than they provided Considering that these regions are in the progress of urbanization and industrialization, while the embodied energy in final consumption expenditure and gross capital formation are either not high (see Fig. 11.6), this may due to relatively high energy use efficiency, which we will discuss in later sections.
4. For some other developing regions, e.g., Shandong, Hebei and Heilongjiang, the per capita embodied energy consumption with Measurement A is higher than that with Measurement B. Compared with Jiangsu, Zhejiang and Guangdong, Shandong is also in a rapid progress of industrialization and urbanization but higher embodied energy consumption in final consumption expenditure, gross capital formation and flow-out (see Fig. 11.5). This indicates lower energy use efficiency, which will be discussed in later sections. For Hebei and Heilongjiang, the per capita embodied energy is relatively higher due to high embodied energy in flow-out. This is because these regions produce and export energy intensive products to other regions.
5. For most under-developed regions whose per capita GDPs are lower than $5000, the per capita embodied energy consumption measured with Measurement B and Measurement A are both low, indicating that these regions have low energy consumption level for residential life and that the industrial structure is dominated by low energy consuming industry. The characteristics of energy consumption in these regions are further discussed with other index in following sections, taking into account of industrial structure and economic development.

11.3.2 Comparison of Energy Consumption Intensity in 30 Regions

In Sect. 11.3.1, we provided an overview of the primary energy consumption of 30 regions in China. By comparing the decomposition results of the 30 regions, we determine that from the view of embodied energy consumption, economically developed regions consume much more energy than indicated by the statistics in the Energy Statistical Yearbook, under-developed regions consume less energy than that indicated by the statistics in the Energy Statistical Yearbook, while the situation of developing regions is complicated and further analysis is required.

To further understand the connection between energy consumption and regional economic development, Fig. 11.8 illustrates the decomposed embodied energy intensity of 30 regions in China. In the following discussions, we define energy intensity as the embodied energy consumption for unit output value.

In Fig. 11.8, embodied energy intensity refers to the embodied energy consumption divided by corresponding output value. The horizontal axis represents the embodied energy intensity of final consumption expenditure, the vertical axis represents the embodied energy intensity of gross capital formation, the size of shaded bubbles represents the embodied energy intensity of flow-in products and services, and the size of solid bubbles represents the embodied energy intensity of flow-out products and services. The dashed line in the figure marks the positions where the embodied energy intensity of final consumption expenditure is equal to that of gross capital formation. By the position and size of the bubbles in Fig. 11.8, we can readily observe the decomposed embodied energy intensity of 30 regions in China.

Generally, the bubbles locate near the dashed line, meaning that in these regions, the energy intensity in residential life and infrastructure construction is close. Generally, 4/5 of the 30 regions' bubbles locate near the dashed line, indicating that

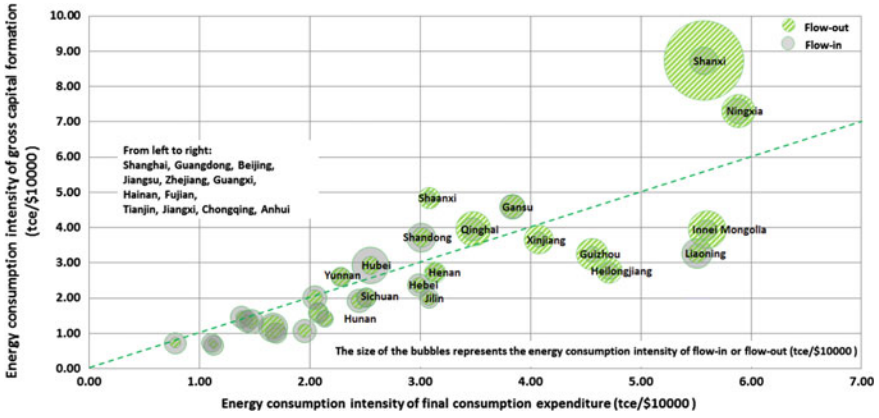

Fig. 11.8 Embodied energy intensity of 30 provinces and regions in China, 2007

for one region, the energy intensity of final consumption expenditure and gross capital formation is at a similar level, and the bubbles located in the top right area have higher energy intensity of final consumption expenditure and gross capital formation than those located in the left bottom area Besides, the flow-out bubble size increases from the left bottom to the right top in the figure, indicating that the embodied energy intensity of flow-out of the regions located in the top right area in Fig. 11.8 is higher than for those regions located in the left bottom area. However, the flow-in bubble size in the middle area in Fig. 11.8 is relatively larger than others, indicating higher embodied energy intensity of flow-in of those regions. This indicates that those regions may be importing energy-intensive products, which will be further discussed next. According to the bubble size and distribution in Fig. 11.8, as well as the regional per capita GDPs, we categorize the 30 regions into four groups and discuss the common characteristics of each group.

There are 9 bubbles located in the left bottom area in Fig. 11.8, of which the energy intensity of final consumption expenditure and gross capital formation is lower than 2.00 tce/$10000 (Shanghai, Guangdong, Beijing, Jiangsu, Zhejiang, Guangxi, Hainan, Fujian, and Tianjin). Of these regions, the per capita GDP of Guangxi and Hainan is lower than $5000, while the others' is higher than $7000. We define Group 1 as the economically developed regions with low embodied energy intensities, of which the regional per capita GDP is higher than $7000 and the energy intensity of final consumption expenditure and gross capital formation is lower than 2.00 tce/$10000.

There are 15 bubbles located in the area of which the energy intensity of final consumption expenditure and gross capital formation is between 1.00–3.00 tce/$10000. 6 of the 15 regions have been grouped into Group 1, and the per capita GDP of the left 9 regions is lower than $5000. Therefore, these 9 regions (Guangxi, Hainan, Jiangxi, Chongqing, Anhui, Yunnan, Hunan, Sichuan and Hubei) are categorized into Group 2 defined as the economically under-developed regions with low energy intensities.

The left 14 bubbles located in the area of which the energy intensity of final consumption expenditure is between 3.00–6.00 tce/$10000, and the energy intensity of gross capital formation is higher than 2.00 tce/$10000. For these regions, the energy intensity is relatively higher than those in Group 1 and Group 2, and the bubble distribution of these 14 regions in Fig. 11.8 is more decentralized. Of these regions, we notice that 8 of them have larger flow-out bubbles than flow-in bubbles, indicating high energy intensity of flow-out products and services. Therefore, we categorized the 8 regions (Shaanxi, Qinghai, Xinjiang, Guizhou, Heilongjiang, Shanxi, Inner Mongolia and Ningxia) into Group 4 defined as the economically developing and energy exporting regions with high energy intensities, and the left 6 regions (Hebei, Shandong, Jilin, Henan, Gansu and Liaoning) into Group 3 defined as the economically developing and energy absorbing regions with high energy intensities.

To benchmark the economic development level of the grouped regions, we select several countries and regions' per capita GDPs in 2007 as references. The per capita GDP range of each group is illustrated in Fig. 11.9.

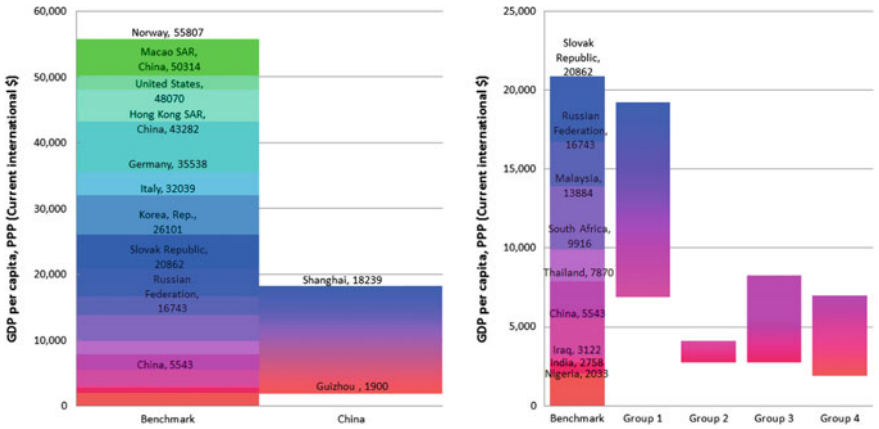

Fig. 11.9 GDP per capita, PPP of benchmark countries and four region groups in China, 2007 (Data source: (WB 2013))

In China, Guizhou demonstrates the lowest regional per capita GDP ($1900, PPP), while Shanghai has the highest regional per capita GDP ($18239, PPP) in 2007, as illustrated in Fig. 11.9. Shanghai is at an economic development level above Russia and below Slovak Republic, while Guizhou is at a level similar to that of Nigeria.

11.3.2.1 Group 1: The Economically Developed Regions with Low Embodied Energy Intensities

The first group includes Shanghai, Guangdong, Beijing, Jiangsu, Zhejiang, Fujian and Tianjin. The embodied energy intensities of the 7 regions are illustrated in Fig. 11.10.

The seven regions in Fig. 11.10 are primarily located in the southeast coastal area of China, with per capital GDPs between $7000 and $20000 (PPP) in 2007. Referencing Fig. 11.9, it is evident that the economic development level of these regions, which are economically developed regions in China, is similar to that of South Africa, Malaysia and Russia. The bubbles in Fig. 11.10 are located near the ordinate origin, indicating that the embodied energy intensity of final consumption expenditure and gross capital formation in these regions is small. Considering the relatively high embodied energy for final consumption expenditure and gross capital formation in these regions (see Fig. 11.6), the small embodied energy intensity is greatly due to high output value, reflecting developed economy and high spending life in these regions. Furthermore, it is determined that the embodied energy intensity of flow-in products and services is 3–6.5 times that of flow-out, indicating that these regions are energy absorbing regions that import cheap, energy-intensive products and export expensive products with little primary energy

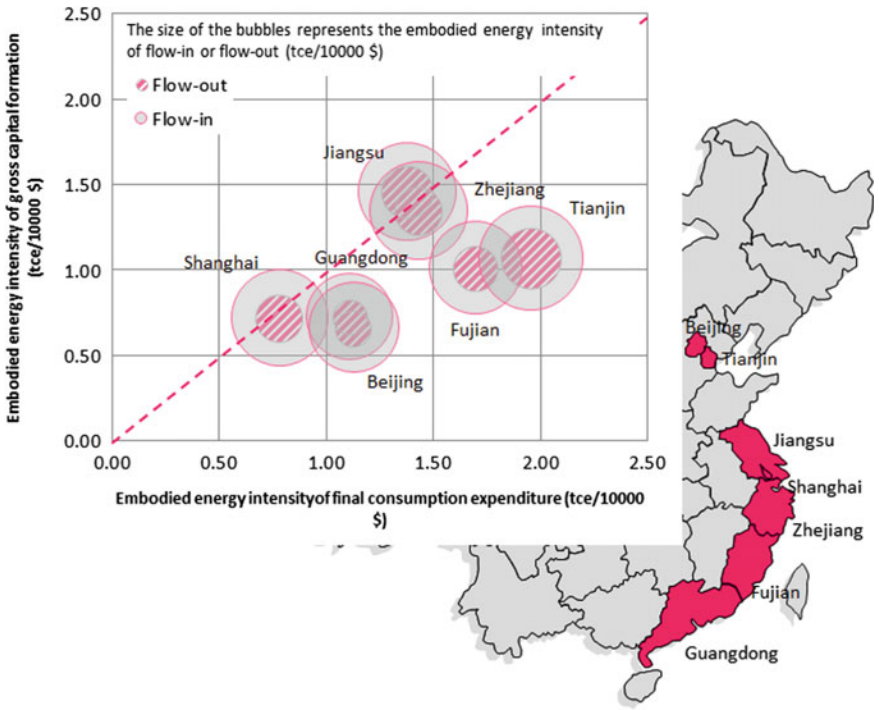

Fig. 11.10 Embodied energy intensity of Shanghai, Guangdong, Beijing, Jiangsu, Zhejiang, Fujian and Tianjin (2007)

embodied. Geographically, these regions are primarily located in the southeast coastal areas, where there is a lack of energy resources. Meanwhile, their industry is dominated by either tertiary industry (for example, the financial intermediate service in Beijing, Shanghai) or light industry (for example, the textile industry in Jiangsu, Zhejiang, etc.). In other words, these regions do not produce energy-intensive products and services, while consume energy-intensive products and services provided by other regions to produce high value-added products and services to support the developed economy and high spending, high energy-consuming life style. For these regions, energy conservation should focus on promoting low energy-consuming life style.

11.3.2.2 Group 2: The Economically Under-Developed Regions with Low Energy Intensities

The second group includes Guangxi, Hainan, Jiangxi, Chongqing, Anhui, Yunnan, Hunan, Sichuan and Hubei. The embodied energy intensities of the 9 regions are illustrated in Fig. 11.11.

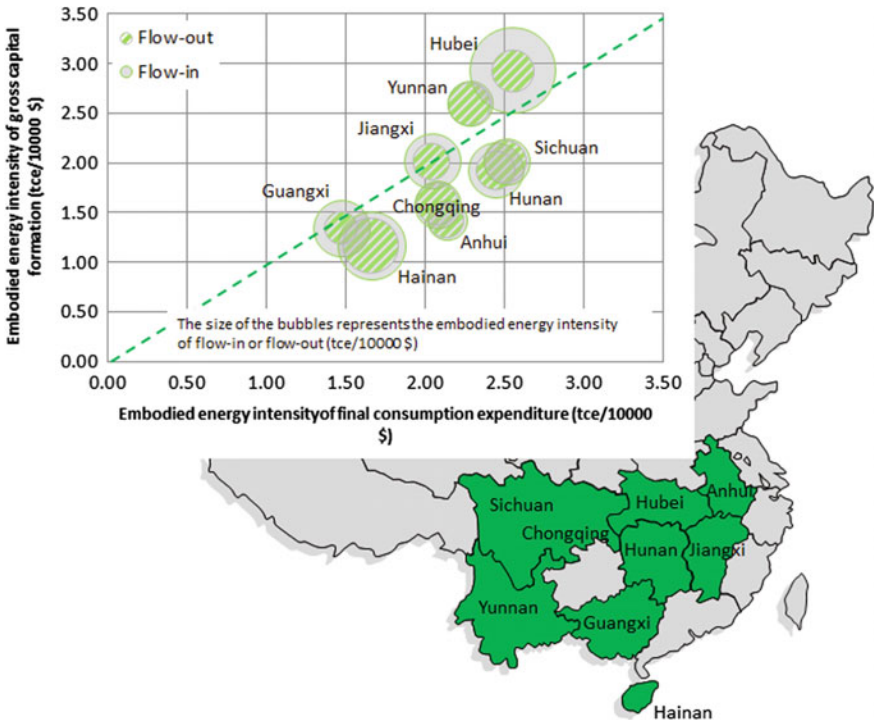

Fig. 11.11 Embodied energy intensity of Guangxi, Hainan, Jiangxi, Chongqing, Anhui, Yunnan, Hunan, Sichuan and Hubei, 2007

The nine regions in Fig. 11.11 are in southern China, and there per capital GDPs are between $2000 and $5000 (PPP) in 2007. Be examining Fig. 11.9, it is evidence that the economic development level of these regions, which are economically under-developed, is similar to that of India and Iraq. The bubbles of these regions are located at the left bottom of Fig. 11.8, indicating that the embodied energy intensity of final consumption expenditure and gross capital formation in these regions is small. Considering the low energy embodied in final consumption expenditure and gross capital formation of the regions in Group 2 (see Fig. 11.6), the small energy intensity is due to low output value for residential life and infrastructure construction, reflecting under-developed economy and rather low-spending life in these regions. In Fig. 11.11, the embodied energy intensity of flow-out is almost equal to that of flow-in, with the exception of Hubei and Guangxi Provinces, thus indicating balanced energy intensity for trade. The per capita embodied energy consumption is also small (see Fig. 11.7) in these regions, indicating that these regions did not produce or consume large quantities of energy-intensive products. Geographically, in these regions, located in southern China, agriculture dominates the regional economy. For Guangxi Province, tobacco exports accounted for a large share of the regional economy, which is high

value-adding and low energy-intensity, and this leads to higher energy intensity for flow-out than flow-in products. Hubei Province is in the progress of changing from an agriculture-dominated economy to an industry-dominated economy. Considering that the per capita embodied energy consumption in Hubei Province is small (see Fig. 11.7), the relatively high energy intensity indicates that Hubei has low energy intensity on both the energy supply and energy consumption sides. For these regions, economy development is more urgent than energy conservation. However, high energy use efficiency during industrialization and urbanization progress should also be taken into account.

11.3.2.3 Group 3: The Economically Developing and Energy Absorbing Regions with High Energy Intensities

The third group includes Hebei, Shandong, Jilin, Henan, Gansu and Liaoning. The embodied energy intensity of these 6 regions is illustrated in Fig. 11.12.

The six regions in Fig. 11.12 are in northern China. In 2007, the per capita GDPs of Shandong and Liaoning are between $7000 and $8000 (PPP), and the per capita GDPs of the other four regions in the figure are between $2000 and $6000 (PPP).

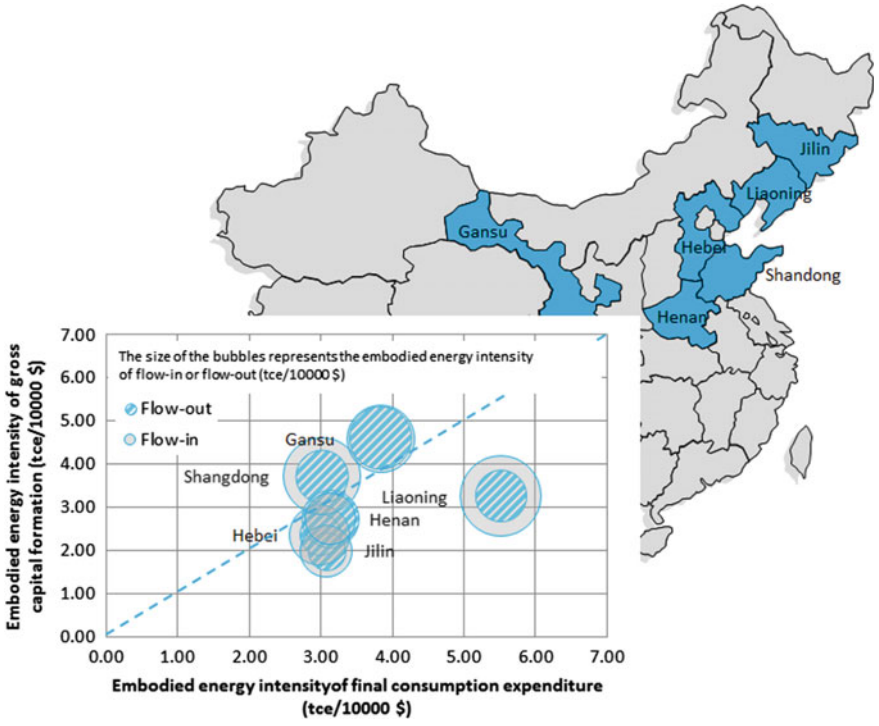

Fig. 11.12 Embodied energy intensity of Hebei, Shandong, Jilin, Henan, Gansu, Liaoning, 2007

Based on Fig. 11.9, it is evident that these regions are on an economic development level similar to Iraq and Thailand. These regions are considered economically developing regions. The bubbles of these regions are located in the central part of Fig. 11.8, meaning that the embodied energy intensity of final consumption expenditure and gross capital formation in these regions are near the average level of that of the 30 regions in China. Regarding Shandong, Liaoning, Hebei and Henan, the per capita embodied energy consumption is large (see Fig. 11.7), and the bubble size of these four regions is also large, indicating that these regions produce and consume large quantities of energy-intensive products. With respect to Liaoning and Shandong, the flow-in energy intensity is more than twice the flow-out energy intensity, indicating that these regions are energy absorbing regions with developing heavy industry. Regarding Jilin and Gansu, the per capita embodied energy consumption was small (see Fig. 11.7), and the large bubble size in Fig. 11.12 indicates low energy intensity. Generally, the regions in Group 3 are in rapid industrialization and urbanization progress, consuming large quantities of energy. At the same time, energy for residential life and infrastructure construction in these regions is not efficiently consumed. For these regions, energy conservation should focus on improving energy use efficiency in industries.

11.3.2.4 Group 4: The Economically Developing and Energy Exporting Regions with High Energy Intensities

The fourth group includes Shaanxi, Qinghai, Xinjiang, Guizhou, Heilongjiang, Shanxi, Inner Mongolia and Ningxia. The embodied energy intensities of the 8 regions are illustrated in Fig. 11.13.

The majority of the eight regions in Fig. 11.13 are located in the north of China, with the exception of Guizhou, which is in southern China. In 2007, the per capita GDP of Inner Mongolia was $6978.5 (PPP), the per capita GDP of Guizhou was $1900 (PPP), and the per capita GDP of the other six regions was between $2000 and $6000 (PPP). Referring to Fig. 11.9, we note that, with the exception of Guizhou, which is the poorest region and at an economic level similar to that of Nigeria, these regions are at an economic level similar to the economic development level with Iraq and Thailand. These eight regions are categorized in the same group because the embodied energy intensity of flow-out was larger than that of flow-in in these regions, a factor that distinguishes these regions from the other 22 regions in China. This phenomenon indicates that these regions produce energy-related products (mainly coal and coal-derived products) and that they are the dominant energy suppliers of the country. Meanwhile, the bubbles of these provinces located in the right top area in Fig. 11.8, indicate that the embodied energy intensity of final consumption expenditure and gross capital formation in these regions is lower than it is in other regions. For the energy supplying regions, except Shanxi, in addition to energy-relevant products, they exported various products that were less energy intensive and higher value-added, which resulted in lower embodied energy intensity of flow-out than Shanxi. For example, Inner

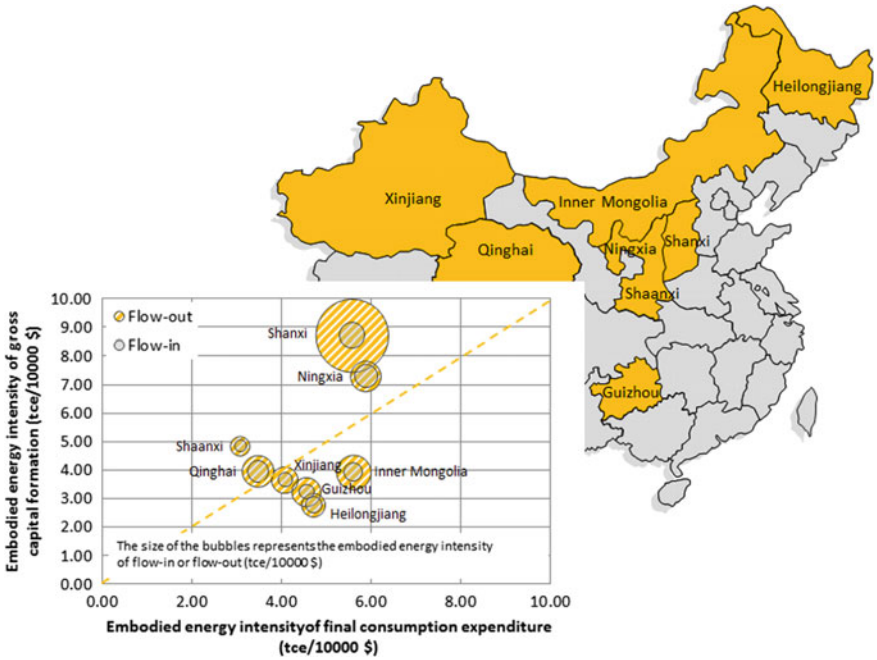

Fig. 11.13 Embodied energy intensity of Shaanxi, Qinghai, Xinjiang, Guizhou, Heilongjiang, Shanxi, Inner Mongolia, Ningxia, 2007

Mongolia exported expensive woolen products, and Ningxia exported precious medicinal plants. However, in Shanxi Province, the economic growth largely depends on exporting energy-related products. Although Shanxi has made attempts to diversify its economy by improving infrastructure construction and residential life, the energy intensity of its industry remained low. In these regions, the regional energy consumption is rather low compared with other more economically developed regions, and the exporting energy-relevant products do not bring about economic benefits. For these regions, transforming primary energy products into high value-added products would be a more economical way than exporting primary energy products directly. At the same time, high energy use efficiency during industrialization and urbanization progress should also be taken into account.

11.4 Conclusions and Suggestions

In this paper, we have described a hybrid energy input-output model to decompose the primary energy consumption of 30 regions in China. Through this approach, we have revealed the regional disparities in energy consumption, using a comparison of per capita consumption and the composition of regional energy consumption using

various decomposition measurements. We have also classified all regions into four groups based on a discussion of their differences in economic development stage, energy intensity and inter-regional trade.

The results indicate that using different measurements of energy decomposition which consider flow-in and flow-out create different perspectives on the calculated amount and composition of regional energy consumption, because the energy embodied in inter-regional trade is hugely significant for most regions. Generally, economically developed regions have relatively high energy consumption and low energy intensity levels, the economically under-developed regions have relatively low energy consumption and low energy intensity levels, and the intermediate developing regions have relatively high energy consumption and high energy intensity levels.

By considering these regional disparities with respect to economic development, embodied energy intensity of final consumption expenditure, gross capital formation, flow-in and flow-out, the 30 regions in China are classified into four groups.

1. **The developed regions with low energy intensities**, such as Shanghai and Beijing, generally have relatively low energy intensity of final consumption expenditure and gross capital formation because their industrialization process has been nearly finished. Additionally, they have a net import of embodied energy because they import energy-intensive products and export high value-added products and services.
2. **The under-developed regions with low energy intensities**, such as Yunnan and Guangxi, generally have relatively low energy intensity of final consumption expenditure and gross capital formation because they are in the preliminary stage of industrialization. Additionally, while they have a relatively small amount of energy flow-in and flow-out, there is a balance between energy flow-in and flow-out embodied in inter-regional trade.
3. **The developing and energy absorbing regions**, such as Shandong and Liaoning, generally have relatively high energy intensity of final consumption expenditure and gross capital formation because their economic development heavily depends on heavy industries. Furthermore, they have more energy flow-in than energy flow-out embodied in inter-regional trade.
4. **The developing and energy exporting regions**, such as Shanxi and Inner Mongolia, generally have relatively high energy intensity of final consumption expenditure and gross capital formation because their economic development heavily depends on their energy industry and export of energy. As a result, it is unsurprising that they have more energy flow-out than energy flow-in embodied in inter-regional trade.

Based on the analyses and conclusions presented herein, it is recommended that regional policies regarding energy conservation in China should be, firstly, based on a comprehensive analysis that decomposes regional energy consumption with respect to regional energy balance and input-output balance. Secondly, the regional disparities in energy consumption, industrial attributes and stage of economic

development of the 4 category groups of regions should be considered when developing governmental policies.

In a next step, the analysis of regional disparities within regional and national environmental constraints should also be considered to further improve the understanding of regional disparities regarding future implementation of energy conservation policies in China.

Acknowledgments The authors gratefully acknowledge the financial support from BP Company in the scope of the Phase II Collaboration between BP and Tsinghua University.

References

Ali Akkemik, K., Göksal, K., & Li, J. (2012). Energy consumption and income in Chinese provinces: Heterogeneous panel causality analysis. *Applied Energy, 99*, 445–454.

Cellura, M., Longo, S., & Mistretta, M. (2011). The energy and environmental impacts of Italian households consumptions: An input–output approach. *Renewable and Sustainable Energy Reviews, 15*, 3897–3908.

Costanza, R. (1980). Embodied energy and economic valuation. *Science, 210*, 1219–1224.

Duchin, F. (1992). Industrial input-output analysis: Implications for industrial ecology. *Proceedings of the National Academy of Sciences, 89*, 851–855.

Fu, F., Liu, H., Polenske, K. R., & Li, Z. (2013). Measuring the energy consumption of China's domestic investment from 1992 to 2007. *Applied Energy, 102*, 1267–1274.

Guo, J., Zhang, Z., & Meng, L. (2012). China's provincial CO2 emissions embodied in international and interprovincial trade. *Energy Policy, 42*, 486–497.

Herrerias, M. J., Joyeux, R., & Girardin, E. (2013). Short-and long-run causality between energy consumption and economic growth: Evidence across regions in China. *Applied Energy, 112*, 1483–1492.

Li, Y., Sun, L., Feng, T., & Zhu, C. (2013). How to reduce energy intensity in China: A regional comparison perspective. *Energy Policy, 61*, 513–522.

Liang, Q. -M., Fan, Y., & Wei, Y. -M. (2007). Multi-regional input–output model for regional energy requirements and CO2 emissions in China. *Energy Policy, 35*, 1685–1700.

Liang, S., Wang, C., & Zhang, T. (2010). An improved input–output model for energy analysis: A case study of Suzhou. *Ecological Economics, 69*, 1805–1813.

Liao, H., Du, J., & Wei, Y. -M. (2013). Energy conservation in China: Key provincial sectors at two-digit level. *Applied Energy, 104*, 457–465.

Liu, H., Guo, J., Qian, D., & Xi, Y. (2009). Comprehensive evaluation of household indirect energy consumption and impacts of alternative energy policies in China by input–output analysis. *Energy Policy, 37*, 3194–3204.

Liu, H., Xi, Y., Guo, J., & Li, X. (2010). Energy embodied in the international trade of China: An energy input–output analysis. *Energy Policy, 38*, 3957–3964.

Ma, L., Liu, P., Fu, F., Li, Z., & Ni, W. D. (2011). Integrated energy strategy for the sustainable development of China. *Energy, 36*, 1143–1154.

Meng, L., Guo, J. E., Chai, J., & Zhang, Z. (2011). China's regional CO2 emissions: Characteristics, inter-regional transfer and emission reduction policies. *Energy Policy, 39*, 6136–6144.

Miller, R. E., & Blair, P. D. (1985). *Input-output analysis: Foundations and extensions*. New Jersey, U.S.: Prentice-Hall Inc.

Park, H. -C., & Heo, E. (2007). The direct and indirect household energy requirements in the Republic of Korea from 1980 to 2000—An input–output analysis. *Energy Policy, 35*, 2839–2851.

PCESY. (2008a). *China energy statistical yearbook 2008*. Beijing: Press of China Energy Statistical Yearbook.

PCESY. (2008b). *Provincial energy statistical yearbook 2008*. Press of China Energy Statistical Yearbook.

PCSY. (2010a). *China statistical yearbook 2010*. Beijing: Press of China Statistical Yearbook.

PCSY. (2010b). *Provincial statistical yearbook 2010*. Beijing: Press of China Statistical Yearbook.

SCC. (2012). 12th Five Year Plan of Energy Conservation & Emission Mitigation. In CHINA, S. C. O. (Ed.). Beijing: State Council of China.

SCIO. (2012). China's Energy Policy (2012) White Paper. In OFFICE, S. C. I. (Ed.). Beijing: Xinhua News Agency.

Su, B., & Ang, B. W. (2010). Input–output analysis of CO2 emissions embodied in trade: The effects of spatial aggregation. *Ecological Economics, 70*, 10–18.

Su, B., & Ang, B. W. (2011). Multi-region input–output analysis of CO2 emissions embodied in trade: The feedback effects. *Ecological Economics, 71*, 42–53.

Su, B., & Ang, B. W. (2014). Input–output analysis of CO2 emissions embodied in trade: A multi-region model for China. *Applied Energy, 114*, 377–384.

Treloar, G. J. (1997). Extracting embodied energy paths from input-output tables: Towards an input-output-based hybrid energy analysis method. *Economic Systems Research, 9*, 375–392.

Wang, Z. -H., Zeng, H. -L., Wei, Y. -M., & Zhang, Y. -X. (2012). Regional total factor energy efficiency: An empirical analysis of industrial sector in China. *Applied Energy, 97*, 115–123.

Wang, K., Lu, B., & Wei, Y. -M. (2013). China's regional energy and environmental efficiency: A range-adjusted measure based analysis. *Applied Energy, 112*, 1403–1415.

WB. (2013). GDP per capita, PPP (purchasing power parity, current international $). *2004–2008.* World Bank.

XNA. (2012). *Report of the Eighteenth National Congress of the Communist Party of China*. Beijing: Xinhua News Agency.

Yu, H. (2012). The influential factors of China's regional energy intensity and its spatial linkages: 1988–2007. *Energy Policy, 45*, 583–593.

Zhang, C., & Lin, Y. (2012b). Panel estimation for urbanization, energy consumption and CO2 emissions: A regional analysis in China. *Energy Policy, 49*, 488–498.

Zhang, C., & Xu, J. (2012a). Retesting the causality between energy consumption and GDP in China: Evidence from sectoral and regional analyses using dynamic panel data. *Energy Economics, 34*, 1782–1789.

Zhang, B., Chen, Z., Xia, X., Xu X., & Chen, Y. (2013). The impact of domestic trade on China's regional energy uses: A multi-regional input–output modeling. *Energy Policy*.

Zhu, Q., Peng, X., & Wu, K. (2012). Calculation and decomposition of indirect carbon emissions from residential consumption in China based on the input–output model. *Energy Policy, 48*, 618–626.

Chapter 12
Life Cycle Analysis of Alternative Fuel Pathways

Peng Tianduo and Ou Xunmin

This chapter is focused on the review of the development and application for alternative fuel pathways Life-cycle Analysis (LCA) models. A summary on the studies of multi-pathway comprehensive analysis and individualized studies of LCA models of the world is also made with emphasis on the methodologies currently being used in life-cycle analysis modeling for transportation fuels. Three major models are examined. And the Life cycle analysis of alternative fuel pathways in China are introduced specifically using self-developed models, including life cycle analysis on energy consumption and GHG emissions of secondary energy, transport sector, six biofuel and some other pathways, which gives out some key results. In addition, research on the method of stating energy consumption of electric vehicle (EV) and plug-in hybrid electric vehicle (PHEV) is introduced. Finally, the future research on life-cycle analysis of energy consumption and GHG emissions on alternative fuel pathways is indicated.

12.1 Review of Life-Cycle Analysis Models for Transportation Energy

This part reviews the methodologies currently being used in life-cycle analysis modeling for transportation fuels (He et al. 2013). Three major models are examined, including the Greenhouse Gases, Regulated Emissions, and Energy Use in

P. Tianduo · O. Xunmin (✉)
Institute of Energy, Environment and Economy (3E), Tsinghua University, Beijing 100084, People's Republic of China
e-mail: ouxm@mail.tsinghua.edu.cn

P. Tianduo · O. Xunmin
China Automotive Energy Research Center (CAERC), Tsinghua University, Beijing 100084, People's Republic of China

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_12

Transportation model and the Conservation of Clean Air and Water in Europe model. Our objective was to evaluate and discuss some central features of these models, including: (1) their advantages and disadvantages; (2) the level of detail; (3) the assumptions, definitions, and system boundary conditions; and (4) the benefits and issues surrounding "grey" energy (the energy used to manufacture the technology and associated infrastructure).

12.1.1 Introduction

12.1.1.1 Life-Cycle Analysis for Energy Pathways and Systems

It is important to identify from a life-cycle perspective the principal sectors in China where overall energy efficiency can be improved and total greenhouse gas emissions (GHG) emissions reduced. Numerous researchers have performed life-cycle analysis (LCA) on energy pathways and systems using process-based methods that feature simplified system boundaries, although upstream system boundaries have been less than complete.

For certain processes of an energy product or system, the output can be used either directly or indirectly in the analyses. However, the result is that it is neither accurate nor even feasible to simply sum the energy requirements of each process to attain the final life-cycle energy consumption. Nevertheless, some iterative methods can be used to tackle this problem. A good example is the Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) model, which was developed at the U.S. Argonne National Laboratory over 10 years ago. The model has been improved continuously since 1999 to include more than 100 transportation fuel production pathways from various energy feedstock. The GREET model consists of about 30 Microsoft® Excel spreadsheets, where both the input and output data are interlinked through iterative methods that provide whole-life and full-cycle perspectives.

The GREET and similar LCA models, including the Life-Cycle Emission Model (LEM), have been widely used by institutions and researchers since the 1990s to analyze a variety of transportation fuels for different regions, particularly in the U.S. and Europe.

Table 12.1 Most popular LCA study or tools related to transport energy

Research	WTT model	Data
GREET 2009	GREET 1.8c.0	PSAT simulation the situation in the US
JRC/Concawe/EUCAR 2008	JRC/LBST database	Europe EUCAR data
UCD/LEM 1997-2006	LEM	Inputs
GHGenius	GHGenius 3.1.5	Various

Currently LCA models have had a primary focus on deriving net energy efficiency and GHG emission for liquid fuels to transport especially based on North America and Europe Union district. The models or studies such as GREET and EUCAR/CONCAWE are the most important ones (See Table 12.1).

12.1.1.2 LCA Studies for Energy Pathways and Model Development in China

Two types of limited LCA studies specific to fuel pathways in China have been performed: (1) comparisons between different fuel pathways (primarily for vehicles) using GREET as the basic model, with some parameters adapted to China; and (2) studies of specific pathways using self-developed models.

LCA modeling for different pathways in China using a unified but comprehensive and systematic computing platform is problematic because: (1) relatively limited data are available, (2) cooperation between related research institutes is weak, and (3) funding for long-term research has been insufficient. In summary, current energy-use research using LCA must make trade-offs between model scale (pathways covered, parameters set, and stages distinguished) and model results (degree of precision and error analysis).

In China, many researchers are dong LCA works based on studies or models listed above. And China Automotive Energy Research Center (CAERC) group in Tsinghua University also has developed Tsinghua-LCAM (TLCAM) which is an integrated computerized model for China's automotive energy supply and demand balance calculation and analysis which is based on China's national conditions and integrates GREET model.

12.1.1.3 The Purpose of This Research

The purpose of this chapter is to review the methodologies used in life-cycle analysis modeling for transportation fuels. The GREET and Conservation of Clean Air and Water in Europe (CONCAWE) models are among the models examined. Central features of the models are evaluated, including: (1) their advantages and disadvantages; (2) the level of detail; (3) the assumptions, definitions, and system boundary conditions; and (4) the benefits and issues surrounding "grey" energy.

12.1.2 Review of Key LCA Models

12.1.2.1 GREET Model

Most of this section is referred from the website of GREET: http://greet.es.anl.gov/main. To fully evaluate energy and emission impacts of advanced vehicle

technologies and new transportation fuels, the fuel cycle from WTW (wells-to-wheels) and the vehicle cycle through material recovery and vehicle disposal need to be considered. Sponsored by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE), Argonne developed this full life-cycle model called GREET. It allows researchers and analysts to evaluate various vehicle and fuel combinations on a full fuel-cycle/vehicle-cycle basis. The first version of GREET was released in 1996. Since then, Argonne has continued to upgrade, update and expand the model. The most recent GREET versions are GREET 1.8c.0 version for fuel-cycle analysis and GREET 2.7 version for vehicle-cycle analysis. GREET was developed as a multidimensional spread sheet model in Microsoft Excel. This public domain model is available free of charge for anyone to use. The current version of GREET includes more than 100 fuel pathways including petroleum fuels, natural gas fuels, biofuels, hydrogen and electricity produced from various energy feedstock sources. GREET outlines LCA modelling from U.S. perspectives.

What's more, by using the iteration calculation method, the upstream energy and material consumption and emissions are also included into fully reveal the situation of life cycle covering exploitation of energy resources, energy resources transportation, energy products processing, and transportation of energy products to end-use energy products. It is worth mentioning yet, GREET model includes upstream energy consumption and emissions related to energy and chemical products, which makes the LCA results more completeness. According to the methodology of the GREET model, for those energy/fuel consumed, they are not only as input for a certain activity, but also can be seen as the output of the upstage. What's more, in the emission calculations, GREET model not only consider the process of combustion emissions, but also considered the other non-combustion emissions such as loss by evaporation and chemical reactions.

12.1.2.2 CONCAWE Study

EUCAR (the European Council for Automotive R&D), CONCAWE and JRC (the Joint Research Centre of the EU Commission) have updated their joint evaluation of the Well-to-Wheels energy use and greenhouse gas (GHG) emissions for a wide range of potential future fuel and power train options, first published in December 2003. This study makes some interesting notes: The study is NOT a Life Cycle Analysis. It does not consider the energy or the emissions involved in building the facilities and the vehicles, or the end of life aspects. It concentrates on fuel production and vehicle use, which are the major contributors to lifetime energy use and GHG emissions. Regulated pollutants have only been considered in so far as all plants and vehicles considered are deemed to meet all current and already agreed future regulations.

12.1.2.3 LEM Model

First published in 1993 by Argonne National Laboratory, researcher Mark Delucchi's LEM has been adopted by government, industry and scientists as one of the most comprehensive tools for calculating energy use, air pollutants and CO_2-equivalent greenhouse-gas emissions from a variety of transportation and energy lifecycles. It models a wide range of passenger and freight transport modes and fuels—and now includes an advanced treatment of biofuels.

The task of developing and evaluating strategies to reduce emissions of urban air pollutants and greenhouse gases is complicated. There are many ways to produce and use energy, many sources of emissions in an energy lifecycle, and several kinds of pollutants (or greenhouse gases) emitted at each source. An evaluation of strategies to reduce emissions of greenhouse gases must be broad, detailed, and systematic. It must encompass the full "lifecycle" of a particular technology or policy, and include all of the relevant pollutants and their effects. Towards this end, UC Davis in U.S. has developed a detailed, comprehensive model of lifecycle emissions of urban air pollutants and greenhouse gases from the use of variety of transportation modes since 1990s.

The LEM estimates energy use, criteria pollutant emissions, and CO_2-equivalent greenhouse-gas emissions from a variety of transportation and energy lifecycles. It includes a wide range of modes of passenger and freight transport, electricity generation, heating, and more. For transport modes, it represents the lifecycle of fuels, vehicles, materials, and infrastructure. It calculates energy use and all regulated air pollutants plus so-called greenhouse gases. It includes input data for up to 30 countries, for the years 1970 to 2050, and is fully specified for the U.S.

For motor vehicles, the LEM calculates lifecycle emissions for a variety of combinations of end-use fuel (e.g., methanol), fuel feedstocks (e.g., coal), and vehicle types (e.g., fuel-cell vehicle). The LEM estimates the use of energy, and emissions of greenhouse gases and urban air pollutants, for the complete lifecycle of fuels, materials, vehicles, and infrastructure for the transportation modes.

12.1.2.4 GHGenius Model

The GHGenius model is based on the 1998 version of Dr. Mark Delucchi's LEM. GHGenius is now a spreadsheet Canadian model that calculates the amount of greenhouse gases generated from the time a fuel is extracted or grown to the time that it is converted in a motive energy vehicle to produce power. Whether the fuel is burned in an internal combustion engine or transformed in a fuel cell, GHGenius identifies the amount of greenhouse gases generated by a wide variety of fuels and technologies, the amount of energy used and provided, and the cost effectiveness of the entire life cycle.

In the current version of the model, there are 45 worksheets which have more than 150,000 cells containing data or results and it contains a multitude of pathways over 140 vehicles, fuel, and feedstock combinations available. With the transition to the Excel platform, GHGenius user-ship has increased to over 600 registered users. A large number of people are familiar with basic/intermediate Excel use and programming.

12.1.3 Key Issues for LCA Models

The word "model" suggests a form of mathematical representation that mimics reality through a type of artificial intelligence. LCA models aim at representing reality, but these models are relatively straightforward, mostly involve linear processes, and are very dependent on assumptions. Two particular issues with these models are the system boundary definitions and the quality of underlying data. Reliable LCA modeling is important because the results are often used to support regulatory schemes and policy objectives, including GHG reductions.

12.1.3.1 System Boundary Issues

System boundaries: The module boundaries of the current LCA models cover cradle-to-gate, or at least gate-to-gate, scenarios. For example, a module could be defined to include only the extraction and crushing of limestone, with boundaries that stop at the quarry or at the crushing facility gate. Any more elaborate processes that require limestone, like cement manufacturing, would then call upon that module and incorporate the life-cycle inventory data for this common unit process.

The boundaries may include activities such as the following:

- Acquisition, beneficiation, storage, and transfer of raw materials, including construction and earth moving, which must be done to gain access to a raw material, and nonpoint emissions from these operations.
- Acquisition, storage, and transfer of energy, which must be calculated from a set of standard processes unless specific data are available for a unit process.
- Processing of raw materials into primary products such as steel and rolls of paperboard.
- Transformation of primary products into secondary products, including steel joists and corrugated boxes.
- Disposal, incineration, recovery of waste materials, recycling, and other end-of-life unit processes.
- Transportation of materials, fuels, and products at all stages.
- Allocated energy requirements of and waste accumulation from pollution control processes that are not integral parts of the industrial processes under study (for example, a central wastewater treatment plant).

- The boundaries typically do not include the following activities:
- Construction of plants, vehicles, or other machinery used for any phase of production.
- Maintenance and administration of plants or equipment.
- Transport of workers and related infrastructure costs.

Boundary Issues: Direct Effects: Direct effects are those readily identifiable as part of a process or production chain. They are mostly under the control of the chain elements, but problems occur when accounting for co-products from system expansion (i.e. substitution). It is obvious that different co-product accounting methods can have significant effects, and the magnitude can be different for different chains. For example, results might be influenced by questions such as: (1) how far should one drill down; (2) should agricultural inputs, fertilizer manufacture, or even sulfuric acid manufacture be taken into account; and (3) what time scale should be considered for crop rotation?

Boundary Issues: Indirect Effects: Indirect effects are more speculative, subject to assumptions and scenarios, and mostly outside the direct control of the chain elements. An example is land use impact.

12.1.3.2 Data Issues

The data selected are representative of real-world diversity. For example, for the biofuel pathway, the data reflect cultivation practices as well as soil and climate characteristics. In this case, specific data are those on yields and processing schemes, and the questions to be asked are: What geographic envelope is being represented? Do most agricultural parameters have a strong local component? How many different pathways are represented in the process? Are reliable data available?

12.1.3.3 LCA as a Regulatory Tool

LCA was developed as a broad analysis tool to improve the understanding of production/supply chains and to guide strategic decisions, and recent legislative initiatives have also attempted to use LCA as a regulatory compliance tool. This has required a major rethinking of all aspects of the LCA process, including:

- Accuracy versus simplicity
- Transparency
- Verification and auditing of data
- Burden on chain elements
- Methodological choices
- Use of default values

Table 12.2 Comparisons of LCA models

Item	GREET	CONCAWE	LEM	GHGenius	Tsinghua-LCAM
Region based	US uniquely	EU uniquely	US majorly	Canada uniquely	China uniquely
Most recent version	2012	2006	2006	2010	2011
Tool used					
Tool copyright holder	US Argon National lab	LBST company	UC DAVIS	Canada government	Tsinghua University
Software open?	Yes	No	No	Yes	No
Easy to master?	Yes	n.a.	n.a.	No	Yes
Popularity	High	Low	Low	Low	Low
Grey energy[a] included or not					
Infrastructure construction	No	No	No	No	No
Machine manufacture in fuel cycle	No	No	No	No	No
Maintenance and administration of plants or equipment?	No	No	No	No	No
Human related?	No	No	No	No	No
Other include item	Direct energy use during upstream stage; Energy use for fertilizer/insect etc. for crop				
Other issues					
Height or complexity?	Over 100 fuel pathways with specific parameters (US)	Nearly 100 fuel pathways with specific parameters (EU)	Fuel/electricity/heat pathways are included	About 80 fuel pathways with specific parameters (Canada)	Just over 10 pathways to formulate 9 type of secondary energy results for China
Overall remarks	Very professional for transportation energy pathways; tools are open and can easily use further	Tools are not free for public; some structures can be referred	Tools are not free for public; some logics can be referred	Tools are open; some structures and logics can be referred	It is easy to expend; with the first step result. the further LC analysis can be done for specific pathway

[a]Grey energy refers to the energy used to manufacture the energy technology and associated infrastructure

12.1.4 Synthesis Assessment

12.1.4.1 Multi-dimension Comparisons

Table 12.2 provides comparisons of LCA models, based on some principles suggested by energy-related and transportation-related models.

12.1.4.2 Concluding Remarks

All of the models have distinct advantages and disadvantages. The level of detail for the pathways analyzed is high for each model, although assumptions, definitions, and system boundary conditions differ among the models to some degree. None of the models take into consideration grey energy for manufacturing the energy technology and associated infrastructure.

12.2 Life Cycle Analysis of Alternative Fuel Pathways in China

Our world is facing an increasing challenge of global climate change associated mainly with the greenhouse gases (GHG) emissions from energy use. A strategic transition to low-carbon energy use is considered to be essential for China, the largest energy consumer and the largest CO_2 emitter, to tackle climate change along with other countries. It is therefore an important and fundamental work to investigate the life cycle analysis of alternative fuel pathways. In China, researchers of Tsinghua University are doing LCA studies of alternative fuel pathways using self-developed models.

12.2.1 Life Cycle Analysis on Energy Consumption and GHG Emissions of Secondary Energy in China

Life-cycle fossil primary energy consumption (FPEC) and greenhouse gas (GHG) emission intensity of nine types of dominant secondary energy (SE) pathways for China in 2010 are calculated with iterative methods, using the TLCAM (Tsinghua Life-cycle Analysis Model) (Li et al. 2013). Three major types of GHG (CO_2, CH_4 and N_2O) are considered for GHG emission intensity, and non-combustion CH_4 leakage during the feedstock production sub-stage is included.

The objectives of this study are to set up a computer model for calculating a life-cycle FPEC and GHG emission intensity inventory of nine types of dominant

SE in the real situation of China, and to pave the way for two types of applications for that situation. Specifically, this encompasses: (1) Providing the multipliers for specific LCA research by multiplying their corresponding process fuel uses to obtain final results; and (2) establishing a platform for sector lifecycle energy use and GHG emission analysis. In other words, life-cycle FPEC and GHG emission intensities (in 2010) are evaluated for nine dominant SE types in China, including raw coal (RC), raw NG (RN), raw oil (RO), final coal (FC), final NG (FN), diesel, gasoline, residual oil and electricity.

It is demonstrated that the distinguishing features of the TLCAM (Tsinghua Life-cycle Analysis Model) and overcome weaknesses described as follows:

(1) Most types of dominant SE in China are covered, including both types used directly after extraction and/or transportation only; these types require further processing;
(2) Interlinks among different energy pathways and consequent impacts for their final life-cycle results are captured using a computerized iterative method based on the TLCAM used in Excel;
(3) Upstream or indirect energy use and GHG emission impacts are analyzed and shown in detail, to describe embodied energy and emissions for SE pathways in China.

It is found that the following. (1) Life-cycle FPEC intensities in units of per MJ SE are obtained and used, in order of magnitude, for: raw coal (recovered only); raw natural gas (NG, recovered and processed only); raw oil (recovered and processed); final coal (finally transported to end-user); final NG (finally transported to end-user); diesel; gasoline; residual oil and electricity. (2) Although their upstream GHG emission intensities are small, their life-cycle intensities are 103.5, 68.3, 81.6, 99.3, 70.0, 101.6, 91.7, 93.5 and 226.4 g CO_2 e/MJ SE, respectively, when direct GHG emissions are included. (3) Life-cycle intensities of both FPEC and GHG emissions for SE in China are higher than those in some other countries, because of the relatively low overall efficiency and high percentage of coal in the national energy mix.

Results of this study (life-cycle intensity inventory of dominant SE pathways) provide comprehensive source data for both specific micro-level energy pathway LCA and macro-level sectoral life-cycle (LC) calculation. The Chinese-specific structure and data, system boundary consistency, and calculation completeness in this study will benefit accurate and feasible micro- and macro-level LCA study in China, through application of the life-cycle intensities inventory derived here: (1) For future micro-level life-cycle energy use and GHG emissions across a wide range of energy-pathway LCA in China, there will be obvious improvement in accuracies of specific results and their comparisons; and (2) For future macro-level sector LC energy use and GHG emission calculation in China, results calculated based on a unified platform will be more meaningful and beneficial for discovering key sectors or stages in which overall energy efficiency can be improved and total GHG emissions reduced, from a life-cycle perspective.

12.2.2 Life-Cycle Analysis on Energy Consumption and GHG Emission Intensities of Transport Sector in China

12.2.2.1 Life-Cycle Analysis on Energy Consumption and GHG Emission Intensities of Alternative Vehicle Fuels

About 10 % of China's total end-use energy is estimated to be consumed by the transport sector, which accounts for about 40 %, 95 %, 60 % and 80 % of petroleum, gasoline, diesel and kerosene consumption in China, respectively. Road vehicle energy consumption is not only the most significant component of transportation energy use but also one of the most important drivers of greenhouse gases (GHG) emissions. China is considering global climate change and national energy security issues and prompting the development and utilization of low-carbon the alternative vehicle fuel (AVF). It is very import to analyze fossil energy consumption (FEC) and GHG emission intensity of China's major AVF pathways by LCA based on one uniform platform.

The Tsinghua-CA3EM is employed to calculate the life-cycle FEC and GHG emissions of AVF pathways. This model is developed by China Automotive Energy Research Center, Tsinghua University (CAERC) based on the famous GREET model (Ou et al. 2012).

Fossil energy consumption (FEC) and greenhouse gas (GHG) emission intensities of major alternative vehicle fuels (AVFs) in China are calculated and compared with conventional fuels by means of full life-cycle analysis. Currently most of the AVFs have not relatively obvious GHG emission reduction when compared to the gasoline pathway: (1) coal-based AVF has higher intensities in terms of both the FEC and GHG emissions; (2) electricity from the average Chinese grid has the GHG emission intensity similar to that of gasoline pathway although relatively lower FEC intensity; and (3) first generation technology bio-fuel has relatively lower GHG emission intensity and substantially lower FEC intensity. It is forecasted that by 2020 when still comparing to the gasoline pathway: (1) coal-based AVF will still have FEC and GHG emission intensities that are 1.5–1.8 and 1.8–2.5 time those of gasoline pathway, and the application of carbon capture and storage technology can reduce the GHG emission intensity of coal-based AVF; (2) electricity will have significantly lower GHG intensity; and (3) second generation technology bio-fuel will have near zero FEC and GHG intensities.

12.2.2.2 Using Coal for Transportation in China: Life Cycle GHG of Coal-Based Fuel and Electric Vehicle, and Policy Implications

Coal is expected to play a crucial role as an abundant energy source in China for the long term. However, both clean and advanced coal technologies are needed to

utilize coal in an environmentally responsible manner while improving utilization efficiency.

This part compares the GHG emissions of coal-to-liquid (CTL) fuels to the GHG emissions of electric vehicles (EVs) powered with coal-to-electricity in China (Ou et al. 2010). A life cycle model is used to account for full fuel cycle and use-phase emissions, as well as vehicle cycle and battery manufacturing emissions. It is found that the reduction of life cycle GHG emissions of EVs charged by electricity generated from coal, without utilizing carbon dioxide capture and storage (CCS) technology can be 3–36 % when compared to petroleum-based gasoline car. The large range in emissions reduction potential is driven by the many different power generation technologies that are and could in the future be used to generate electricity in China. When CCS is employed in power plants, the GHG emission reductions increase to 60–70 % compared to petroleum-based gasoline car. However, the use of coal to produce liquid transportation fuels (CTL fuels) will likely lead to significantly increased life cycle GHG emissions, potentially 30–140 % higher than petroleum-based gasoline. When CCS is utilized in the CTL plant, the CTL fueled vehicles emit roughly equal GHG emissions to petroleum-based gasoline vehicles from the life cycle perspective. The authors conclude that policies are therefore needed in China in order to accelerate battery technology and infrastructural improvements for EV charging, increased energy efficiency management, and deployment of low-carbon technologies such as CCS.

12.2.2.3 Life Cycle GHG of NG-Based Fuel and Electric Vehicle in China

The percentage of coal in China's total primary energy consumption was 68.8 % and coal is expected to play a crucial role as an abundant energy source in China for the long term. Though the share of NG in China's total energy consumption was very low in the past years (2.6 % in 2001 and 4.6 % in 2011), the rates of increase were significant: the average annual rate of increase was 13 % and 16 %, respectively, for NG production and consumption in China in the period from 2001 to 2011. China aims to increase the share of NG in China's total energy consumption to 7–8 % by 2015, according to the specific plan for NG development in China (2011–2015). It is estimated that share of NG in China's total energy consumption can reach to 13 % and 15 % by 2030 and 2050, respectively, by expanding the domestic production and international imports at the same time.

This part compares the greenhouse gas (GHG) emissions of natural gas (NG)-based fuels to the GHG emissions of electric vehicles (EVs) powered with NG-to-electricity in China (Ou et al. 2013a). A life-cycle model is used to account for full fuel cycle and use-phase emissions, as well as vehicle cycle and battery manufacturing.

(1) EVs look promising as a pathway for reducing GHG emissions when NG is used to power/fuel transportation. Even if NG electricity without CCS is used, EVs reduce life cycle GHG emissions by 36–47 % compared to petroleum-based gasoline in conventional ICE vehicles.
(2) CNG and LNG pathways both can decrease life cycle GHG emissions by 10 % compared to petroleum-based gasoline. GTL fuels, on the other hand, will likely increase the GHG emissions associated with transportation fuels when CCS technology is not employed. It actually emit life cycle GHG emissions roughly equal to diesel fueled HEVs with CCS applied in GTL plant. To simultaneously assist with the goals of enhancing oil security while reducing GHG emissions in the passenger transportation sector, EVs (including EVs and PHEVs) are better than GTL fuels.
(3) Almost all NG-based pathways are being developed in China and they are based mainly on the implementation of a portfolio energy strategy that promotes industrial development. However, the GHG emissions will continue to increase. If the goal is to substitute petroleum with NG, then the energy savings dilemma (in particular petroleum saving) and associated increases in GHG emissions, must be carefully comprehended.

12.2.2.4 Research on the Method of Stating Energy Consumption and GHG Emissions of Electric Vehicle (EV) and Plug-in Hybrid Electric Vehicle (PHEV)

Tsinghua University has developed Tsinghua-LCAM (TLCAM) which is an integrated computerized model for China's automotive energy supply and demand balance calculation and analysis which is based on China's national conditions and integrates GREET model. Life-cycle analysis on energy consumption and GHG emission intensities of EV and PHEV can be conducted based on the model.

The emissions from EVs depend on their own energy consumption and on the CO_2 intensity of the power generation mix from which the EV's energy should obtained. The energy consumption is the amount of energy used per unit distance traveled. The CO_2 intensity of a power generation mix is the average amount of CO_2 emitted per unit of electrical energy generated by all of the power production processes in a mix weighted by the amount of power obtained from each of those processes. Thus the upstream stage of power supply should be covered in the assessment of EV energy consumption.

Methods overview:

(1) The methods are based on EXCEL tools to get life cycle analysis results. Electricity chains and vehicle running are considered in the calculation, that is, upstream and operation stages are both covered in life cycle consumption and emissions.

(2) Three kinds of fossil fuels including Coal, Oil and Natural gas are used as feedstock in power generation. Energy consumption and emissions include the upstream stages, such as feedstock exploration, transportation, fuel production, and transportation, in addition to the energy consumption emissions occurring in the fuel utilization; but the facility construction and vehicle manufacturing stages are excluded for their little effect on the life cycle energy consumption and emission.
(3) Non-fossil fuels includes Hydro, Nuclear, Solar, Wind and other types. The energy consumption and emission during facility manufacturing and factory construction stages are allocated to the total power supplying during the whole life time of those power stations for they account for a very large proportion. Life cycle analysis of GHG emissions situation of different power.

Calculation methodology:

(1) Life cycle energy consumption per unit distance driven by EV and PHEV can be calculated based on some data collected for the specific region.

$$EN_{EV} = \left[\frac{E_{LC,Coal}}{\eta_{Coal-to-ele}} * SH_{Coal} + \frac{E_{LC,NG}}{\eta_{NG-to-ele}} * SH_{NG} + \frac{E_{LC,Oil}}{\eta_{Oil-to-ele}} * SH_{Oil} + \sum_{k=1}^{i} E_{LC,k} * SH_k\right] * \frac{1}{1-\eta_{Loss}} * \frac{E_{Ele,EV}}{\eta_{Charge}} * \frac{3.6}{100}$$

$$EN_{PHEV} = \left[\frac{E_{LC,Coal}}{\eta_{Coal-to-ele}} * SH_{Coal} + \frac{E_{LC,NG}}{\eta_{NG-to-ele}} * SH_{NG} + \frac{E_{LC,Oil}}{\eta_{Oil-to-ele}} * SH_{Oil} + \sum_{k=1}^{i} E_{LC,k} * SH_k\right] * \frac{1}{1-\eta_{Loss}} * \frac{E_{Ele,PHEV}}{\eta_{Charge}} * \frac{3.6}{100} * SH_{Ele} + (1-SH_{Ele}) * E_{LC,Gasoline} * V_{Gasoline} * Q_{Gasoline} * \frac{1}{100}$$

where:

- $E_{LC,Coal}$: Life cycle energy consumption for coal production and distribution (MJ/MJ fuel obtained)
- $\eta_{Coal-to-ele}$: Coal power electricity generation efficiency (%)
- SH_{Coal}: The share of coal power in the total electricity supplying of regional electrical grids (%)
- k: The type of non-fossil fuel power from 1 to j mean: Hydro, Nuclear, …
- $E_{LC,K}$: Life cycle energy consumption for electricity generation and supply of type k (MJ/MJ power supplying)
- SH_k: The share of non-fossil fuel type k in the total electricity supplying (%)
- η_{Loss}: Electricity transmission loss rate (%)
- η_{Charge}: Charging Efficiency (%)
- $E_{Ele,EV}$: Direct energy consumption of EV (kWh/100 km)
- $E_{Ele,PHEV}$: Direct energy consumption of PHEV (kWh/100 km)
- SH_{Ele}: The range share by electricity (%)

- $Q_{Gasoline}$: Calorific value of gasoline (32 MJ/L)
- $E_{LC,\,Gasoline}$: Life cycle energy consumption for gasoline production and utilization (MJ/MJ)
- $V_{Gasoline}$: Energy consumption of PHEV driven by gasoline in running stage (Liter/100 km)

(2) Life cycle GHG emission per unit distance driven by EV and PHEV

$$EM_{EV} = \left[\frac{EM_{LC,\,Coal}}{\eta_{Coal-to-ele}} * SH_{Coal} + \frac{EM_{LC,\,NG}}{\eta_{NG-to-ele}} * SH_{NG} + \frac{EM_{LC,\,Oil}}{\eta_{Oil-to-ele}} * SH_{Oil} + \sum_{k=1}^{i} EM_{LC,k} * SH_k\right] * \frac{1}{1-\eta_{Loss}} * \frac{E_{Ele,\,EV}}{\eta_{Ch\,arg\,e}} * \frac{3.6}{100}$$

$$EM_{PHEV} = \left[\frac{EM_{LC,\,Coal}}{\eta_{Coal-to-ele}} * SH_{Coal} + \frac{EM_{LC,\,NG}}{\eta_{NG-to-ele}} * SH_{NG} + \frac{EM_{LC,\,Oil}}{\eta_{Oil-to-ele}} * SH_{Oil} + \sum_{k=1}^{i} EM_{LC,k} * SH_k\right] * \frac{1}{1-\eta_{Loss}} * \frac{E_{Ele,\,PHEV}}{\eta_{Ch\,arg\,e}} * \frac{3.6}{100} * SH_{Ele} + (1-SH_{Ele}) * EM_{LC,\,Gasoline} * V_{Gasoline} * Q_{Gasoline} * \frac{1}{100}$$

where:

- $EM_{LC,\,Coal}$: Life cycle GHG emission for coal production and distribution (g CO_2, e/MJ)
- $\eta_{Coal-to-ele}$: Coal power electricity generation efficiency (%)
- SH_{Coal}: The share of coal power in the total electricity supplying of regional electrical grids (%)
- k: The type of non-fossil fuel power from 1 to j mean: Hydro, Nuclear, …
- $EM_{LC,\,K}$: Life cycle GHG emission for electricity generation and supply of type k (g CO_2, e/MJ power supplying)
- SH_k: The share of non-fossil fuel type k in the total electricity supplying (%)
- η_{Loss}: Electricity transmission loss rate (%)
- η_{Charge}: Charging Efficiency (%)
- $E_{Ele,\,EV}$: Direct energy consumption of EV (kWh/100 km)
- $E_{Ele,\,PHEV}$: Direct energy consumption of PHEV (kWh/100 km)
- SH_{Ele}: The range share by electricity (%)
- $Q_{Gasoline}$: Calorific value of gasoline (32 MJ/L)
- $EM_{LC,\,Gasoline}$: Life cycle GHG emission for gasoline production and utilization (67.91 g CO_2, e/MJ)
- $V_{Gasoline}$: Energy consumption of PHEV driven by gasoline in running stage (Liter/100 km)

The data mentioned in the calculation formula can be collected with clear sources such as statistical book or formal report. And the results can be calculated based on the model and presented as follows:

(1) Labelling together

- ** kWh/100 km
- ** Liter (gasoline equivalent)/100 km

(2) Comparing energy consumption by primary energy (**MJ/km)
(3) Comparing GHG emissions to conventional gasoline vehicle

- Total
- By stages

12.2.3 Energy Consumption and GHG Emissions of Six Biofuel Pathways by LCA in (the) People's Republic of China

In this part, the Well-to-Wheels (WTW) analysis module of the Tsinghua-CA3EM (China Automotive Energy, Environment and Economy Model) model is used, which is an integrated computerized model module for China's automotive energy supply and demand balance calculation and analysis (Ou et al. 2009). The model is based on China's national conditions with the integration of the widely known transportation energy micro-level computing GREET model. Part of the GREET model structure has been adjusted to Chinese specific situations, such as the dominance of coal utilization. Therefore, a majority of the parameters have been modified with local Chinese data.

This study presents life-cycle-analysis (LCA) energy consumption (EC) and greenhouse gas (GHG) emissions of China's current six biofuel pathways, which are: corn-derived ethanol (CE); cassava-derived ethanol (KE); sweet sorghum-derived ethanol (SE); soybean-derived bio-diesel (SB); jatropha fruit-derived bio-diesel (JB); and used cooking oil (UCO)-derived bio-diesel (UB). The tool utilized here is the WTW (Well-to-Wheels) module of Tsinghua-CA3EM model covering the entire lifecycle including: raw materials cultivation (or feedstock collection); fuel production; transportation and distribution; and application in automobile engines, compared with Conventional Petroleum-based gasoline and diesel Pathways (CPP).

The results indicate: (1) the fossil energy inputs are about 1.0–1.5 times the energy contained in the fuel for the CE, SE and SB pathways, but 0.5–0.9 times for the KE, UB and JB pathways; (2) compared with CPP, the JB, KE and UB pathways can reduce both fossil fuel consumption and GHG emissions; the CE and SB pathways can only reduce fossil fuel consumption, but increase GHG emission; the SE pathway increases not only fossil fuel consumption but also GHG emission; and (3) the main factors inducing high EC and GHG emission levels include: high EC levels during the fuel production stage and high fertilizer application rates during the planting of raw feedstocks. Conclusions are that of the aforementioned biofuel

pathways in (the) People's Republic of China: (1) only the JB, KE and UB pathways have energy-saving merits as indicated by the LCA energy inputs and outputs; (2) compared with CPP, all but the SE pathway reduces fossil fuel consumption. However, the SB and CE pathway increase GHG emission; (3) all six displace petroleum by utilizing more coal; and (4) feedstock productivity levels must be increased, and there must be a reduction in fertilizer utilization and EC consumption during the cultivation and transportation stages in order to achieve the goals of energy balance and GHG emission reduction.

12.2.4 Energy Consumption and GHG Emissions of Other Pathways by LCA in China

Some other studies have been conducted except for the above. For Example, the life-cycle analysis (LCA) of energy use and greenhouse gas emissions from the LanzaTech process has been developed for a Chinese setting using the original Tsinghua China Automotive LCA model along with a customized module developed principally for the process (Ou et al. 2013b). The LCA results demonstrate that LanzaTech gas-to-liquid (GTL) processing in China's steel manufacturing is favorable in terms of life-cycle fossil energy and can reduce greenhouse gas emissions by approximately 50 % compared with the conventional petroleum gasoline. The LanzaTech process, therefore, shows advantages in both energy-savings and a reduction in greenhouse gas emissions when compared with most bio-ethanol production pathways in China.

And the study on Life-Cycle Energy Use and Greenhouse Gas Emissions Analysis for Bio-Liquid Jet Fuel from Open Pond-Based Micro-Algae under China Conditions is also conducted using the Tsinghua University LCA Model (TLCAM) (Ou et al. 2013). Attention was paid to energy recovery through biogas production and cogeneration of heat and power (CHP) from the residual biomass after oil extraction, including fugitive methane (CH_4) emissions during the production of biogas and nitrous oxide (N_2O) emissions during the use of digestate (solid residue from anaerobic digestion) as agricultural fertilizer. Analyses were performed based on examination of process parameters, mass balance conditions, material requirement, energy consumptions and the realities of energy supply and transport in China (i.e., electricity generation and heat supply primarily based on coal, multiple transport modes). LCA result of the BJF pathway showed that, compared with the traditional petrochemical pathway, this new pathway will increase the overall fossil energy use and carbon emission by 39 % and 70 %, respectively, while decrease petroleum consumption by about 84 %, based on the same units of energy service. Moreover, the energy conservation and emission reduction benefit of this new pathway may be accomplished by two sets of approaches: wider adoption of low-carbon process fuels and optimization of algae cultivation and harvest, and oil extraction processes.

12.3 Summary and Prospect of This Chapter

The chapter gives a brief introduction about the development and application for alternative fuel pathways Life-cycle Analysis (LCA) models, especially the methodologies currently being used in life-cycle analysis modeling for transportation fuels, including the Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation model and the Conservation of Clean Air and Water in Europe model. Tsinghua University has developed the GREET model for the comparative study on different pathways. Life-cycle analysis on energy consumption and GHG emission intensities of transport sector, secondary energy, NG-Based Fuel and Electric Vehicle, and biofuel pathways in China are introduced respectively. It is an important and significant work to do life cycle analysis of alternative fuel pathways.

In the future, different LCA approaches will be harmonized for key sectors, including building, transport, and agricultural, industrial and residential sectors. Analysis and assessment of energy supply and utilization need to expand into normal pollutants discharge and economy analysis and can be done via real life-cycle analysis approaches that consider both full life stages and indirect consumption and emission implications owing to direct consumption and emission in the cycle methods.

References

He, X., Ou, X., Zhang, X., Zhang, Q., & Xing, Y. (2013). Review of life-cycle analysis models for transportation energy. *International Journal on Advances in Information Sciences and Service Sciences, 5*(9), 375–382.

Li, X., Ou, X., Zhang, X., Zhang, Q. & Zhang, X. (2013). Life-cycle fossil energy consumption and greenhouse gas emission intensity of dominant secondary energy pathways of China in 2010. *Energy, 5015–23*.

Ou, X., Yan, X., Zhang, X., & Liu, Z. (2012). Life-cycle analysis on energy consumption and GHG emission intensities of alternative vehicle fuels in China. *Applied Energy, 90*(1), 218–224.

Ou, X., Yan, X., & Zhang, X. (2010). Using coal for transportation in China: Life cycle GHG of coal-based fuel and electric vehicle, and policy implications. *International Journal of Greenhouse Gas Control, 4*(5), 878–887.

Ou, X., Zhang, X., Zhang, X. & Zhang, Q. (2013a). Life cycle GHG of NG-based fuel and electric vehicle in China. *Energies, 6(*5), 2644–2662.

Ou, X., Zhang, X., Chang, S. & Guo, Q. (2009). Energy consumption and GHG emissions of six biofuel pathways by LCA in (the) People's Republic of China. *Applied Energy, 86S197–S208*.

Ou, X., Zhang, X., Zhang, Q. & Zhang, X. (2013b). Life-cycle analysis of energy use and greenhouse gas emissions of gas-to-liquid fuel pathway from steel mill off-gas in China by the LanzaTech process. *Frontiers in Energy, 7*(3), 263–270.

Ou, X., Yan, X., Zhang, X., & Zhang, X. (2013c). Life-cycle energy use and greenhouse gas emissions analysis for bio-liquid jet fuel from open pond-based micro-algae under China conditions. *Energies, 6*(9), 4897–4923.

Chapter 13
Power Generation System Optimization with Emission Co-benefits Analysis: A Case Study of Shanghai

Zheng Chang, Ke-xi Pan and Han-xiong Zhu

Abstract Power generation industry will become the key air emission (such as CO_2 and $PM_{2.5}$) control sector for China in the mid to long-term. To study the least-cost and maximizing co-benefits of emission reduction, a Reginal Power-generation System Optimization (RPSO) model is built using LEAP platform in this study. Taking Shanghai as a case study, multiple scenarios are designed to simulate the optimal technology development pathways and to evaluate mitigation co-benefits under different air emission targets. The results show that, fuel consumption is reduced while the share of clean generation unit is increased in all emission control scenarios, however, system cost is raised up as well. Ultra-supercritical units take the biggest share in all scenarios, natural gas combined cycle (NGCC) units and gas distributed units develop better in $PM_{2.5}$ constrained scenarios. Hydro and nuclear units reach the maximum capacity limits in all scenarios. Wind and solar units grows fast but account for very little share. Reduction co-benefits of $PM_{2.5}$ from carbon constrained scenario is larger than that of $PM_{2.5}$ constrained scenario which indicates carbon migration policy has greater influence on $PM_{2.5}$ reductions.

13.1 Introduction

As climate change became increasingly serious, IPCC stated in the 5th Assessment Report that only in the maximum reduction effort which is RCP2.6 scenario that the rising temperature of atmosphere could be controlled under 2 °C (IPCC 2014).

Z. Chang
Fudan Development Institute, Fudan University, Shanghai 200433, China

K. Pan (✉)
School of Social Development and Public Policy, Fudan University, Shanghai 200433, China
e-mail: pankx@fudan.edu.cn

K. Pan · H. Zhu
Fudan Energy Research Center, Fudan University, Shanghai 200433, China

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_13

China has surpassed the US becoming the biggest GHG emitter in the world. Though not has taken mandatory emission reduction responsibilities so far, China is now facing growing international mitigation pressure ever since. Considering that GHG emission in China will still increase in a short-term subjected to coal dominated energy structure, it is of highly importance to make accelerating mitigation policies at once. Moreover, *2014 China Environment Status Communique* showed that atmosphere pollutions in Jing-Jin-Ji region, the Yangtze Delta region, the Pearl Delta region and other main cities in China are still serious, among which over 90 % of 161 Chinese main cities' air quality index failed to reach the national standard (MEPC 2015). With the rapid economic development and improvement of living conditions, residents' appeals for clean air had increased simultaneously. Local air pollutants (LAPs) reduction has become a substantial aspect of livelihood issues.

As one of the influential GHG, a portion of CO_2 emission has the same origin with LAPs. Primary energy (coal, oil, natural gas and biomass) is the carrier of multiple elements, e.g. carbon, nitrogen and sulfur. The energy utilization is also the process that transported these elements to air emissions. More than 95 % of carbon emission is led by fossil fuel combustion (IPCC 2006) which LAPs could be derived as well. Meanwhile, more than 90 % anthropology activities related NO_x emission comes from fuel combustion. Thus, it is believed that reducing CO_2 emission could cut down LAPs at the same time. Or in other words, there is both co-benefits effect of reducing CO_2 and LAPs. Hence, how to coordinate multiple air emission reductions in an efficient and economical way that not only fulfills the promise of climate mitigation to the world but also satisfy the needs of improvement air quality by residents, need to be considered as a whole.

Power generation system is one of the most primary sectors that energy consumed, transferred and also air emission generated. China consumed 4.24 billion metric tons of coal in 2014, of which power generation system was the biggest consumer contributing 55 % of total. Meanwhile, power generation system emitted more 40 % of LAPs. Air emissions of power generation present the characteristic of intensively emission, easily control and great reduction potential. For instance, a research of global GHG reduction potential suggested that power generation system made up 27 % of total global reduction potential in 2030 (McKinsey & Company 2013). Considering energy plays a driving role in economic growth, as China's economy would keep a relative high speed growth with urbanization progress continuing, power generation system will play a key role in reducing Chinese multiple air emissions.

Existing studies of emission reduction co-benefits has focused on developed countries (Bollen 2015; Menikpura et al. 2014; Mrkajic et al. 2015), and yet, co-benefit studies aiming at China emerged in recent years. The aims of those studies has shifted to multiple air emissions reductions and the co-benefits including GHGs (mainly CO_2) and LAPs (mainly SO_2, NO_x, CO and PM_s) (Zhang et al. 2015; He et al. 2010; Jiang et al. 2013; Chen et al. 2006). With regard to industry co-benefits assessment, several studies evaluated co-benefits of energy efficiency and air pollution abatement in China's cement industry (Zhang et al. 2015; Yang et al. 2013; Xi et al. 2013). Mao et al. (2012) compared different policy

instruments for CO_2 emission reduction and the co-benefits of local air pollution abatement in the transportation sector of China. Ma et al. (2014) assessed the emission reduction potentials and the co-benefits for advanced technologies in the iron and steel industry based on China-TIMES model. For power generation industry, Ma et al. (2013) computed the co-benefits of CO_2, LAPs and water savings of wind industry in Xin-jiang province. Xue et al. (2015) conducted a similar analysis for wind power sector by using life cycle methods.

Along with the reduction and control target allocated to industry and enterprise from national level, policy makers begin to focus on mitigation strategies and technologies for power generation system that suitable for the specific industry and region. Their concerns include reduction allocation, mitigation cost-benefits, technology selection and policy measurement choices e.g. subside and tax. Nevertheless, as many existing studied focusing on national level, there is still a gap between researchers and regional policy makers to transfer academic achievement to policy-makings.

The study aims to assess power generation technology combination pathways with emission co-benefits under various policy scenarios. The authors organized the study into six parts. Study background and review of emission co-benefits were introduced first. The methodology is then constructed before the case study of Shanghai is presented. The fourth and fifth part contains the results of six scenarios alongside a sensitivity analysis. The final part includes conclusions and policy recommendations for future studies.

13.2 Methodology

In order to assess power generation technology combination pathways and quantify the emission mitigation co-benefit effects, we construct a Regional Power Generation System Optimization (RPSO) model using Long-range Energy Alternatives Planning System (LEAP) system. This section describes the existing LEAP studies and the model structure of RPSO model we built.

13.2.1 Long-range Energy Alternatives Planning System (LEAP) System

LEAP (the Long range Energy Alternatives Planning) System is an application software for energy environment planning developed by the Stockholm Environment Institute. It is an integrated modeling tool that can be used to simulate mid to long-term energy supply and demand for air pollutions and GHGs stem from energy utilization (SEI 2015). LEAP has been widely used by researchers and organizations for national and regional energy demand forecast and scenario analysis (Suhono and Sarjiya 2015; Andrade Guerra et al. 2015; McPherson and Karney 2014). It also has

been adopted to analyze energy related GHG and LAPs emission (Hong et al. 2016). For power generation system, Perwez and Sohail (2014) assessed financial and environmental cost alternative scenarios for power sector of Pakistan. Dias et al. (2014) analyzed impact on electricity demand and emissions due to the introduction of electric cars in the San Paulo power system. And Cai et al. (2007) constructed current policy scenario and the new policy scenario to estimated CO_2 emissions reduction potential in China's power generation sector.

As an important transformation sector of energy system, power generation system connects both energy supply and demand side producing air emissions as well. RPSO model built on LEAP can better simulate optimizing energy technologies and air emissions together calculating the co-benefits by setting up policy scenarios.

13.2.2 *Structure of Regional Power Generation System Optimization (RPSO) Model*

RPSO is a multi-stage closure model which takes power generation as the research boundary. The time step of the model is 1 year and planning period covers 2015–2030 including CO_2, SO_2, NO_x, PM_{10}, $PM_{2.5}$ etc. As shown in Fig. 13.1,

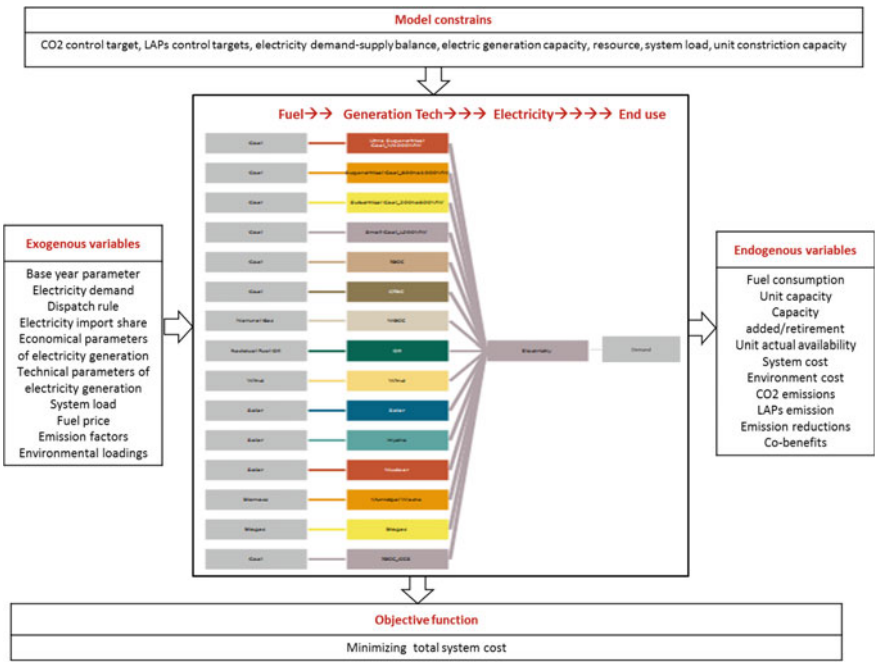

Fig. 13.1 Structure of Regional Power Generation System Optimization (RPSO) model

RPSO model simulates fuel utilization process from production to electricity generation units and ultimately to end use sector. A set of variables are defined to describe the process namely exogenous variables, model constrains, objective function and endogenous variables that calculated by model.

Electricity generation technologies that taken into account of this study include:

a. Coal-fired power generation technology including less than 300 MW unit, pulverized coal fired boiler of power plants (subcritical, supercritical, ultra-supercritical), IGCC (integrated gasification combined cycle) and CFB (combined circulation boiler).
b. Other fossil fuel combustion technology such as gas power generation and oil power generation.
c. Nuclear.
d. Renewable power generation includes biofuel, wind power, hydro and solar power generation.

13.2.3 Objective Function

The optimizing objective of RPSO is minimizing regional electric generation system total cost within the planning period, which is composed by initial capital cost (*CAP*), fixed OM cost (*FOM*), variable OM cost (*VOM*) and fuel cost (*FUFL*). VOM is the cost that all the other variable cost excludes fuel cost. *i* presents discount rate, *t* is year, *n* is electric generation method:

$$COST = min\left\{ \sum_{t} \left[\sum_{n} (CAP_{n,i} + FOM_{n,i} + VOM_{n,i} + FUEL_{n,i}) / (1+i)^{t-1} \right] \right\} \quad (13.1)$$

Initial capital cost is the input in construction period. In order to compare different technologies, initial cost here is shared to life cycle of the unit. *l* is the service year of tech *n*. *caper* is the initial capital cost per capacity. Thus, capital cost per capacity in year *t* is:

$$costcap_{n,t} = caper_{n,t} \times i \times (1+i)^{l} / \left[(1+i)^{l} - 1 \right] \quad (13.2)$$

Running cost can be divided into two parts, *FOM* and *VOM*. *QCAP* is capacity, *costfix* is *FOM* per capacity. *GELEC* represents electricity production, *costvar* is VOM per electricity production:

$$FOM_{n,t} = QCAP_{n,t} \times costfix_{n,t} \quad (13.3)$$

$$VOM_{n,t} = GELEC_{n,t} \times costvar_{n,t} \quad (13.4)$$

$capPR_{n,t}$, $fixPR_{n,t}$, and $varPR_{n,t}$ are the change rate of *caper*, *costfix* and *costvar*:

$$caper_{n,t} = caper_{n,0}(1 + capPR_{n,t})^t \tag{13.5}$$

$$costfix_{n,t} = costfix_{n,0}(1 + fixPR_{n,t})^t \tag{13.6}$$

$$costvar_{n,t} = costvar_{n,0}(1 + varPR_{n,t})^t \tag{13.7}$$

Fuel cost *FUEL* is computed according to *GELEC*, energy consumption per electricity production *efuel* and fuel price *costfuel*. *CFUEL* is the consumption of fuel *j*:

$$DFUEL_{n,t} = GELEC_{n,t} \times efuel_{n,j} \tag{13.8}$$

$$FUEL_{n,t} = CFUEL_{n,t} \times costfuel_{n,j} \tag{13.9}$$

costfuel changes year by year in *fuelPR*:

$$costfuel_{n,t} = costfuel_{n,0}(1 + fuelPR_{n,t})^t \tag{13.10}$$

13.2.4 Emission Calculations

Air emission *s* in year *t* of the reginal electric generation system is calculated as follows, in which *EF* is the emission factor of process namely emission per energy consumed.

$$EMIS_{s,t} = \sum_n CFUEL_{n,s} \times EF_{n,s} \tag{13.11}$$

13.2.5 Operational Constrains

13.2.5.1 Electricity Balance

GELEC in year *t* is the total electricity production of all types of power generation technologies. Reginal electricity consumption *CELEC* should be no more than *GELEC*.

$$CELEC_t \le GELEC_t = \sum_n GELEC_{n,t} \tag{13.12}$$

It should be noted that as imported electricity is treated as one category of unit, there was no need to consider electricity transfer across regions in this study.

α is planning reserve margin rate, *AVA* is availability of peak load, *HCAP* is maximum load of reginal electric generation system:

$$\sum_n (QCAP_{n,t} \times AVA_{n,t}) \geq HCAP_t(1+\alpha) \tag{13.13}$$

13.2.5.2 Electricity Production

Availability of each electricity generation technology should not more than a maximum value *AMAX*. As the capacity in the beginning of the year is different from that of the end of the year due to the construction or retirement during the year, availability is mean value of end year value of last year and this year:

$$GELEC_{n,t} \leq AMAX_n \times (QCAP_{n,t} + QCAP_{n,t-1})/2 \tag{13.14}$$

13.2.5.3 Capacity Addition/Retirement

Take several factors into considerations such as unit construction capacity, technology maturity, construction cycle, space limitation and policy inclination, capacity addition/retirement AQCAP should not less than minimum capacity addition *LQCAP*, and no more than maximum capacity addition *HQCAP*:

$$LQCAP_{n,t} \leq AQCAP_{n,t} \leq HQCAP_{n,t} \tag{13.15}$$

13.2.5.4 Total Capacity

Similarly, total capacity *TQCAP* should not less that minimum total capacity *TLQCAP*, and no more than maximum total capacity *THQCAP*:

$$TLQCAP_{n,t} \leq TQCAP_{n,t} \leq LHQCAP_{n,t} \tag{13.16}$$

13.2.5.5 Power Generation Technology Trend

This model assumes there is no capacity addition of less than 300 MW coal-fired unit (n') for the reason that is a backward technology.

$$QCAP_{n',t} \leq QCAP_{n',t-1} \tag{13.17}$$

The advanced/clean technologies including supercritical, ultra-supercritical, CFB, IGCC, NGCC, wind, solar, hydro, and nuclear and bio-fuel electric generation technologies (n'') are the main force of China's future power generation

system. Therefore, capacity of these technologies would no more decrease assumed in this study.

$$QCAP_{n'',t} \geq QCAP_{n'',t-1} \tag{13.18}$$

And all types of capacity should satisfy the nonnegative request.

$$QCAP_{n,t} \geq 0 \tag{13.19}$$

13.2.5.6 Air Emission Cap

Air emission cap *EMIT* is defined as follows, where *REMIT* is the reduction rate compared to reference year.

$$EMIT_{s,t} \leq EMIT_0 \times (1 - REMIT_{s,t}) \tag{13.20}$$

13.3 Description of Case Study

13.3.1 General Structure of Shanghai Power Generation System

Located in China's south-eastern of the Yangtze Delta, Shanghai covers 6340 square kilometers with 24.26 million populations. Shanghai is one of central four municipalities of China with its GDP ranking the first in China and second in Asia in 2014. Power generation system plays a supportive role in Shanghai's economic development and residence daily life.

By 2014, there are 48 power plants in Shanghai. The total capacity reached 30211 MW, of which 15690 MW comes from outside the region. Local units generated 561.175 billion KWh electricity, and that number of imported electricity is 135.903 billion KWh in 2014. Figure 13.2 demonstrates the general structure of Shanghai power generation system.

13.3.2 Data Input

13.3.2.1 Planning Period and Discount Rate

Taking 2014 as the base year and 2012–2014 as historical years, the planning period of this study covers 2015–2030. Considering the discount rate used by IEA to calculate power generation cost was 5 and 10 %, this study adopts 8 % as the

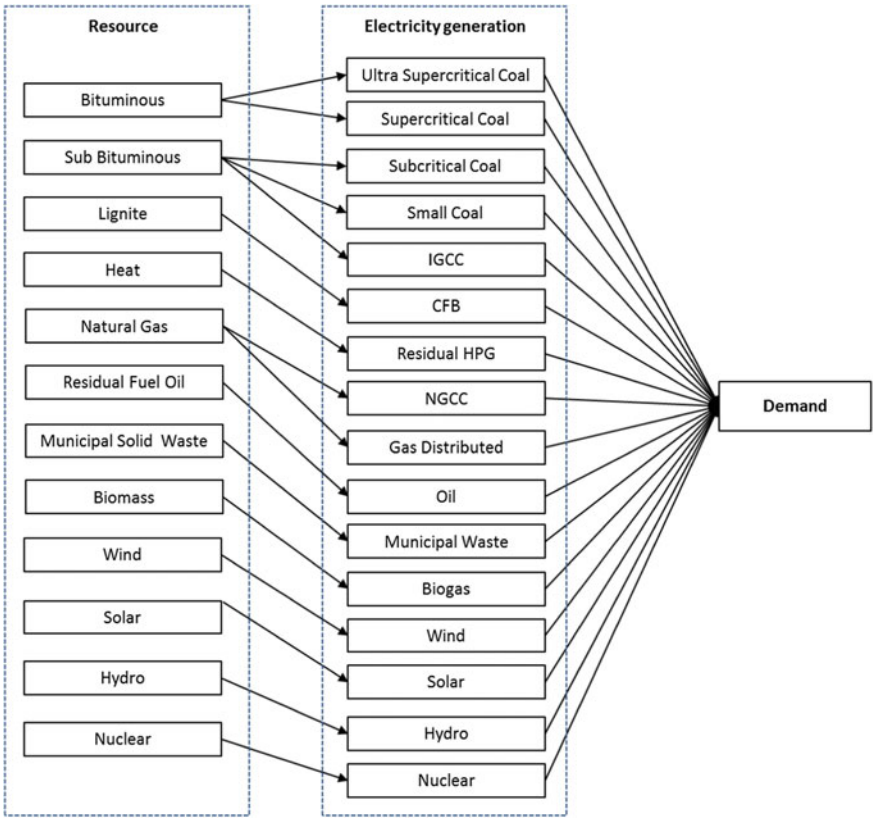

Fig. 13.2 General structure of Shanghai power generation system

discount rate. Since distribution and grid are less affected by policy change within region, the boundary of the study only includes power generation but not distribution and grid system.

13.3.2.2 Demand Projection

Electricity consumption is closely related with economic development. In the past two decades, especially 2000–2010, Shanghai's electricity consumption increase rapidly with the economic growth. As the economic growth has slowed down after 2010, electricity consumption growth weakened respectively. With China's economic steps into slow-growth period and electricity consumption would further decouple with economic growth, it can be predicted that electricity demand increase will keep in a relatively low level. Taking the study results of *Shanghai Mid to Long-term Energy Development Strategy* by Shanghai Development and Reform Committee and our previous study as reference, this study assumes electricity

Table 13.1 Comparison of demand projection of this study with other studies

Thousand Giga-watt hour	2014	2015	2020	2025	2030
LEAP-Shanghai	136.9	133.3	155.4	177.6	200.6
This study	136.9	141.0	163.5	180.5	199.3
SDRC		171.5			228.0

demand growth rate is 3 % in 2015–2020 and slows to 2 % in 2020–2030. As shown in Table 13.1, the result lies in between the above two studies.

13.3.2.3 Technical Characteristics and Cost Data

Technical characteristics and cost data of existing power plant units are shown in Tables 13.2 and 13.3. Data collected from China Power Yearbook editorial board, Chinese Renewable Energy Institute, China Electricity Council, National Energy Berea and National Department of Environment Protection, IEA, etc.

Cost decrease rate of power generation units are shown in Table 13.4, in which capital cost presents the changing rate of initial capital cost which is same with fixed OM cost. Due to the rise of environment cost and migration compensation, as well as

Table 13.2 Technical characteristics of existing power generation units

Power unit	Existing capacity (MW)	Electricity production (Gigawatt-hour)	Efficiency (%)	Lifetime (year)
Ultra-supercritical coal	5320	24,004	46.5	30
Supercritical coal	3000	12,870	44.5	30
Subcritical coal	6529.7	27,163	42.5	30
Small coal	767	4617	34	30
Residual HPG	586.14	1344	100	30
IGCC	0	0	50.5	30
CFBC	33	13	40.5	30
NGCC	4224	8734	48.5	30
Gas distributed	47	14	38	25
Oil	864	196	39	30
Wind	284.35	736	100	20
Solar	215	6	100	20
Municipal waste	136	868	40.5	30
Biogas	15	102	40.6	20
Imported thermal coal	4000	14230.52565	44.5	30
Imported hydro	11,000	39489.70868	100	70
Imported nuclear	690	2454.765674	100	60
Total	37711.19	136,842	–	–

Table 13.3 Power generation unit cost

	Capital cost (RMB/KW)	Fixed OM cost (RMB/KW)	Variable OM cost (RMB/KWh)
Ultra-supercritical coal	3540	106.2	0.02778
Supercritical coal	3901	117	0.02778
Subcritical coal	4443	133.3	0.02778
Small coal	4410	132.3	0.03473
Residual HPG	8000	133.3	0.02778
IGCC	8981	269.3	0.03056
CFBC	4800	144	0.03125
NGCC	3360	100.8	0.031
Gas distributed	5520	100.8	0.031
Oil	3680	110	0.01945
Wind	9606	310	0.014
Solar	12,500	216	0.00049
Municipal waste	4443	133	0.02778
Biogas	9700	390	0.0484
Imported thermal coal	3901	117	0.02778
Imported hydro	6870	105	0.007
Imported nuclear	15,000	600	0.028

Table 13.4 Cost decrease rate

	Capital cost (%)	Variable OM cost (%)
Ultra-supercritical coal	0	0
Supercritical coal	0	0
Subcritical coal	0	0
Small coal	0	0
Residual HPG	0	0
IGCC	0.9	0.75
CFBC	0	0
NGCC	0.5	0
Gas distributed	0.7	0
Oil	0	0
Wind	1.1	0.5
Solar	5	0.5
Municipal waste	0	0
Biogas	1.5	1
Imported thermal coal	0	0
Imported hydro	0	0
Imported nuclear	0	0

Table 13.5 Fuel characteristics and price growth rate

Fuel type	Heating value	Price in 2014	Growth rate (%)
Bituminous	21.766 GJ/t	500 RMB/t	1
Subbituminous	20.092 GJ/t	430 RMB/t	1
Lignite	15.906 GJ/t	330 RMB/t	1
Natural gas	34.2 MJ/m^3	2.72 RMB/m^3	0.5
Residual fuel oil	40.19 GJ/t	2580 RMB/t	0.2
Biomass	13.985 GJ/t	700 RMB/t	0.5

the increasing difficulty for explore, capital cost of hydro is considered to increase. However, as technology improvement helps to reduce cost, it is hard to anticipate the cost trend of hydro. Hence, captital cost of hydro set to unchanged here. Similarly, nuclear power capital cost stays unchanged due to the rising-up security demand.

13.3.2.4 Fuel Characteristic and Price

Fuels involved in this study and price growth rate are shown in Table 13.5. And types of coal for power generation units are shown in Table 13.6.

13.3.2.5 Environment Loading Factor

Emission factors by unit type are shown in Table 13.7. It should be noted that, to simplify the study, only $PM_{2.5}$ is considered for LAPs in this study. Data are collected from *the first national pollution census industrial pollution source emission coefficient manual*, Li (2012), and IPCC (2006). According to an investigation on LAPs emission control measurement adoption in Shanghai (Ding et al. 2015), it is assumed all coal-fired units emission factors are SCR+ESP+FGD based in this study.

Table 13.6 Power generation unit and the corresponding fuel type

Power generation unit	Fuel type
Ultra-supercritical coal	Bituminous
Bituminous	Bituminous
Subcritical coal	Sub-bituminous
Small coal	Sub-bituminous
IGCC	Sub-bituminous
CFBC	Lignite
NGCC	Natural gas
Gas distributed	Natural gas
Oil	Residual fuel oil
Municipal waste	Municipal solid waste
Biogas	Biomass
Imported thermal coal	Bituminous

Table 13.7 Emission factors for different types of power generation unit

Per fuel consuming	CO_2 (t/TJ)	$PM_{2.5}$ (kg/t)
Ultra-supercritical coal	94.0087	0.3657
Supercritical coal	94.0087	0.3657
Subcritical coal	95.1577	0.3657
Small coal	95.1577	0.8
IGCC	95.1577	0.3657
CFB	100.4004	0.8
NGCC	55.7811	–
Gas distributed	55.7811	–
Residual fuel oil	76.5402	0.4533
Municipal waste	123.5638	0.1258
Biomass	109.5578	0.1975
Import thermal coal	94.0087	0.3675

13.3.2.6 Dispatch Rule

As the principle of energy savings, the dispatch rule for power generation system in this study is set as follows:

a. Wind, solar, hydro that cannot be adjusted;
b. Bio-fuel generation unit that is environmental friendly renewables;
c. Nuclear;
d. CHP, residual heat, gas and pressure unit, gangue unit etc.;
e. Natural gas and coal gasification unit;
f. Other coal-fired units;
g. Oil units.

13.3.2.7 Maximum Capacity

It is assumed that despite ultra-supercritical unit, capacity of other coal-fired units will keep unchanged or reduced in this study. Oil units will retire gradually during the planning period. Maximum capacity for different types of units is shown in Table 13.8.

13.3.3 Scenario Design

13.3.3.1 Scenario Contents and Relations

RESO model built here could also been used for policy simulation of carbon tax, carbon price, pollutant discharge and other environment protection incentive

Table 13.8 Maximum capacity of different types of unit

Unit	Maximum capacity (MW)	Unit	Maximum capacity (MW)
Ultra-supercritical coal	Unlimited	Residual fuel oil	864
Supercritical coal	3000	Wind	6500
Subcritical coal	6530	Solar	500
Small coal	767	Municipal waste	600
Residual HPG	586	Biomass	200
IGCC	Unlimited	Import thermal coal	5400
CFB	Unlimited	Import hydro	12,700
NGCC	Unlimited	Import nuclear	1010
Gas distributed	Unlimited		

policies. Planning capacity addition and air emission cap are the two main simulating variables that this study concerned about. Table 13.9 summarize the contents of six scenarios and their relations within this study:

a. Baseline Scenario (BASE) assumes no capacity planning and emission cap policy in the planning period. It simulates the power generation system change under current situation.
b. Advanced Technology Scenario (ADVT) simulates system response under existing planning. The existing plan of ADVT is mainly focus on advanced/clean units such as ultra-supercritical substitute for subcritical and small coal-fired unit.
c. Optimization Scenario (OPTI) is built based on ADVT simulating capacity addition under minimizing cost target. However, there is no emission constrain imposed on this scenario.

Table 13.9 Scenario content and relations

Scenarios	Optimization	CO_2 cap	$PM_{2.5}$ cap	Content
BASE	No	No	No	Remain unchanged
ADVT	No	No	No	Shutdown small coal unit, early retirement of sub ultra-supercritical, increase capacity of gas distributed, wind, solar
OPTI	Yes	No	No	System cost minimization
CO2C	Yes	Yes	No	Carbon dioxide cap
LAPC	Yes	No	Yes	Local air pollutant emission cap
CAEC	Yes	Yes	Yes	Carbon dioxide + local air pollutant emission cap

d. CO_2 Cap Scenario (CO2C) adds CO_2 cap constrains. It simulates system response and mitigation co-benefits.
e. Local Air Pollutant Cap Scenario (LAPC) adds LAP cap constrains. It simulates system response and mitigation co-benefits.
f. Combined Air Emission Cap Scenario (CAEC) simulates system influence that controls CO_2 and LAP emission at the same time.

13.3.3.2 Exogenous Capacity Added/Retirement

According to scenario design, capacity addition is endogenous under the four optimizing scenarios which are CO2C, LAPC and CAEC. Thus, exogenous capacity under ADVT and the above four scenarios are designed according to existing plan in Table 13.10. Other technologies that not mentioned are assumed to have no exogenous capacity addition.

13.3.3.3 Air Emission Cap

China has not launched mandatory carbon emission reduction targets for power generation sector, yet proposed a target that reaching peak emission in 2030 and reducing carbon intensity by 60–65 % by 2030 compared to 2005 level on the 2015 climate conference in Paris. Hence, it is assumed that CO_2 emission of Shanghai's power generation system peaks in around 2025 and is reduced by 5 % compared to 2025 level in this study. According to *Shanghai Clean Air Act (2013–2017)*, by the end of 2017, Shanghai's air quality will be improved substantially, $PM_{2.5}$ concentration decline 20 % than that of 2012. Thus, in LAPC scenario, $PM_{2.5}$ is suggested to reduce 20 % than 2030 and 35 % than 2012. Emission caps in CAEC scenario is the caps of both CO2C scenario and LAPC scenario.

Table 13.10 Exogenous capacity (*Unit* MW)

MW	2015	2016	2020	2025	2030
Supercritical coal		5229.7	3949.7		1500
Small coal			400		0
IGCC				800	
Oil					0
Wind	869.35		1869.35		6000
Solar			300		

13.4 Results and Discussion

13.4.1 Electricity Production

There is a shortage of electricity under BASE scenario and ADVT scenario due to scenario design. Model results show that from the year of 2015, the gap between production and demand expends over time in the above two scenarios. As shown in Fig. 13.3, by 2030, production shortage in baseline is 26.8 billion KWh, and 42.1 billion KWh in ADVT scenario. With the assumption that electricity demand is fully met, the production gap in the other four optimizing scenarios is zero.

Electricity production constitutions of four optimizing scenarios are shown in Fig. 13.4. In OPTI scenario which had no emission constrain, ultra-supercritical produces 52.5 % of total electricity production and the portion lowers to 34.4 % in CO_2 scenario, 33.6 % in LAPC scenario and 29.6 % in CEAC scenario. Thanks to the advantages of economic and environmental friendly, imported hydro and nuclear has rapidly reached their maximum generating capacity in all four optimizing scenarios. NGCC and gas distributed units rank highest in LAPC and CAEC scenario account for 26.19 % and 25.65 % respectively. Coal-fired units including imported thermal power are still the main force of Shanghai's electricity production, accounting for 63.3 % in OPTI scenario and 36.4–43.1 % in the other three optimizing scenarios.

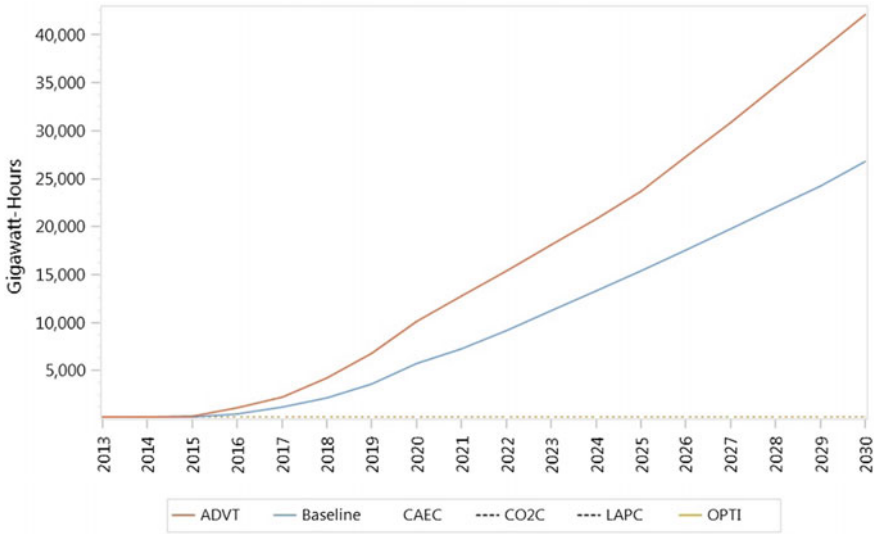

Fig. 13.3 Electricity demand shortage under different scenarios (2015–2030)

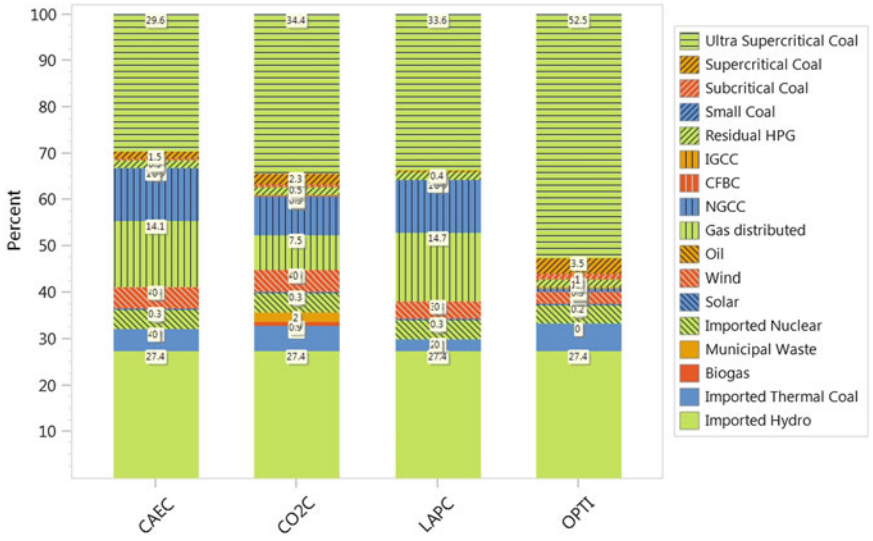

Fig. 13.4 Electricity production constitution in four optimizing scenario in 2030

13.4.2 Installed Capacity and Capacity Additions

Total capacity installed in BASE scenario is unchanged according to scenario setting. Solar and wind capacity growth could not compensate the retirement of subcritical and small coal-fired units, thus installed capacity in ADVT come down slightly. Installed capacity in the other four optimizing scenarios grows over time. As shown in Fig. 13.5, installed capacity in BASE scenario reaches 58177 MW.

Ultra-supercritical unit takes the biggest share of installed capacity. In 2030, capacity of ultra-supercritical unit accounted more than one third in OPTI scenario and about a quarter in CAEC scenario (see Fig. 13.6).

Capacity addition of the four optimizing scenarios in short-term varies little before in short-term (Fig. 13.7). CO_2 restriction policy has an influence on capacity addition from the year of 2017 with the capacity increase of 978 MW in CO2C scenario. During 2018–2021, capacity addition in OPTI and LAPC scenarios is the same, while that bulk in CO2C and CEAC scenarios is the same. Thus, it is indicated that not LAP control policy but CO_2 restriction policy has a reduction effect during that period of time. Until 2022, LAP control policy has begun to work, exogenous capacity addition reaches 971 MW. However, as not affected by emission control policies, endogenous capacity of OPTI is the smallest in all planning year.

As shown in Fig. 13.8, electricity generation system prefers imported nuclear, hydro and solar as capacity addition in no emission cap scenario namely OPTI scenario and ultra-supercritical units as the second but the biggest source. Similarly, wind power becomes an important supplement since the year of 2017 under CO2C scenario. And NGCC and gas distributed units take better place in the long run under CO2C and LAPC scenario.

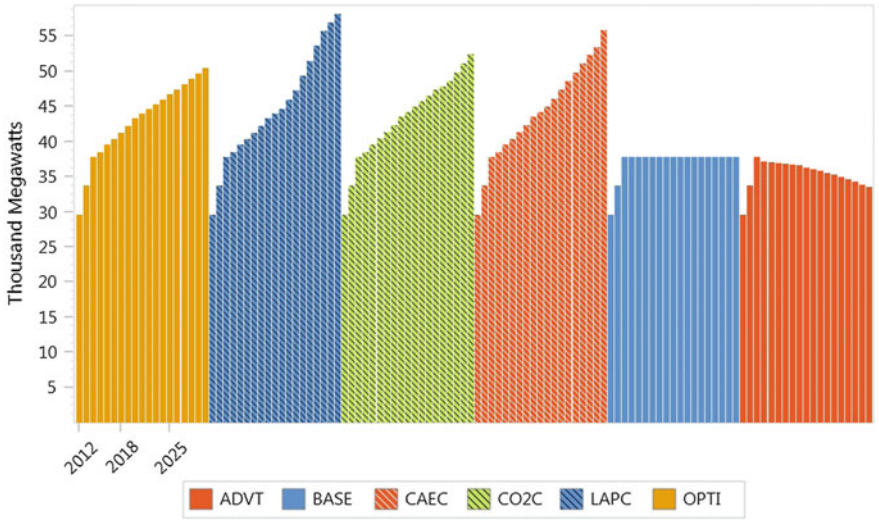

Fig. 13.5 Installed capacity in different scenarios (2012–2030)

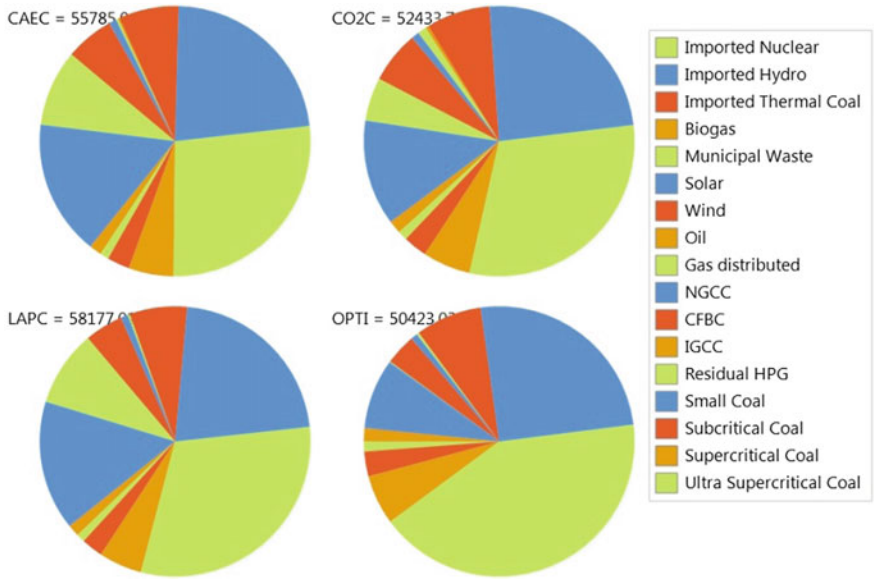

Fig. 13.6 Constitution of installed capacity by electricity generation technology in 2030

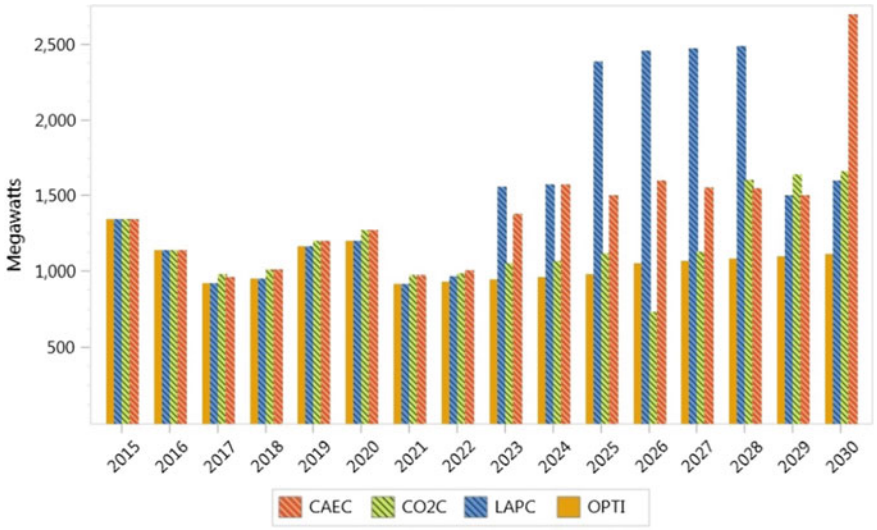

Fig. 13.7 Capacity addition under four optimizing scenarios (2015–2030)

13.4.3 Fuel Consumption

Energy consumption witnesses an increasing trend throughout the planning period from 862.65 million GJ in 2015 to 1261.15 million GJ in 2030 under BASE scenario (Fig. 13.9). The growth rate of energy consumption is consistent with electricity demand which is 3.84 % in 2020 and slowed down to 2.25 % ever since.

Energy-saving impact by emission control policy is shown in Fig. 13.10. Taking OPTI scenario as baseline, energy savings of CO2C, LAPC and CAEC scenarios increased year by year, among which CO2C and CAEC scenario climbs to 43.74 million GJ in 2029.

13.4.4 CO_2 Emissions Co-benefits

Without CO_2 constrain (under OPTI scenario and LAPC scenario), CO2 emission increases continuously reaching 94,000 thousand tons in 2030 (Fig. 13.11). CO_2 emission in LAPC scenario peaks at 89,520 thousand tons in 2029 then declined to 72,370 thousand tons in 2030. CO_2 emission in CO2C and CAEC scenario grow to 80,000 thousand tons in around 2025 then falls to 75,000 thousand tons due to scenario settings.

Taking OPTI scenario as reference, CO_2 reduction in other three emission control scenarios is shown in Fig. 13.12. It can be observed that, LAPs constrain policy has little co-benefit on CO_2 reductions with the largest reduction of 21.62 million tons in 2030.

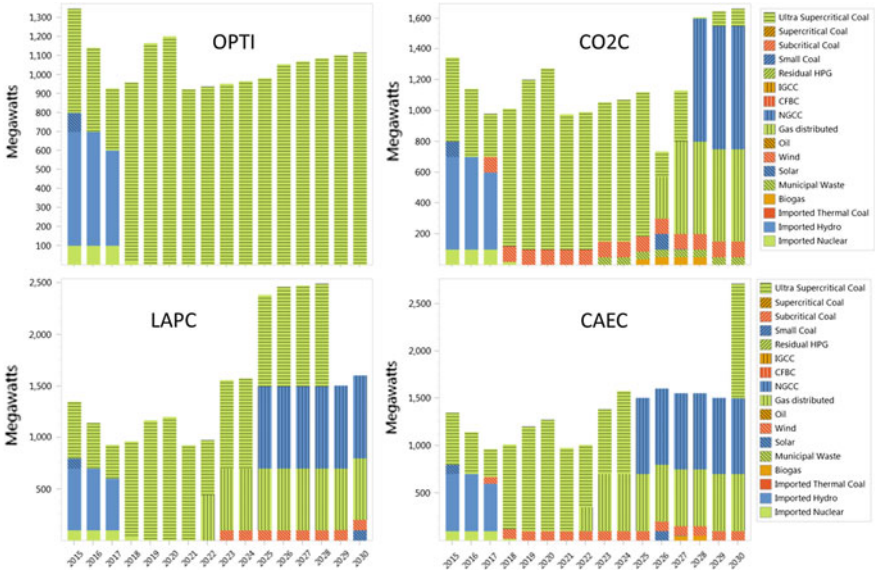

Fig. 13.8 Constitution of capacity added by power generation technology in four optimizing scenario

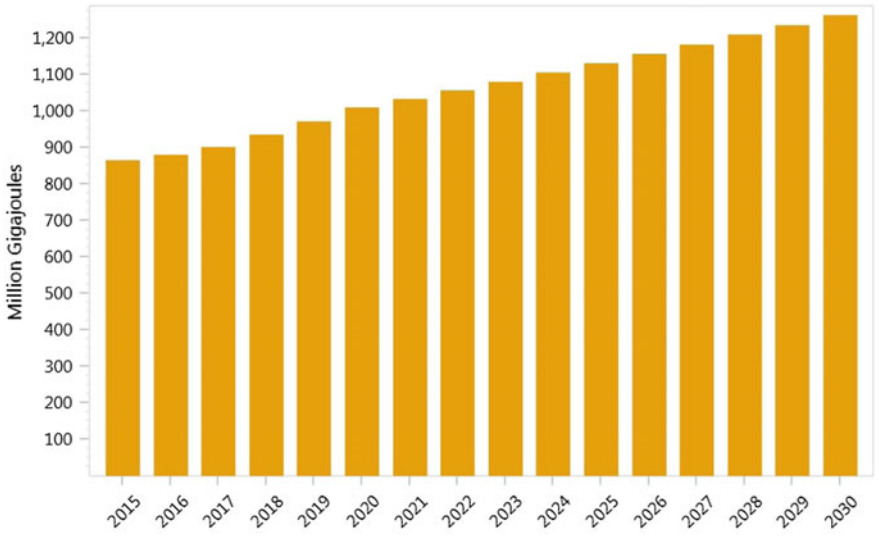

Fig. 13.9 Energy consumption under BASE scenario (2015–2030)

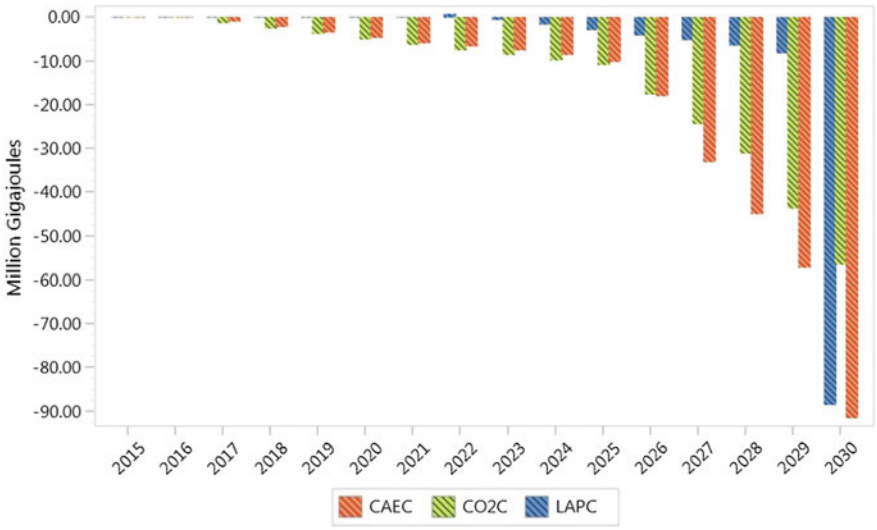

Fig. 13.10 Energy savings of three emission control scenarios versus OPTI scenario (2015–2030)

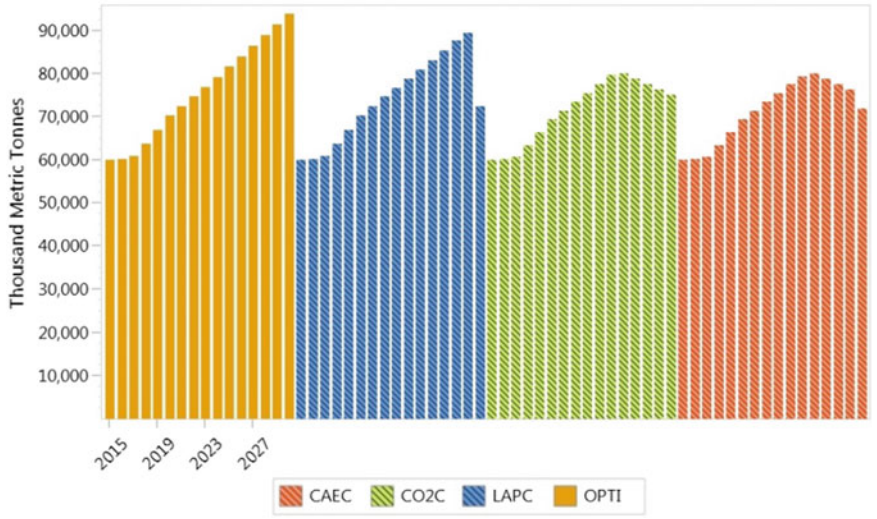

Fig. 13.11 CO_2 emissions in the four optimizing scenarios (2015–2030)

Co-benefits of carbon emission mitigation on $PM_{2.5}$ reach 4579 tons in 2030 under CO2C scenario (Fig. 13.13). Ultra-supercritical units contribute most reductions before 2025, which however, lead to more carbon emissions from municipal waste and bio-fuel units in the long-term.

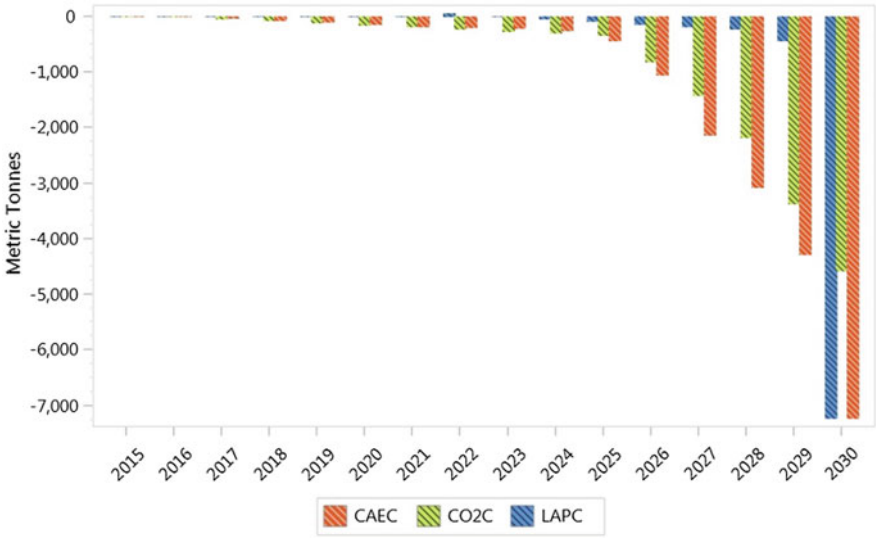

Fig. 13.12 CO_2 reductions in other three emission control scenarios (2015–2030)

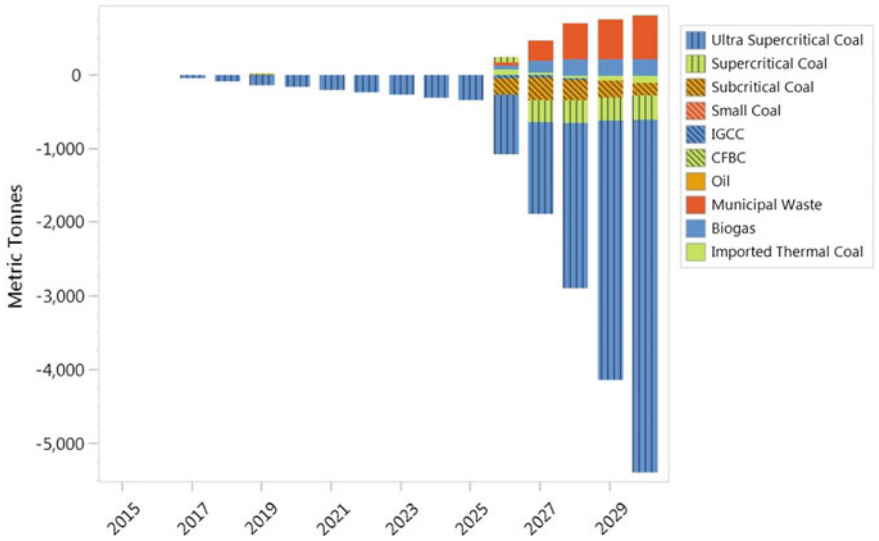

Fig. 13.13 Co-benefits of carbon emission mitigation on $PM_{2.5}$

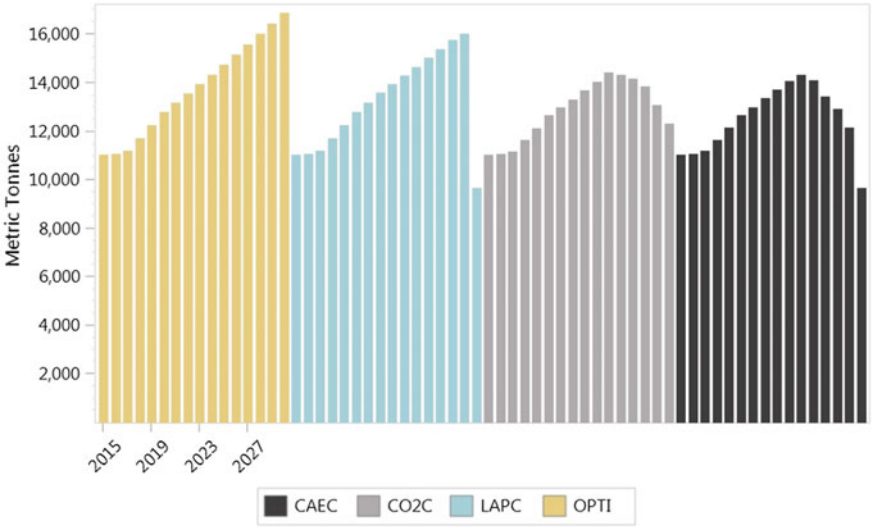

Fig. 13.14 $PM_{2.5}$ emission in the four optimizing scenarios (2015–2030)

13.4.5 LAP Emissions Co-benefits

$PM_{2.5}$ emission keeps continuously increasing throughout the planning period reaching 16,892 tons in 2030, whereas it is constrained under LAPC and CAEC scenario according to scenario settings (Fig. 13.14).

Taking OPTI scenario as reference, co-benefits of $PM_{2.5}$ from CO2C is larger than that of LAPC throughout the planning period except the year of 2030. It means that carbon migration policy has greater influence on $PM_{2.5}$ reductions than LAP control policy (Fig. 13.15).

Co-benefits of carbon reductions from LAPs controls are apparent from the year of 2022 among which ultra-supercritical units attribute most (Fig. 13.16). Whereas, co-benefits of one power generation technology is not always positive, NGCC units performance positive effect during 2024–2027 but negative during 2028–2030 for instance.

13.4.6 Reduction Cost

There is a slightly increase of cumulative system cost before 2025 (Fig. 13.17). Cumulative system cost is 693.6 billion rmb in OPTI which is the lowest, and the other three optimizing scenarios are 732.1 billion rmb (LAPC), 738.6 billion rmb (CO2C) and 751.5 billion rmb (CAEC), respectively.

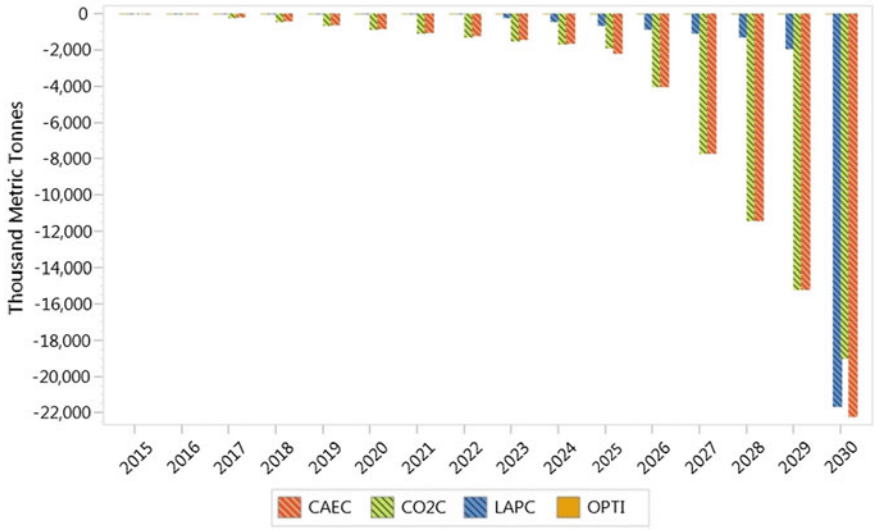

Fig. 13.15 $PM_{2.5}$ reductions in other three emission control scenarios (2015–2030)

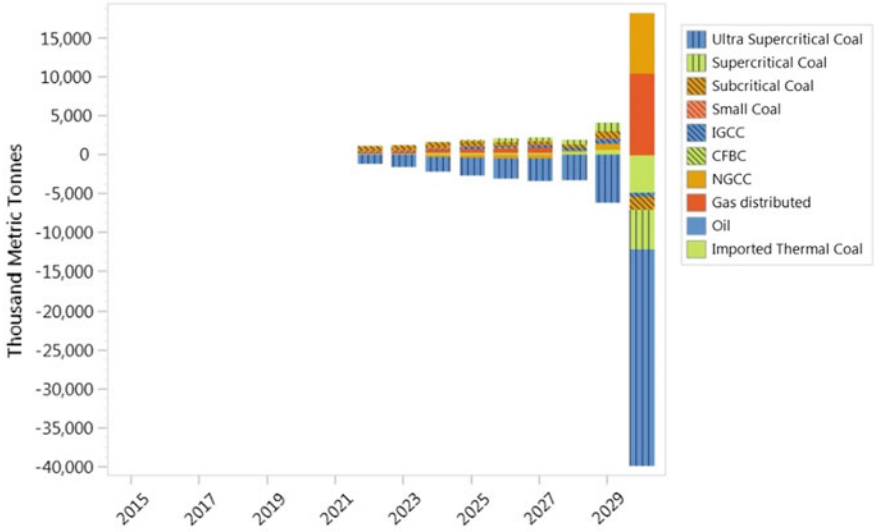

Fig. 13.16 Co-benefits of $PM_{2.5}$ reduction policy on CO_2 emission mitigation

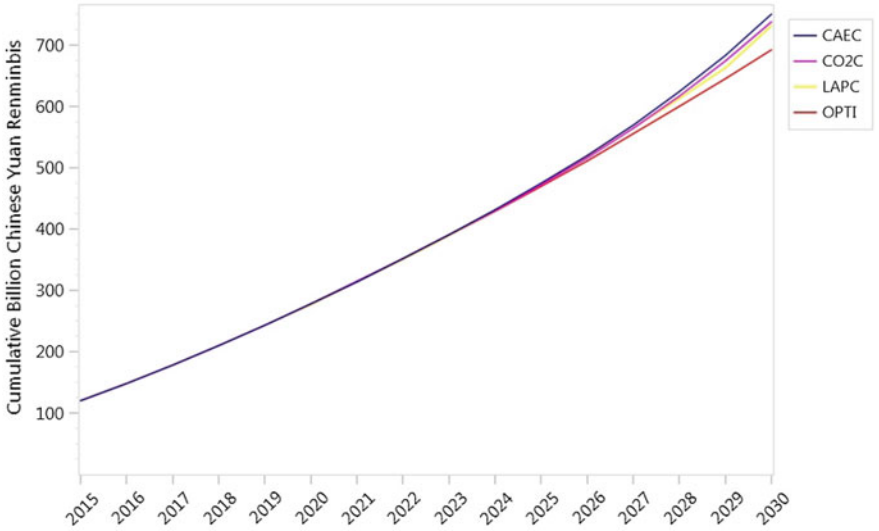

Fig. 13.17 Cumulative system cost in four optimizing scenarios (2015–2030)

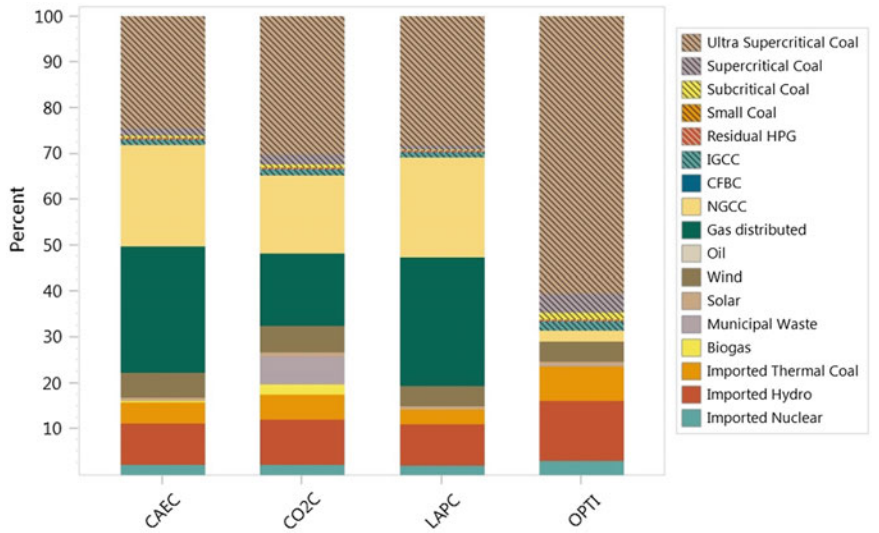

Fig. 13.18 Constitution of total system cost under the four optimizing scenarios

As shown in Fig. 13.18, the share of ultra-supercritical units cost ranks highest in OPTI scenario. NGCC and gas distribution cost more in the three emission constrain scenarios.

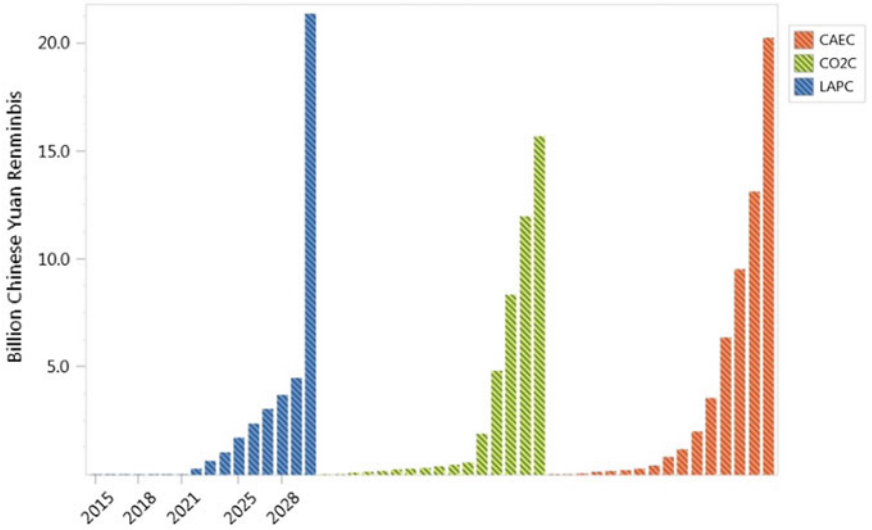

Fig. 13.19 Cost increment of three emission control scenarios compared to OPTI scenario (2015–2030)

Carrying out emission constrain policies would rise up system cost (Fig. 13.19). Taking OPTI scenario as reference, $PM_{2.5}$ control policy does not increase system cost until 2022. The effect of cost increment appears clearly since 2025 which grows to 15.68 trillion rmb in 2030.

Average reduction cost per ton per year is acquired from cumulative cost derived by total reductions and by planning years (16 year). The result shows that, CO_2 reduction cost is 42.2 rmb/ton/year in CO2C scenario, and 29.04 thousand rmb/ton/year for $PM_{2.5}$ reduction.

13.4.7 Energy Flow

As an important method for energy utilization analysis, energy flow chart is a visualized way to demonstrate the direction and volume of energy. It can be used to reflect the process of energy used by electric generation system from supply to transformation and to end use from which energy utilization characteristics such as energy structure and efficiency could be estimated. Figures 13.20, 13.21, 13.22 and 13.23 are energy flow charts of electric generation system in 2030 under four optimizing scenarios.

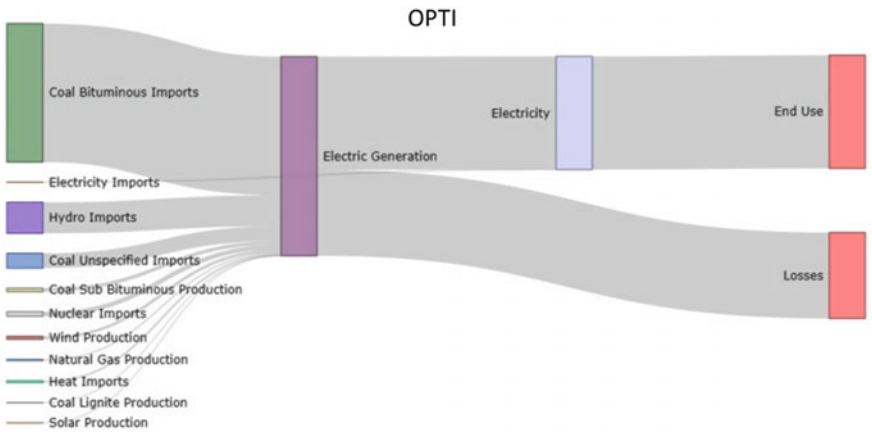

Fig. 13.20 Energy flow charts of electric generation system in 2030 under OPTI scenario

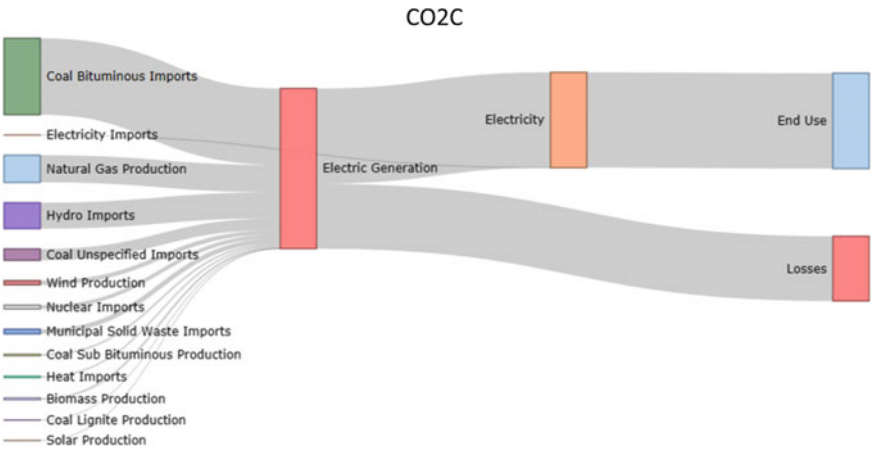

Fig. 13.21 Energy flow charts of electric generation system in 2030 under LAPC scenario

13.5 Sensitivity Analysis

The robustness of RPSO model should be tested through sensitivity analysis, since the model results may be heavily dependent on parameter assumptions. Taking CAEC scenario in reference case as the baseline, a sensitivity analysis is conducted to examine the impact of variable changes on model results. As there are no limitations on the expansion of ultra-supercritical, IGCC, NGCC and gas distributed units according to scenario design, this part will mainly focus on the sensitivity analysis of these technologies to the variable change.

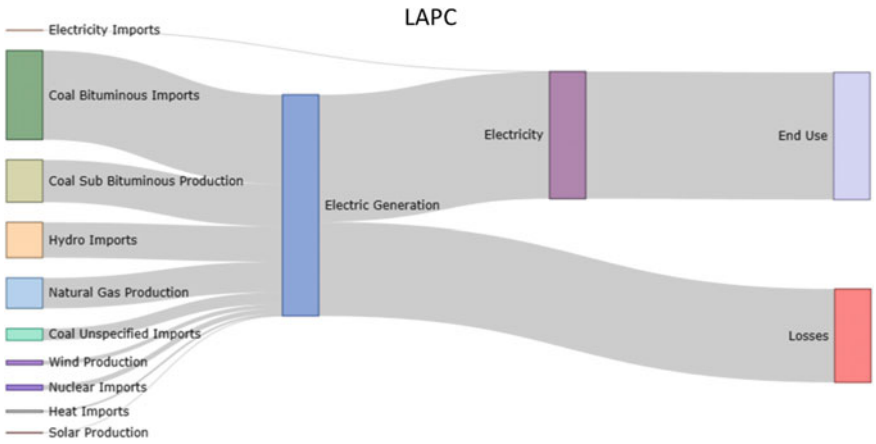

Fig. 13.22 Energy flow charts of electric generation system in 2030 under LAPC scenario

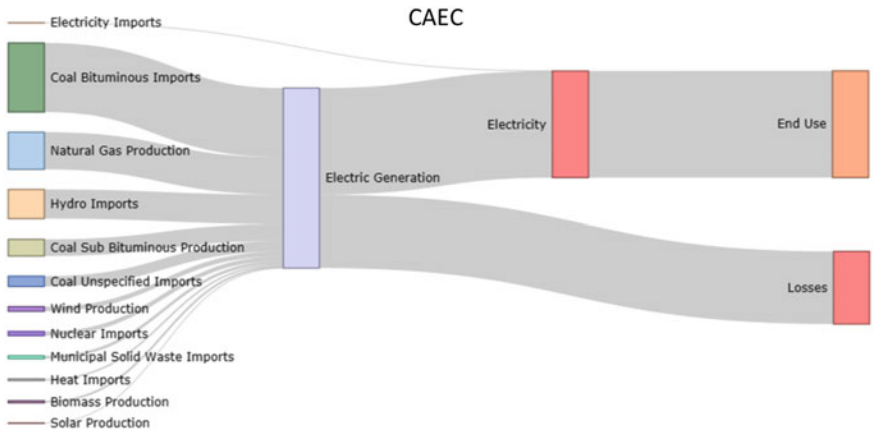

Fig. 13.23 Energy flow charts of electric generation system in 2030 under CO2C scenario

13.5.1 Impact of Electricity Demand

A 2.5, 5 and 7.5 % increase or drop from the reference case value is examined to see the influence on total system cost and capacity added of above technologies (Table 13.11). Electricity demand have a significant influence on total system cost and capacity addition. The higher electricity demand is, the more system cost will be charged and capacity added. Total system cost would increase 15.39 % if electricity demand climbs 7.5 %, while declining 12.37 % when electricity drops 7.5 %.

Table 13.11 System cost (2015–2030) of the electricity generation system expansion and power capacity (2030) for different

Increase rate of the electricity demand (%)	Electricity demand in 2030 (Thousand gigawatt-hour)	Total Cost (billion rmb)	%	Ultra-supercritical capacity in 2030 (MW)	NGCC capacity in 2030 (MW)	Gas distributed capacity in 2030 (MW)
−7.5	184323.10	658.52	87.63	13998	7424	3391
−5	189304.80	686.67	91.38	13351	8224	3867
−2.5	194286.51	717.76	95.51	14093	9024	4271
0 (Reference case)	199268.22	751.48	100.00	14042	9025	5098
2.5	204249.92	788.13	104.88	14443	9024	5598
5	209231.63	826.09	109.93	15132	10624	6051
7.5	214213.33	867.10	115.39	14604	11097	6692

13.5.2 Impact of Fuel Price

The sensitivity analysis is imposed on fuel price as well. We take natural gas price as the case. The results shows that, a 1 % natural gas price drop would lead to a 1.55 % decline of total cost, while a 2 % price rise leads to a 2.44 % total cost increase (Table 13.12). Similarly, electricity production of natural gas units shrinks when the price goes up.

13.5.3 Impact of Carbon Emission Factor

Air emission is calculated by fuel consumption and emission factor. Thus, the uncertainty of emission factor would have an impact on emission and electricity generation pathways. The emission factors of oil and natural gas in China are

Table 13.12 System cost (2015–2030) and cumulative power production by 2030 for different technologies

Increase rate of the fuel price (%)	Fuel price in 2030 (rmb/m^3)	Total cost (billion rmb)	%	Total natural gas electric production (Gigawatt-hour)	Gas distributed electricity production by 2030 (Gigawatt-hour)
−1	2.32	739.82	98.45	51662	28737
−0.5	2.51	744.45	99.06	51389	28464
0	2.72	749.59	99.75	51310	28385
0.2 (Reference case)	2.81	751.48	100.00	51061	28136
1	3.19	759.75	101.10	50781	27856
1.5	3.45	764.62	101.75	50452	27527
2	3.73	769.79	102.44	50445	27520

Table 13.13 System cost under the fluctuation of coal emission factors for different types of power generation unit

Unit type	CO_2 (t/TJ)			$PM_{2.5}$ (kg/t)		
	−5 %	Reference case	5 %	−5 %	Reference case	5 %
Ultra-supercritical coal	89.31	94.01	98.71	0.3474	0.3657	0.38
Supercritical coal	89.31	94.01	98.71	0.3474	0.3657	0.38
Subcritical coal	90.40	95.16	99.92	0.3474	0.3657	0.38
Small coal	90.40	95.16	99.92	0.7600	0.8000	0.84
IGCC	90.40	95.16	99.92	0.3474	0.3657	0.38
CFB	95.38	100.40	105.42	0.7600	0.8000	0.84
System cost	739.76	751.48	762.23	739.38	751.48	761.55

basically the same with IPCC default value. However, due to the big difference of category standard, the uncertainty of China's emission factor by coal rank is huge. The sensitivity analysis for coal emission factors (Table 13.13) shows that the rise of emission factor for both CO_2 and $PM_{2.5}$ would lead to higher system, vice versus. When emission factors of coal fluctuates ±5 %, total system cost would rise or drop in between −1.56 and 1.43 %, that range for $PM_{2.5}$ is −1.61 and 1.34 %.

13.6 Conclusions

RPSO model built in this study can be applied for regional power generation system planning under multiple air emission constrains. The optimization pathways for power generation system are acquired through establishing various policy scenarios. Baseline scenario is set to examine whether it is necessary to add capacity. Advanced technology scenario estimates whether existing plan could meet future electricity demand, based on which four optimizing scenario are designed to simulated optimized system pathways and emission co-benefits according to different emission constrains.

Compared to no emission constrain, emission cap would increase system cost and installed capacity, but also improve construction of system and reduce energy consumptions. Specifically, system cost in CAEC scenario is the highest. LAPC scenario installs the most new capacity, while share of non-fossil capacity accounts the highest in CAEC scenario with electricity production accounting for 64 %. Energy savings is larger in CO2C and LAPC scenario.

When no emission constrain is imposed, ultra-supercritical is the first choice for capacity addition while still accounted relatively high shares under emission-constrained scenarios. With the characteristics of environmental friendly and economy, hydro and nuclear becomes the optimal source of new capacity. However, restricted to resource endowment, there is difficulty for those to

continuously support capacity added. NGCC and gas distributed unit get greater development under strict emission policies with electricity production accounting 26.2 and 25.6 % in LAPC and CAEC scenario in 2030. Wind and solar power plays a cleaning role in structure improvement that should be brought to more policy incentives. Under CAEC scenario, they attributed 4.9 % of total electricity production.

For the aspect of emission reduction co-benefits, ultra-supercritical units attributed most of the $PM_{2.5}$ reduction co-benefits in CO2C scenario. However, $PM_{2.5}$ restriction policy came into play in the long-term. Therefore, it is believed that the co-benefits of carbon mitigation policy are bigger than that of $PM_{2.5}$ reductions.

References

Andrade Guerra, J. B. S. O., Dutra, L., Schwinden, N. B. C. O., & Andrade, S. F. D. (2015). Future scenarios and trends in energy generation in Brazil: Supply and demand and mitigation forecasts. *Journal of Cleaner Production, 103*, 197–210.

Bollen, J. (2015). The value of air pollution co-benefits of climate policies: Analysis with a global sector-trade CGE model called WorldScan. *Technological Forecasting and Social Change, 90*, 178–191.

Cai, W., Wang, C., Wang, K., Zhang, Y., & Chen, J. (2007). Scenario analysis on CO_2 emissions reduction potential in China's electricity sector. *Energy Policy, 35*, 6445–6456.

Chen, C., Wang, B., Fu, Q., Green, C., & Streets, D. G. (2006). Reductions in emissions of local air pollutants and co-benefits of Chinese energy policy: A Shanghai case study. *Energy Policy, 34*, 754–762.

Dias, M. V. X., Haddad, J., Horta Nogueira, L., Costa Bortoni, E. D., Passos Da Cruz, R. A., Akira Yamachita, R., et al. (2014). The impact on electricity demand and emissions due to the introduction of electric cars in the São Paulo Power System. *Energy Policy, 65*, 298–304.

Ding, Q., Wei, W., Shen, Q., & Sun, Y. (2015). Major air pollutant emissions of coal-fired power plant in Yangtze River Delta. *Environmental Science, 36*, 2389–2394. Retrieved December 11, 2015, from http://d.g.wanfangdata.com.cn/Periodical_hjkx201507009.aspx.

He, K., Lei, Y., Pan, X., Zhang, Y., Zhang, Q., & Chen, D. (2010). Co-benefits from energy policies in China. *Energy, 35*, 4265–4272.

Hong, S., Chung, Y., Kim, J., & Chun, D. (2016). Analysis on the level of contribution to the national greenhouse gas reduction target in Korean transportation sector using LEAP model. *Renewable and Sustainable Energy Reviews, 60*, 549–559.

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R. K. Pachauri and L. A. Meyer (Eds.)]. IPCC, Geneva, Switzerland, 151 pp. in IPCC AR5 Synthesis Report website.

IPCC. (2006). 2006 IPCC guidelines for national greenhouse gas inventories. *Energy*, *2*. Retrieved October 05, 2015, from http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol2.html.

IPCC. (2016). *Emission Factor Database*. Retrieved October 25, 2015, from http://www.ipcc-nggip.iges.or.jp/EFDB/main.php.

Jiang, P., Chen, Y., Geng, Y., Dong, W., Xue, B., Xu, B., et al. (2013). Analysis of the co-benefits of climate change mitigation and air pollution reduction in China. *Journal of Cleaner Production, 58*, 130–137.

Li, B., Li, J., Lu, J., Sheng, J., Zhang, Q., Yao, G., et al. (2012). Status of larger scale circulating fluidized bed boiler operation in China. *Boiler Technology, 43*, 22–28. Retrieved November 02, 2015, from http://d.g.wanfangdata.com.cn/Periodical_gljs201201006.aspx.

Ma, D., Wang, L., et al. (2014). Co-benefits of reducing carbon emissions in China's iron and steel industry. *Energy Procedia, 61*, 1557–1560.

Ma, Z., Xue, B., Geng, Y., Ren, W., Fujita, T., Zhang, Z., et al. (2013). Co-benefits analysis on climate change and environmental effects of wind-power: A case study from Xinjiang, China. *Renewable Energy, 57*, 35–42.

Mao, X., Yang, S., Liu, Q., Tu, J., & Jaccard, M. (2012). Achieving CO_2 emission reduction and the co-benefits of local air pollution abatement in the transportation sector of China. *Environmental Science & Policy, 21*, 1–13.

McKinsey & Company. (2013). *Pathways to a Low-Carbon Economy: Version 2 of the Global Greenhouse Gas Abatement Cost Curve*. Retrieved November 10, 2015, from http://www.mckinsey.com/business-functions/sustainability-and-resource-productivity/our-insights/pathways-to-a-low-carbon-economy.

McPherson, M., & Karney, B. (2014). Long-term scenario alternatives and their implications: LEAP model application of Panama's electricity sector. *Energy Policy, 68*, 146–157.

Menikpura, S. N. M., Santo, A., & Hotta, Y. (2014). Assessing the climate co-benefits from Waste Electrical and Electronic Equipment (WEEE) recycling in Japan. *Journal of Cleaner Production, 74*, 183–190.

Ministry of Environmental Protection of the People's Republic of China (MEPC). (2015). *2014 China Environment Status Communique*. Retrieved November 20, 2015, from http://www.zhb.gov.cn/gkml/hbb/qt/201506/t20150604_302855.htm.

Mrkajic, V., Vukelic, D., & Mihajlov, A. (2015). Reduction of CO_2 emission and non-environmental co-benefits of bicycle infrastructure provision: The case of the University of Novi Sad, Serbia. *Renewable and Sustainable Energy Reviews, 49*, 232–242.

Perwez, U., & Sohail, A. (2014). GHG emissions and monetary analysis of electric power sector of Pakistan: Alternative scenarios and it's implications. *Energy Procedia, 61*, 2443–2449.

SEI. (2015). *An Introduction to LEAP*. Retrieved December 12, 2015, from http://www.energycommunity.org/default.asp?action=47.

Suhono, S. (2015). Long-term electricity demand forecasting of Sumatera system based on electricity consumption intensity and Indonesia population projection 2010-2035. *Energy Procedia, 68*, 455–462.

Xi, Y., Fei, T., & Gehua, W. (2013). Quantifying co-benefit potentials in the Chinese cement sector during 12th Five Year Plan: An analysis based on marginal abatement cost with monetized environmental effect. *Journal of Cleaner Production, 58*, 102–111.

Xue, B., Ma, Z., Geng, Y., Heck, P., Ren, W., Tobias, M., et al. (2015). A life cycle co-benefits assessment of wind power in China. *Renewable and Sustainable Energy Reviews, 41*, 338–346.

Yang, X., Teng, F., & Wang, G. (2013). Incorporating environmental co-benefits into climate policies: A regional study of the cement industry in China. *Applied Energy, 112*, 1446–1453.

Zhang, S., Worrell, E., & Crijns-Graus, W. (2015). Evaluating co-benefits of energy efficiency and air pollution abatement in China's cement industry. *Applied Energy, 147*, 192–213.

Chapter 14
Model-Based Predictive Control of Integrated Fuel Cell Systems—From Design to Implementation

Chrysovalantou Ziogou, Simira Papadopoulou, Efstratios Pistikopoulos, Michael Georgiadis and Spyros Voutetakis

Abstract Fuel cell systems are a promising alternative to traditional power sources for a wide range of portable, automotive and stationary applications and have an increasing potential for wider use as the demand for clean energy is increasing and the focus is shifting towards renewable energy generation. This chapter has a multidisciplinary scope, the design of a computer-aided framework for monitoring and operation of integrated fuel cell systems and the development of advanced model-based control schemes. The behavior of the framework is experimentally verified through the online deployment to an automated small-scale fuel cell unit, demonstrating excellent response in terms of computational effort and accuracy with respect to the control objectives.

14.1 Introduction and Motivation

Fuel cell and hydrogen technologies have the potential to contribute to the ambitious energy and climate objectives of the European Union for 2020 which are the reduction of the greenhouse gas emissions by 20 %, the increase of the share of renewable energy to 20 % and the improvement of the energy efficiency by 20 %

C. Ziogou (✉) · S. Voutetakis
Centre for Research and Technology Hellas (CERTH), Chemical Process and Energy Resources Institute (CPERI), Thessaloniki, Greece
e-mail: cziogou@cperi.certh.gr

S. Papadopoulou
Department of Automation Engineering, Alexander Technological Educational Institute of Thessaloniki, Thessaloniki, Greece

E. Pistikopoulos
Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, USA

M. Georgiadis
Department of Chemical Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_14

(EC 2001). At longer term, they could play a significant role in supporting Europe and industrialized countries meeting the 2050 targets of 80–95 % reductions in CO2 emissions. Hydrogen and fuel cells, by enabling the so-called hydrogen economy, hold great promise for meeting in a quite unique way, concerns over security of supply and climate change. The last decade, significant research efforts have been allocated (Fuel Cell Today 2012) to the development of fuel cell components and integrated systems, since they constitute an efficient energy conversion technology for transforming hydrogen, and other fuels, into electricity. Overall fuel cell (FC) systems are part of a prominent key enabling technology for achieving carbon free electricity generation and can be used for stationary, mobile and portable applications. During the past few years, the fuel cell market continued to grow, especially in the stationary applications area, including utility-scale fuel cells, fuel cells for industrial and commercial buildings, and fuel cells for residential power. The demand for FCs is driven by an increasing awareness and the need for zero emission energy sources. Also it is predicted that the global fuel cell market size is estimated to reach $40.0 billion in 2022. Besides the market potential, FCs are a vital component for the successful implementation of the hydrogen economy and comprise an important initiative of EU's Strategic Energy Technology Plan (SET Plan) (COM 2007) towards the change of the energy system.

In general a fuel cell is a versatile and efficient electricity generation system that can be applied in a wide range of industries—from vehicles and primary energy systems to autonomous back-up power systems and portable consumer electronics devices. Although there are several types of fuel cells, all of them are structured around a central design, which includes two electrodes, named anode and cathode, and between them there is an electrolyte that facilitates the movement of the electrons. This electrolyte can be either solid or liquid and carries the ions between the electrodes. Also, in order to enhance and accelerate the reaction a catalyst is used. The electrodes and the electrolyte constitute the basic structure of a fuel cell and when a number of individual cells are composed they form the fuel cell stack. A fuel cell system has a number of subsystems and peripherals that control the flows of fuel and oxidant, the produced power, water and heat. However even though all types of fuel cells have the same structure, each one is suitable for different applications, uses different materials and requires specific fuel. Each fuel cell type also has its own operational characteristics, which makes them a very versatile and flexible technology.

14.1.1 Polymer Electrolyte Membrane (PEM) Fuel Cell

Among the various FC types, Polymer Electrolyte Membrane (PEM) FCs have some very appealing characteristics. Their low operating temperature, size and weight are only few of the features that make them more suitable for use in vehicles and portable devices. Thus, they have an increased role compared to other fuel cell types, derived by the fact that they are selected for a number of applications that

currently are in an early market entrance stage (material handling equipment—MHE) or are expected to enter the market in the near future (such as FCEVs). Also, PEMFCs can be widely used in a number of small portable devices or small stationary applications to provide primary power (telecommunication stations) or backup power (autonomous power units).

PEMFCs dominate the market in terms of shipments per year (Fuel Cell Today 2012), as their flexibility makes them appropriate for a range of markets, such as the transport sector, small stationary power applications and lately consumer electronics. The main features of a PEMFC is its fast response, the low operating temperature, high efficiency and low corrosion (Stefanopoulou and Suh 2007). Overall the operation of a fuel cell involves various phenomena related to electrochemical reactions, mass and heat transport. Therefore it is necessary to design and develop a control system able to handle a number of issues regarding power management, in the context of a safe operation which is achieved by proper fuel/air delivery and temperature control. To achieve this objective, more effectively, the system and material developments must be complemented by computer-aided modeling, control and field demonstrations. Moreover, the need for control strategies is regarded as an important component of the Balance of Plant (BOP) for an integrated fuel cell system (Bavarian et al. 2010).

14.1.2 Key Challenges and Issues of PEMFCs System Behavior

The development and deployment of optimum control and energy management is necessary as most of the available fuel cell systems contain only basic controllers with no or little regard towards an optimum state of operation. Thus, the full potential of the system is not currently utilized. Based on that, the main objectives that are targeted by this chapter are to:

- Accurately capture the behavior of the system through the development of a dynamic and detailed fuel cell mathematical model.
- Improve the performance of the control method by exploiting features of the dynamic optimization problem and the development of new algorithms.
- Design and implement an integrated industrial automation system able to incorporate advanced model predictive controllers.
- Demonstrate the applicability and efficiency of the newly developed algorithms and tools in the operation of an experimental PEM fuel cell unit.
- Monitor and evaluate the system's performance at real time.

In order to achieve these objectives various interdisciplinary actions were necessary, related to the analysis, design and development of a fuel cell system along with advanced model-based control techniques. Furthermore, emphasis is placed on the structure of the optimization problem which is solved online at each time

interval. The latter objective is realized through the development of a framework with the following desired features:

- Fast calculation of the optimal control actions while taking into account the physical and operating constraints.
- Flexibility to adapt to changing fuel cell response under the influence of disturbances or during start-up and shutdown.
- Incorporation of a multitude of performance criteria under strict computational time demands.
- Easily deployable and maintainable control and optimization solutions.

The issues that are addressed in this chapter are related to the fuel cell process into consideration while the challenges that are tackled by this work are related to advanced control issues.

14.1.3 Importance of Control to Process/System Behavior

Continuous research efforts are imperative for the realization of system-level goals for performance, predictability, stability, and other properties through appropriate analysis, design and implementation. In this context, the proper control structure and methods can function as a catalyst that transforms technological innovation to systems engineering and process novelties. Overall, control engineering provides the scientific foundation and technology for dynamically evolving systems by integrating concepts from computer science, mathematics, and systems engineering. The impact of control technology is evident in a wide range of application areas, including fuel cells, as it is the necessary facilitator for achieving desired objectives and fulfilling application-specific goals (CSS 2011). Fuel cell systems exhibit fast dynamics, nonlinearities and uncertainties that constitute challenges requiring appropriate control in order to be confronted effectively. The use of efficient control strategies would not only increase the performance of these systems, but would increase the number of operational hours as their lifetime is preserved by operating at optimal levels and also reduce the cost per produced kilowatt-hour. Overall control can be considered as a key enabling technology for the deployment of fuel cell systems as well as renewable energy systems.

More specifically, during their operation various phenomena are evolving and their behavior is affected by many variables such as temperature, partial pressures, gas utilization and humidity. Therefore, it is necessary to be able to understand qualitatively and predict quantitatively the behavior of an integrated fuel cell system in order to protect its longevity and preserve its long-term performance. Driven by this motivation their optimum operation is of great importance. Thus, it is imperative to develop appropriate control strategies and algorithms that optimize their response so that they can accomplish certain intended functions and utilize the available resources, e.g. consumption of fuel, in an efficient manner and satisfy operating and physical constraints.

14.1.4 Development Methodology—From Design to Implementation

Although the selection of the control strategy is very important, it is only one of the various steps that are necessary to achieve the desired outcome, which is a fully functional system for a specific process. These steps constitute a generic development procedure (Fig. 14.1) and they are categorized into two main sets of actions related to research and development and system engineering.

The first set is developed using a top-down approach while the later a bottom-up approach. Initially the objective is to specify a well determined set of requirements and specifications that will be used throughout the entire procedure. Also, it is important to select the appropriate topology and tools as by their integration the system structure will result. Finally, a decisive element to the success of the whole procedure is the automation and control system, since it constitutes the interface between the process and the rest of the world. Figure 14.1 illustrates a procedure where an Advanced Process Control (APC) methodology is present and more specifically Model Predictive Control (MPC). This procedure involves the development of a model of the process and a simulation study that explores the response of the developed model prior to the development of the MPC controller which is based on the dynamic model.

The necessary actions for the analysis and design of the control system are:

- Determine the requirements and the boundaries of the system derived by the process operating specifications.
- Select appropriate topology where the various components will be placed and also define the interactions between the components and the subsystem of the process.

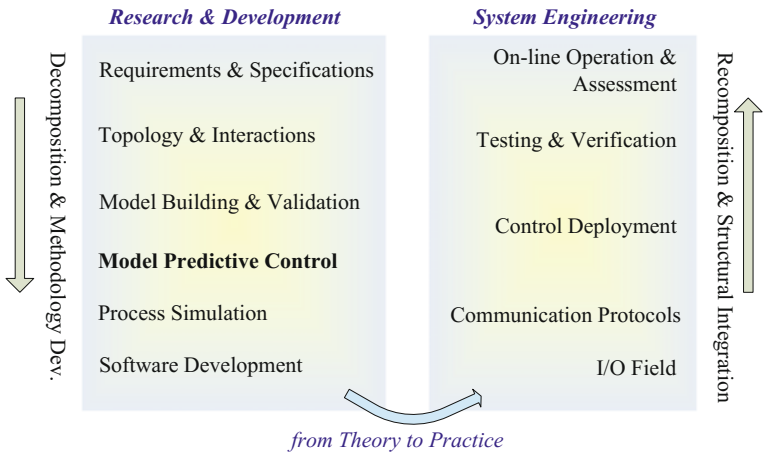

Fig. 14.1 Development procedure

- Build a dynamic mathematical model of the process and validate its behavior against either experimental data supplied by the manufacturer or from the literature.
- Analyze and design the control architecture based on the requirements and the control objectives of the process. When feasible use rapid prototyping tools to speed up the procedure.
- Simulate the behavior and response of the process.
- Develop the software that will be used by the automation system in order to deploy the selected advanced controller to the process.

Besides the definition of the boundaries and structure of the system, the outcome of these actions is a software platform able to communicate with the rest of the system according to the predefined specifications of the process operation. After these actions the rest of the procedure involves a number of steps which are necessary to develop and implement the selected topology and control structure:

- Setup the Input/Output (I/O) field by connecting the sensors and actuators of the process to the selected data acquisition system and configure the equipment in order to communicate using the appropriate industrial protocols.
- Deploy the control infrastructure to the process and the software installation provided by the automation system's vendors.
- Perform a set of tests to verify the control, including the signal tracing.
- Commission the process for nominal operation.
- Assess the behavior of the process after an initial period of operation and tune the control loops if necessary.

The fine tuning of the control system is of high importance since the overall response of the process depends on it. In the aforementioned procedure a number of challenges like computational issues, uncertainties and communication problems are of major concern that are confronted during software and hardware implementation. The ultimate aim is to select, according to needs of the system, a flexible, expandable, efficient and well defined control structure.

In order to study the behavior of fuel cell systems and demonstrate the potential of the advanced controllers a small-scale experimental process unit is designed and constructed (Sect. 14.2). Also, the automation infrastructure and the architecture of the Supervisory Control and Data Acquisition (SCADA) system is presented which is used as a platform for the verification of a number of advanced controllers. Furthermore, a dynamic nonlinear mathematical model (Sect. 14.3) is briefly analyzed that describes the behavior of the PEMFC which is experimentally validated using a formal systematic estimation procedure for the determination of the empirical parameters. Finally a thorough analysis of model-based predictive controllers (Sect. 14.4) is included and the results from the online testing to the experimental unit is described (Sect. 14.5).

14.2 Small-Scale Automated PEM Fuel Cell Unit

The scope of Sect. 14.2 is to present the overall design of the system and to specify the operation requirements and the technical features that are necessary for the implementation of an integrated supervisory control framework for a small scale automated PEM fuel cell unit. Furthermore, the structure of the automation system with its control topology is outlined and the architecture of the software that was utilized is presented. In the case of the PEM fuel cell unit and its automation system the main requirements are:

- Ability to test various fuel cells and stacks.
- Modular and scalable architecture (e.g. bypass of hydrators).
- Modifiable operating conditions related to temperature, humidity and pressure.
- Consistent startup and safe shutdown procedures.
- Supervisory monitoring and data archiving capabilities.
- Flexible automation framework able to incorporate different model-based control schemes.

Considering these requirements a small scale fully automated plant is designed and constructed at the laboratory of Process Systems Design and Implementation (PSDI) at CPERI/CERTH. The unit is able to measure all the necessary input signals, control the appropriate variables and adjust several system parameters. A dedicated hardware for the data acquisition and system control of the unit is configured and successfully integrated with the unit's components.

14.2.1 Experimental Setup of the Unit

Overall there are five distinct subsystems, the power, the gas supply, the temperature, the pressure and the water management subsystem. However, each subsystem interacts with the others although it has its own control objective. An integrated fuel cell system is equipped with various electrochemical and electronic components which are combined to form an integrated unit. Figure 14.2 illustrates the simplified process and instrumentation diagram (P & ID) of the unit.

A number of sensors measure a variety of signals from the unit along with the respective actuators that control each of the unit's subsystems. A subset of these signals is of interest for the control problem formulation and they are considered as variables at the APC framework which is developed in the subsequent sections. These signals are categorized into analog input and analog output. The available analog input measurements as shown in Fig. 14.2 are:

- The inlet flows of air and hydrogen ($\dot{m}_{air,\,in}, \dot{m}_{H2,\,in}$),
- The temperature of the hydrators ($T_{h,\,ca}, T_{h,\,an}$),
- The line temperature, before and after the fuel cell ($T_{ca,\,in}, T_{an,\,in}, T_{ca,\,out}, T_{an,\,out}$),
- The fuel cell temperature (T_{fc}),

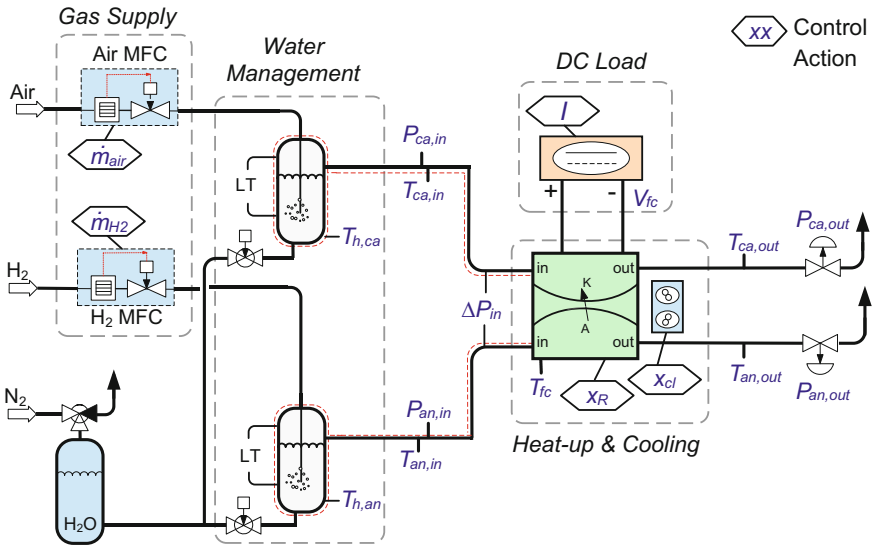

Fig. 14.2 Analog input and output variables with respect to the unit flowsheet

- The inlet/outlet pressure at the anode/ cathode ($P_{ca,in}, P_{an,in}, P_{ca,out}, P_{an,out}$),
- The pressure difference between the anode and the cathode (ΔP),
- The DC load which provides the measurement of fuel cell voltage (V_{fc}).

Furthermore, a number of analog output signals exist in the unit:

- Flow of the gases ($\dot{m}_{air}, \dot{m}_{H2}$),
- Heat-up and the cooling percentage (x_{ht}, x_{cl}),
- Current (I_{fc}) or the voltage (V_{fc}) applied by the DC electronic load depending on the mode of operation,
- Percentage of operation for the temperature of the hydrators and the heated lines ($x_{ht,an}, x_{ht,ca}, x_{\ln,an}, x_{\ln,ca}$).

Finally, there are a number of digital signals (with Boolean state, open/close, high/low) related to the electro-valves of the unit, e.g. there are valves after the MFCs, at the top and at the bottom of the hydrators, valves at the inlet and at the outlet of the fuel cell and indication of the status level of the water at the hydrators (high/low).

14.2.2 Automation System—Supervisory Control and Data Acquisition (SCADA)

Based on the aforementioned requirements the online monitoring of the unit and the flexible control structure is a prerequisite that the automation system must fulfill.

This objective is accomplished by the use of an industrial Supervisory Control and Data Acquisition (SCADA) system. Although there are various alternatives for the selection of the automation system, such as simple data acquisition systems or PLCs, the most appropriate solution that can accomplish the requirements of this chapter, is a SCADA system. This choice was made based on the SCADA's features in conjunction with the operation objectives of the fuel cell unit. The automation system of the developed unit is based on the industrial platform Proficy iFIX from General Electric. The main features that made this type of system attractive for the evaluation of the behavior of PEMFCs are:

- Open architecture able to incorporate a wide range of process equipment;
- Ability to send data to information systems for archiving purposes;
- Communicate with heterogeneous software platforms based on OPC protocol;
- Supervisory functions that support rapid software prototyping;

Apart from these interesting features the selected type of system is implemented on a computer-based architecture which means that the software platform is installed, developed and parameterized on a PC. The use of a computer-based system has many advantages as it allows the use of a common place for software development including I/O drivers and programming platforms such as Matlab, gPROMS or simple programming languages such as Fortran. Particular emphasis was placed on being able to integrate advanced process controllers.

14.2.3 Interactive Monitoring—Human Machine Interface (HMI)

All the system components (mass flow controllers, heaters, valves, etc.) are controlled by digital commands which are initiated by an interactive monitoring interface which is connected with the server part of the SCADA. This monitoring interface constitutes the client part of the SCADA and it is responsible for presenting process data to the user by interacting with the database system of the server part and it is often referred as the Human Machine Interface (HMI). A representative screen of the HMI for the monitoring of the fuel cell unit and the control of the various set-points is illustrated at Fig. 14.3.

In a system that works exclusively with sensors, controls, and process hardware, the HMI provides a graphical representation of the overall process. The development of user friendly graphical interfaces enables us to monitor, supervise, apply control actions to the system and be aware of any alarm conditions. On these interfaces real-time data are presented. This graphic representation is an important feature as it minimizes the deployment cycle compared to traditional programming of user interfaces.

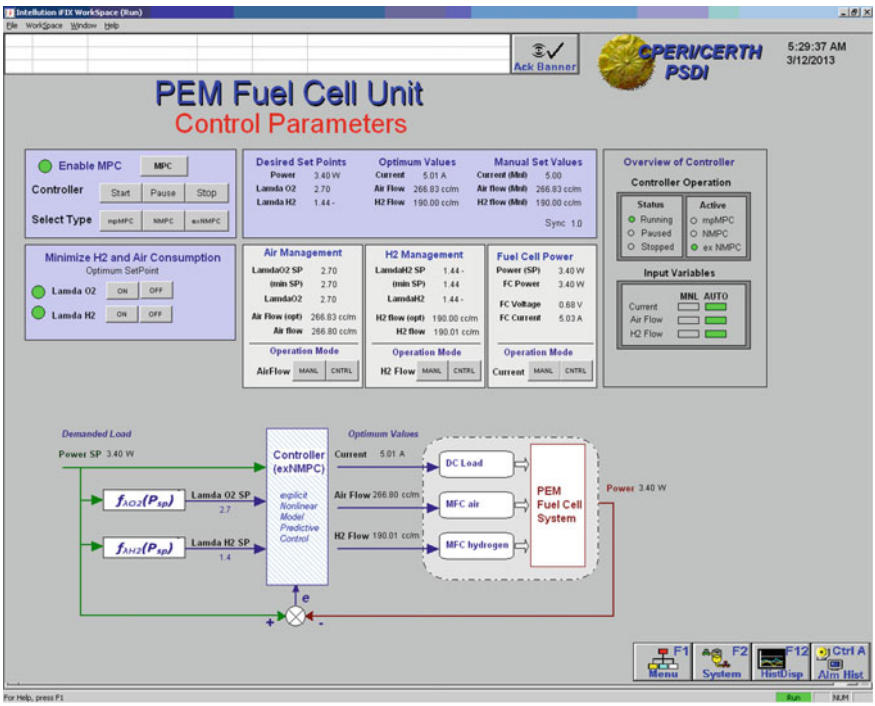

Fig. 14.3 Indicative HMI of the control parameters and MPC activation at the PEMFC system

14.2.4 Flow of Information and Archiving

Overall the control framework used in this work, has a number of interacting entities that form a common computer-aided platform for the monitoring and control of the PEM fuel cell unit. The path of the process measurements from all the devices starts from the sensor signaling and finishes at the archiving system. This information flow, from the I/O field to the end user, is shown in Fig. 14.4. Furthermore Fig. 14.4 illustrates the interconnection of the control system's entities and the I/O field including the conceptual flow of information from the signal acquisition to the final presentation at the user level through the development of a graphical interface.

Besides the control and the data acquisition, it is important to be able to monitor the long-term system behavior. A comprehensive way to view the evolution of each experiment and the response of each subsystem is the use of diagrams based on historically acquired data. In our case a Process Information Management System (PIMS) from OSISoft is used in order to archive the online data of the PEM fuel cell unit through the SCADA. The use of a PIMS increases the visualization of the unit's behavior and enhances the decision making at the supervisory level.

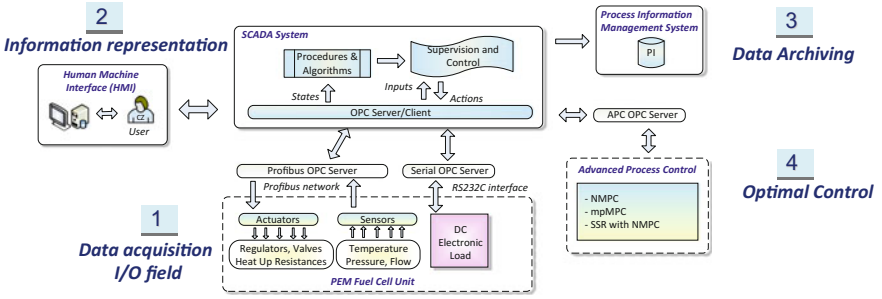

Fig. 14.4 Computer-aided framework for the integrated PEM fuel cell unit and flow if information

Although the aforementioned application involves a specific system, the computer-aided platform is a generic one and can be applied to:

- Other PEM fuel cells by adjusting the physical parameters of the mathematical model and by determining the empirical parameters based on the developed estimation procedure.
- Various process systems using respective linear or nonlinear models while keeping the structure of the MPC methodology.
- Other industrial automation systems due to its open architecture and interoperability features.

Furthermore, the middleware that interconnects the MPC-based controllers to the automation system can be used for other type of controllers (e.g. adaptive controllers) or software platforms as long as they have OPC connectivity.

14.3 Control Objectives and Modeling of an Integrated Fuel Cell System

As stated earlier the durability and performance of a PEM fuel cell are influenced by the operating conditions. Therefore, it is of vital importance to control the various subsystems responsible for maintaining a stable operating environment while ensuring an economically attractive operation. Overall the objectives for the control system are to effectively address the issue of power generation in an optimum manner. In this context the optimality is defined by the following three terms:

- Operation at a safe region regardless of the load fluctuations.
- Minimization of the fuel consumption and air supply.
- Maintenance of stable temperature conditions ensuring proper gas humidification.

One of the most important considerations for the control of the fuel cell is to guarantee that the operation is within a safe region which is expressed by avoiding fuel and oxidant starvation. It is important to prevent such phenomenon as it affects the longevity of the fuel cell and can cause irreversible damage to the membrane (Schmittinger and Vahidi 2008). Moreover this is very critical when abrupt changes on the load occur. Such case can cause a sudden increase for gas supply demand since the oxygen and hydrogen react instantaneously and the level of the gases are reduced drastically.

From the fuel point of view, it is desirable to supply the amount of hydrogen close to the required one. When a recycling line is not present at the fuel cell unit the unreacted hydrogen will be released to the vent. Therefore, one of the control objectives is to minimize the supplied hydrogen in conjunction with the previous objective for safe operation.

Another significant factor that should be considered is the proper handling of the operating temperature, as it affects the long-term performance of the fuel cell. The operation at an elevated temperature accelerates the degradation phenomena and can influence the durability of the fuel cell. Furthermore, the temperature directly affects the rate of chemical reactions and transport of vapor and reactants (Ahn and Choe 2008). As the hydrators, used for the humidification of the gases, operate at a stable temperature point, it is important to maintain the fuel cell operating temperature, in order to keep the water content of the gases at a desired level.

14.3.1 Variables and Control Configuration

In the considered system there are four distinct control objectives, one direct external and three indirect internal objectives. The main control objective is to fulfill the varying power demand at acceptable response time, avoiding oxygen starvation while minimizing the hydrogen consumption at stable temperature conditions. Once the environment of fuel cells has been clearly determined, it is necessary to identify the controlled and corresponding manipulated variables which will be used to device the appropriate control scheme that make the system to respond as requested. More specifically the desired power (P_{sp}) is delivered by properly manipulating the current (I) which is applied to the fuel cell by the converter (DC electronic load) connected to the system. In our case the mode of operation for the electronic load is set to constant current (CC) since the boundaries of the system were identified by the experimental study performed during the activation of the fuel cell.

The safe operation is maintained by controlling the reactants at a certain excess ratio level in order to avoid starvation caused by sub-stochiometric reaction conditions at the cathode and the anode. The safe operating region for the cathode and the anode is defined by two unmeasured variables, the oxygen and hydrogen excess ratios (λ_{O2}, λ_{H2}), expressed as the ratios of the input flow of each gas to the consumed quantities per unit time due to the reaction (Pukrushpan et al. 2004):

$$\lambda_{o2} = \frac{\dot{m}_{o2,\ cach,\ in}}{\dot{m}_{o2,\ caGDL}} \tag{14.1}$$

$$\lambda_{H2} = \frac{\dot{m}_{H2,\ anch,\ in}}{\dot{m}_{H2,\ anGDL}} \tag{14.2}$$

where $\dot{m}_{o2,\,cach,\,in}, \dot{m}_{H2,\,anch,\,in}$ are the oxygen and hydrogen input flows at the channels while $\dot{m}_{o2,\ caGDL}, \dot{m}_{H2,\ anGDL}$ are the respective reacted quantities. In order to reach the required excess ratio set-point the air and hydrogen flows ($\dot{m}_{air}, \dot{m}_{H2}$) are used as manipulated variables. The safety of the operation is ensured by maintaining the excess ratios above one ($\lambda_{O2,SP} > 1$, $\lambda_{H2,SP} > 1$).

The power generation and the starvation avoidance objective can be achieved by control actions that aim at an accurate set-point tracking of P_{SP}, $\lambda_{O2,\ SP}$, $\lambda_{H2,SP}$. On the other hand the temperature control ($T_{fc,SP}$) involves two mutually exclusive subsystems, one for the heat-up and another for the cooling; therefore a number of requirements are considered besides the set-point tracking ability of the controller:

- Maintain the temperature at the desired level having an acceptable deviation
- Avoid concurrent operation of the heat-up and cooling
- Exhibit stable operation
- Avoid large overshoot after step changes

The fuel cell temperature is controlled by manipulating the operating percentage of the heating resistance (x_{ht}) and of the cooling fans (x_{cl}) of the system, respectively. Based on these variables the resulted conceptual control configuration is shown by Fig. 14.5.

Figure 14.5 illustrates the entities of the system and the flow of information related to the controller and the fuel cell including the measured variables from the unit.

14.3.2 Dynamic Nonlinear Semi-empirical

The behavior of the aforementioned unit is described by an experimentally validated nonlinear dynamic model that couples a set of first principles equations with a

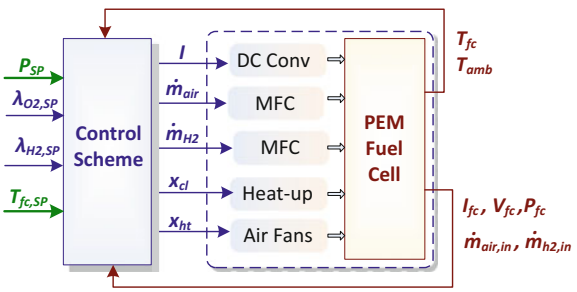

Fig. 14.5 Generic control configuration for the PEMFC unit

set of semi-empirical equations. For reasons of completeness a brief description of the main equations is presented here while a detailed analysis can be found in our previous work (Ziogou et al. 2011b). The mass balances at the cathode's gas flow channel are:

$$\frac{dm_{o2,\,cach}}{dt} = \dot{m}_{o2,\,cach,\,in} - \dot{m}_{o2,\,cach,\,out} - M_{o2}\frac{I}{4F} \tag{14.3}$$

$$\frac{dm_{N2,cach}}{dt} = \dot{m}_{N2,\,cach,\,in} - \dot{m}_{N2,\,cach,\,out} \tag{14.4}$$

$$\frac{dm_{v,\,cach}}{dt} = \dot{m}_{v,\,cach,\,in} - \dot{m}_{v,\,cach,\,out} + \dot{m}_{evap,\,cach} + A_{fc}M_{H2O}N_{v,k} \tag{14.5}$$

where $m_{o2,\,cach}, m_{N2,cach}, m_{v,cach}$ are the oxygen, nitrogen and water vapor masses at the cathode, $\dot{m}_{k,\,cach,\,in}, \dot{m}_{k,\,cach,\,out}, k = [O_2, v]$, are the input and output mass flows, $\dot{m}_{evap,\,cach}$ is the rate of evaporation, $N_{v,ca}$ is the vapor molar fluxes between GDL and the cathode, M_{O2}, M_{H2O} are the oxygen and water molar mass, I is the current, A_{fc} is the membrane active area and F is the Faraday number. Similarly the mass balances for the anode channel are derived. The partial pressure of water vapor in the cathode GDL ($p_{v,caGDL}$) satisfies the respective mass balance equation:

$$\frac{dp_{v,\,caGDL}}{dt} = RT_{fc}\left(\frac{I}{2FA_{fc}} + N_{v,\,mem} - N_{v,\,ca}\right)/\delta_{GDL} \tag{14.6}$$

where δ_{GDL} is the thickness of the diffusion layers and N_{mem} is the vapor molar flux at the membrane affected by the electro-osmotic drag ($N_{v,\,osm}$) and the back diffusion ($N_{v,\,diff}$) between the cathode and the anode:

$$N_{v,\,mem} = N_{v,\,osm} - N_{v,\,diff} \tag{14.7}$$

The dynamics of the temperature (T_{fc}) results from the overall energy balance equation of the fuel cell:

$$m_{fc}Cp_{fc}\frac{dT_{fc}}{dt} = \Delta\dot{H}_{an} + \Delta\dot{H}_{ca} + \Delta\dot{H}_{chem} - \dot{Q}_{rad} - \dot{Q}_{amb} + \dot{Q}_R - \dot{Q}_{cl} - P_{elec} \tag{14.8}$$

where m_{fc} denotes the mass of the fuel cell and Cp_{fc} is the specific heat calculated for the system into consideration. The above equation takes into account the differences of the energy flow rates between the input and output streams at the anode ($\Delta\dot{H}_{an}$) and the cathode ($\Delta\dot{H}_{ca}$), the rate of energy produced by the chemical reaction ($\Delta\dot{H}_{chem}$), the rate of energy which is released to the environment through radiation ($\dot{Q}_{rad}$) and the rate of heat losses to the environment ($\dot{Q}_{amb}$). Also the heat supplied by the heating resistance ($\dot{Q}_R$) is included along with the heat which is

removed by the air cooling system ($\dot{Q}_{cl}$). The last term P_{elec} is the amount of energy which is converted to electrical power.

To determine the voltage (V_{fc}) and subsequently the produced power of the fuel cell, the following equations are used:

$$V_{fc} = E_{nerst} - V_{act} - V_{ohm} - V_{conc} \tag{14.9}$$

$$V_{act} = \xi_1 + \xi_2 T_{fc} + \xi_3 T_{fc} \ln(I) \ + \xi_4 T_{fc} \ln(c_{O2}) \tag{14.9a}$$

$$V_{ohm} = (\xi_5 + \xi_6 T_{fc} + \xi_7 I) I \tag{14.9b}$$

$$V_{conc} = \xi_8 \exp(\xi_9 I) \tag{14.9c}$$

where E_{nerst} is the ideal Nernst voltage, V_{act}, V_{ohm}, V_{conc} are the activation, ohmic and concentration losses and c_{o2} is the oxygen concentration. In (14.9a), (14.9b), (14.9c) $\xi_k, k = 1..9$ represent experimentally defined parametric coefficients (Ziogou et al. 2011b). This nonlinear dynamic model is used at the core of the proposed advanced model-based control scheme which is analyzed in Sect. 14.4.

The described dynamic model is not system dependent as it can be adjusted to describe any other PEM fuel cell system by performing a number of sequentially executed actions. These actions are related to the determination of the physical characteristics of the PEM fuel cell, the unit specific parameters and finally the empirical parameters utilized by the electrical subsystem of the model. The value of the empirical parameters are determined by a systematic parameter estimation procedure as long as experimental data or data from the manufacturer are provided. Finally the resulted estimated values are included to the model and a different set of data is necessary to explore the validity of the model against the fuel cell unit.

14.4 Advanced Model-Based Control Strategies

The objective of control is to achieve a set of predefined conditions for the process and maintain the operation at the desired or optimal values (Qin and Badgwell 2003). In general process control refers to the technologies that are necessary to design, develop and implement control systems for a process. Advanced process control methodologies are applied to an even wider range of processes and applications as over the past few years there were significant improvements to the control theory and the computational requirements of such methodologies (Bauer and Craig 2008). This section presents two advanced model-based control methodologies for the efficient real-time control of PEM fuel cell and a new method that combined features from the main advanced MPC approaches.

14.4.1 Model Predictive Control

Model predictive control (MPC) also known as receding horizon control (RHC), is becoming a preferred control strategy for a large number of industrial processes. The main reasons for the increased popularity include the ability to explicitly handle constraints and dynamics of the system. Moreover, the use of MPC can simultaneously consider economic and operating objectives based on an optimizing strategy. Nevertheless, despite its well-known acceptance both in a theoretical and industrial level, the various variants of MPC are still subject of in-depth analysis and evaluation. Especially, nonlinear MPC is currently in a top position in the evolving area of advanced control as its use can lead to more accurate results, increased robustness and flexibility in system development, compared to conventional approaches.

In general MPC refers to a methodology which makes explicit use of a process model to optimize the future predicted behavior of a process. The main objective is to obtain control actions that minimize a cost function related to selected objectives or performance indices of the system. At each sampling time an optimal control problem is solved using measurements acquired from the system and it yields the appropriate control inputs for the system.

The main objective is to obtain a control action minimizing a cost function related to selected objectives or performance indexes of the system. At each sampling time a finite horizon optimal control problem is solved over a prediction horizon (T_p), using the current state of the process as the initial state. The optimization yields an optimal control sequence ($u_k..u_{k+Np}$), and only the first control action (u_k) for the current time is applied to the system while the rest of the calculated sequence is discarded. At the next time instant the horizon is shifted by one sampling interval and the optimization is restarted with the information of the new measurements acquired from the system. The structure of the MPC framework is presented at Fig. 14.6 while the concept of the receding horizon is illustrated at Fig. 14.7.

MPC is very flexible regarding its implementation and can be deployed and adjusted to a wide range of systems since its development is based on a model of the system. Nevertheless, despite its well-known acceptance MPC is still subject to

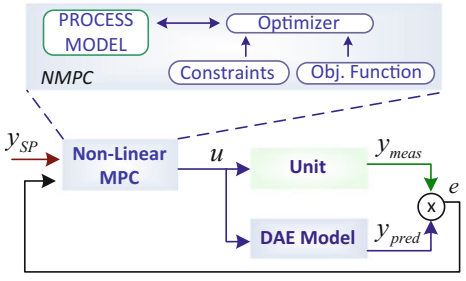

Fig. 14.6 Model predictive control structure

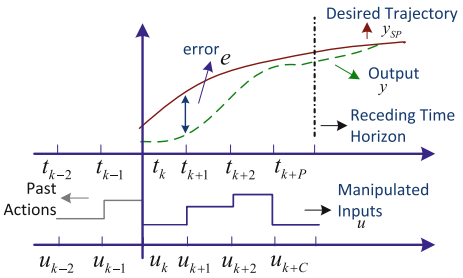

Fig. 14.7 Receding horizon control concept

in-depth analysis and evaluation. Especially, nonlinear MPC is currently in a top position in the evolving area of advanced control as its use can lead to more accurate results, increased robustness and flexibility in hardware development, compared to the linear case of MPC and other conventional approaches. In this chapter MPC will be used as the selected method of APC that will be deployed to the PEM fuel cell unit under consideration.

14.4.2 Nonlinear Model-Predictive Control

Motivated by the impact of MPC and urged by the need for near boundary operation and highly nonlinear behavior of many processes Nonlinear MPC (NMPC) has gained significant attention over the past decade (Findeisen et al. 2007; Magni et al. 2009). Moreover recent advances in optimization enable the move towards direct online optimizing control (Engell 2007). Various theoretical and practical aspects (e.g. stability, reliability, robustness, computational burden) have been recently explored and on-going research is progressing in the area of NMPC towards the industrial implementation of the methodology (Rawlings and Mayne 2009). The NMPC formulation requires the online solution at each time step of an optimization problem to determine the manipulated inputs. The optimization problem derived by the control formulation generally is nonconvex and consequently, the major practical challenge associated with NMPC is the online solution of the nonlinear program (NLP). The potential and benefits of NMPC are progressively penetrating to complex processes and furthermore research is encouraged.

14.4.2.1 Dynamic Constrained Optimization

As stated earlier the NMPC formulation includes the solution of an optimization problem at each sampling instance. But the online application of the NMPC framework faces a challenging dilemma (Diehl et al. 2002), either the nonlinear iteration procedure is performed until a pre-specified convergence criterion is met, which might introduce considerable feedback delays, or the procedure is stopped

prematurely with only an approximate solution, so that a pre-specified computation time limit can be met. Recently NMPC controllers are based on nonlinear programming (NLP) sensitivity with reduced online computational costs and can lead to significantly improved performance (Zavala and Biegler 2009). Overall the application of dynamic optimization in conjunction with fast optimization solvers allows the use of first-principles models for NMPC (Diehl et al. 2009). In general, a DAE constrained optimization problem is considered which includes the continuous-time counterpart of the NMPC problem:

$$\min_{u(t)} J = \varphi(x(t), z(t), u(t)) \tag{14.10}$$

$$\text{s.t. } \frac{dx(t)}{dt} = f_d(u(t), x(t), z(t)) \tag{14.10a}$$

$$0 = f_d(u(t), x(t), z(t)) \tag{14.10b}$$

$$x(0) = x_0 \tag{14.10c}$$

$$x^L \le x(t) \le x^U,\ z^L \le z(t) \le z^U, u^L \le u(t) \le u^U \tag{14.10d}$$

where t is the scalar independent dimension defined in the fixed domain [0, t_f], x is the vector of differential (state) variables, z is a vector of algebraic variables, u is the vector of manipulated variables and x_0 are the initial conditions of the state variables, f_d and f_a are the differential and the algebraic equations. Finally Eq. (14.10d) denotes the bounds. Since the implemented NMPC algorithm involves inequality constraints, direct optimization methods are used for the optimization problem which is transformed into a NLP problem.

In principle two main numerical approaches exist for the solution of the open-loop optimal control problem, the indirect and the direct methods (Biegler and Grossman 2004). The first relies on Pontryagin's Maximum Principle, it is suitable for problems that have only equality constraints. The second approach, the direct approach, transforms the optimization problem into an NLP problem. Direct methods can handle inequality constraints and find suitable initial guesses for state variables. Mainly two direct strategies exist, the sequential and the simultaneous. In the sequential or single-shooting approach the control trajectory of the manipulated variables is finitely parameterized or discretized, the optimization steps are performed and afterwards the NLP problem is solved. In each evaluation of the performance index at the solution of the NLP, the process model is integrated with a DAE solver that integrates the model over the entire horizon in a single call.

Simultaneous Approach

In the direct simultaneous approach the solution of the differential equations and the optimization is obtained concurrently. For this purpose the differential equations are

discretized and enter the optimization problem as additional constraints. Typical simultaneous approaches use multiple shooting or direct transcription to parameterize/discretized the equations. The direct multiple shooting approach discretizes the optimization horizon (prediction horizon) into a number of elements (time intervals) with discretization of the manipulated variables.

The direct transcription method explicitly discretizes all the variables (differential, algebraic, input and output) and generates a large scale but sparse NLP problem. This discretization is based on orthogonal collocation on finite elements (OCFE), which can be treated as a special calls of implicit Runge-Kutta type method (Betts 2001). Initial values for the whole state trajectory are required which may be an advantage if such knowledge is available. The direct transcription method is usually selected due to its accuracy and numerical stability properties (Diehl et al. 2002; Kameswaram and Biegler 2008). In addition, this approach does not require a DAE integrator, as the discretized model is solved once at the optimal point. The overall procedure for the development of an NMPC controller based on direct transcription method and the various conceptual stages for from the design to the online deployment are presented in Fig. 14.8.

As this is a modular design the place of the direct transcription method could take any other optimization method (single shooting or multiple shooting). The time horizon is divided in finite elements (*NE*) and each equally spaced finite element is partitioned in collocation points (N_{cop}). The residuals of state and algebraic equations are assumed to be exactly satisfied only at the collocation points. The position of the collocation points is determined as the shifted roots of orthogonal polynomials, usually Legendre or Radau orthogonal polynomials. The solution is approximated with Lagrange polynomials at each finite element (Finlayson 1992; Biegler et al. 2002):

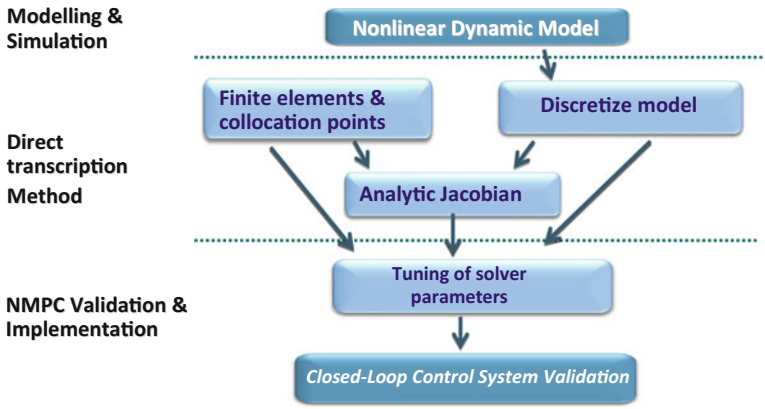

Fig. 14.8 Development procedure of NMPC framework

$$x(t) \approx \sum_{j=0}^{Ncop} x^{i,j}\Omega_j(t),\ i = 1..NE,\ j = 0..N_{cop},\ t \in [t_i, t_{i+1}] \tag{14.11}$$

$$\Omega_j(t) = \prod_{k=0, k\neq j}^{Ncop} \frac{(t - t_{i,k})}{(t_{i,j} - t_{i,k})} \tag{14.12}$$

where N_{cop} is the total number of the internal collocation points of each element, NE is the number of the finite elements, $x^{i,j}$ is the value of the state vector at collocation point j of the *ith* finite element. Respectively the algebraic variables ($z^{i,j}$) and input (manipulated variables) variables (u^i) are approximated. The length of each element is $h_i = t_i - t_{i-1}$. The basis function (Ω) is normalized over each element having time $\tau \in [0, 1]$ and $t = t_{i-1} + h_i \cdot \tau$. Ω_j is calculated using the shifted roots of the Legendre polynomials. After the discretization of the DAE model, the constrained optimization problem (14.10) is expressed as an NLP problem in the form (Biegler et al. 2002):

$$\min_{x^{i,j}, z^{i,j}, u^i} \sum_{i=1}^{NE} \sum_{j=1}^{Ncop} w_{i,j}\phi\left(x^{i,j}, z^{i,j}, u^i\right), \tag{14.13}$$

$$\text{s.t.}: \sum_{k=0}^{Ncop} \Omega_k(\tau_{i,j}) x^{i,k} = h_i f_d\left(u^i, x^{i,j}, z^{i,j}\right) \tag{14.13a}$$

$$0 = f_a\left(u^i, x^{i,j}, z^{i,j}\right) \tag{14.13b}$$

$$x^{1,0} = x_0, \quad x(t_f) = \sum_{j=0}^{Ncop} x^{NE,j}\Omega_j(1) \tag{14.13c}$$

$$x^{i,0} = \sum_{j=0}^{Ncop} x^{i-1,j}\Omega_j(1),\ \text{i} = 2..NE \tag{14.13d}$$

$$x_l \leq x^{i,j} \leq x_u, z_l \leq z^{i,j} \leq z_u, u_l \leq u^i \leq u_u,\ i = 1..NE, \quad j = 1..N_{cop} \tag{14.13e}$$

To enforce zero-order continuity of the state variables at the element boundaries the connecting equations are used (14.13d). Overall this method exhibits fast convergence rates and can deal with unstable systems in a straightforward manner since it allows direct enforcement of state and control variable constraints (Biegler et al. 2002). In this work the simultaneous direct transcription method is selected for the NMPC framework.

14.4.3 Multi-parametric Model Predictive Control

The computation of an MPC law is derived by the solution of an optimization problem at each sampling instant. However this inserts a computational issue which must be carefully handled in order to avoid loss of performance due to delays. An alternative approach to classic MPC is the explicit or multi-parametric MPC (mpMPC) method that avoids the need for repetitive online optimization (Pistikopoulos 2012). This method is suitable for linear constrained state space system with low complexity (Bemporad et al., 2002a,b; Pistikopoulos et al. 2002). The development of an mpMPC controller is realized into two main steps:

- Off-line optimization: Derivation of the critical regions which are explored by an optimal look-up function.
- Online implementation: Based on the system measurements the critical regions are traversed and the corresponding optimal control action is determined.

In mpMPC the online optimization problem is solved off-line with multi-parametric quadratic programming (mpQP) techniques to obtain the objective function and the control actions as functions of the measured state/outputs (parameters of the process) and the regions in the state/output space where these parameters are valid i.e. as a complete map of the parameters. Online control is then applied as a result of simple function evaluations since the computational burden is shifted offline. The overall procedure of development an mpMPC controller, from design to online deployment (Pistikopoulos et al. 2007; Pistikopoulos 2012), is presented in Fig. 14.9.

The following section presents a brief overview of the main ideas, necessary for the formulation and implementation of the mpMPC.

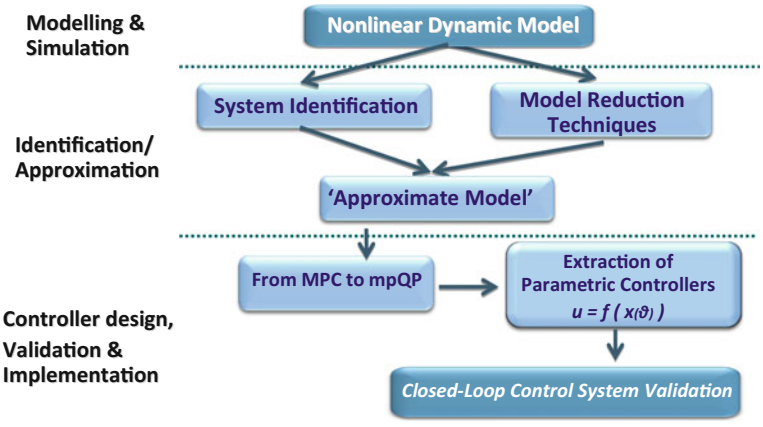

Fig. 14.9 Development procedure of mpMPC framework

14.4.3.1 Parametric Programming

Parametric programming is a generic mathematical technique that for a given objective function to optimize, a vector of optimization variables and a vector of parameters, provides the optimization variables as a set of functions of the parameters and the corresponding regions in the space of parameters, called critical regions (*CR*). Consider the following general parametric programming problem:

$$\begin{array}{ll} z(\theta) = & \min\limits_{x} f(x,\theta) \\ s.t. & g(x,\theta) \le 0 \\ x \in X \subseteq \Re^n, & \theta \in \Theta \subseteq \Re^s \end{array} \tag{14.14}$$

where x is the vector of continuous variables and θ is the vector of parameters bounded between certain upper and lower bounds. The substitution of $x(\theta)$ into $f(x, \theta)$ result to $z(\theta)$ which is the parametric profile of the objective function. When the parameter is a vector instead of a scalar we refer to multi-parametric programming. The solution of (14.14) is given by Dua et al. (2002):

$$x(\theta) = \begin{cases} x^1(\theta) & \text{if} \quad \theta \in CR^1 \\ x^2(\theta) & \text{if} \quad \theta \in CR^2 \\ \vdots & \\ x^i(\theta) & \text{if} \quad \theta \in CR^i \\ \vdots & \\ x^N(\theta) & \text{if} \quad \theta \in CR^N \end{cases}$$

such that $CR^i \cap CR^j = \varnothing$, $i \neq j$, $\forall i, j = 1,\ldots,N$ and CR^i denotes a critical region. In order to obtain the critical regions and $x^i(\theta)$ a number of algorithms have been proposed in the literature. The selection of the appropriate one depends on the nature of the problem, if f and g are convex, differentiable, linear, quadratic or nonlinear. For the formulation of the mpMPC f is convex and quadratic and g is linear. Thus, the resulting parametric approach is a multi-parametric quadratic program (mpQP). An algorithm for the solution of the mpQP has been proposed by Dua et al. (2002) where there is an iterative determination of a set of unique active constraints and the corresponding critical regions by writing the KKT optimality conditions of the mpQP.

14.4.3.2 From MPC to mpQP

The use of mpQP enables the derivation of the objective and optimization variable as functions of the varying parameters and the regions in the space of parameters where these functions are valid. Thus, the solution of the optimal control problem (MPC problem) with mpQP is a multi-parametric problem. A brief outline of the standard MPC formulation is presented along with its transformation to the mpQP

problem. Consider the following constrained discrete-time linear time invariant model of the system:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \\ y_t &= Cx_t \end{aligned} \qquad (14.15)$$

$$\text{s.t.} \quad \begin{aligned} x_{\min} \le x_t \le x_{\max} \\ u_{\min} \le u_t \le u_{\max} \\ y_{\min} \le y_t \le y_{\max} \end{aligned}$$

where $x_t \in \Re^n$ is the state vector, $u_t \in \Re^m$ is the vector of input variables, $y_t \in \Re^p$ is the vector of the output variables, $A \in \Re^{n \times n}$, $B \in \Re^{n \times m}$ and $C \in \Re^{p \times n}$ are the system matrices, t is the current time interval and the subscripts *min* and *max* denote the lower and upper bounds respectively. The receding horizon open-loop optimal control problem regulating (14.15) to the origin has the following formulation and it is used to derive the explicit control law (Mayne et al. 2000) for x_t:

$$\begin{aligned} \min_{U} \quad & J(U, x(t)) = x_{t+Ny|t}^T P x_{t+Ny|t} + \sum_{k=0}^{N_y - 1} \left[x_{t+k|t}^T Q x_{t+k|t} + u_{t+k}^T R u_{t+k} \right] \\ s.t. \quad & x_{t|t} = x(t) \quad and: \\ & x_{t+k+1|t} = A x_{t+k|t} + B u_{t+k}, k \ge 0 \\ & u_{t+k} = K x_{t+k|t}, N_u \le k \le N_y \\ & x_{\min} \le x_{t+k|t} \le x_{\max}, k = 1, \ldots\ldots, N_c \\ & u_{\min} \le u_{t+k} \le u_{\max}, k = 1, \ldots\ldots, N_c \end{aligned} \qquad (14.16)$$

where $x_{t+k|t}$ is the prediction of x_{t+k} at time t, Q and R, are the tuning parameters which are constant, symmetric and positive definite matrices, N_y, N_u and N_c are the prediction, control and constraint horizons respectively with $N_u \le N_y$ and K is some feedback gain. The weight $R\psi$ penalizes the use of control action u. The sequence $U \triangleq \left[u_t^T, \ldots., u_{t+N_u-1}^T \right]^T$ contains the future control inputs that yield the best predicted output with respect to the performance index for the prediction horizon. Once this is determined, the first control input u_t is applied to the system. Considering the linear model (14.15), the MPC problem (14.15) can be recast as a mpQP which can be solved with standard multi-parametric programming techniques and involves a systematic exploration of the parameter space. From (14.16) the following can be derived (Pistikopoulos et al. 2002):

$$x_{t+k|t} = A^k x(t) + \sum_{j=0}^{k-1} A^j B u_{t+k-1-j} \qquad (14.17)$$

The optimization problem (14.16) with the aid of (14.17) can be rewritten in a QP problem.

$$V(x(t)) = \frac{1}{2}x^T(t)Yx(t) + \min_U \left\{ \frac{1}{2}U^T HU + x^T(t)FU \right\} \tag{14.18}$$

$$s.t.\, GU \leq W + Ex(t)$$

where U is the vector of optimization variables and H, F, Y, G, W and E are obtained from Q and R.

14.4.3.3 Critical Regions and Feedback Control Law

The mpQP problem can be solved with any available QP solver or by the use of software packages like POP (POP 2007) or MPT (Kvasnica et al. 2004) and the solution is a set of convex non-overlapping polyhedra on the parameter space, each corresponding to a unique set of active constraints. The solution of the mpQP problem consists of several steps (Pistikopoulos et al. 2007):

Find a local optimum $z(x)$ by solving the QP problem for $x = x_0$ and identifying the active constraints.

- Find the set in the space of $x(t)$ (critical regions) where $z(x)$ is valid.
- Proceed iteratively until the $x(t)$-space is covered.

The mpQP (14.18) is solved by treating z as the vector of optimization variables and x_t as the vector of parameters to obtain z as a set of explicit functions of x_t. The optimizer z(x) is continuous and piecewise affine so will be U. Subsequently only the first element of U is applied and the control action $u(t)$ is also piecewise affine and continuous and it is expressed as an explicit function of the state variable $x(t)$ for the different critical regions, obtained though an affine mapping:

$$u(t) = f(t) = \begin{cases} K_1 x + c_1 & if \quad D_1 x \leq b_1 \\ & : \\ K_{Ncr} x + c_{Ncr} & if \quad D_{Ncr} x \leq b_{Ncr} \end{cases} \tag{14.19}$$

where N_{CR} is the number of critical regions, K,c,D,b are constants defining each region CR and the derived optimal control action within. The online effort is thus reduced to the evaluation of (14.19) of the current state and the determination of the region (point location problem) in which the current state x belongs.

The offline preprocessing of the optimal solution allows the implementation of mpMPC with high sampling requirements whereas the look-up function offers an easy and computationally cheap implementation. The advantages of mpMPC are widely recognized and the on-going research efforts are constantly enhancing its features and characteristics.

14.4.4 Reduced Search Space Nonlinear Model Predictive Control

The presented advanced model-based control methodologies have many advantages and some limitations that affect their applicability. As previously stated in NMPC which is very appealing due to its ability to handle dynamic nonlinearities of the process under consideration, whereas mpMPC is based on a reduced order linear model that provides the optimal solution in real-time via a simple look-up function, as the optimization problem is computed offline. However these approaches have also some limitations. The mpMPC approach can be used to linear and relatively low dimensional systems as the complexity of the solution grows with the problem size, along with the fact that the response of the controller depends heavily on the accuracy of the derived reduced order model. In the case of NMPC the main barrier for its wider applicability arises by the requirements imposed by the online solution of the optimization problem at every iteration.

During the last decade significant effort is devoted by the research community to develop algorithms and methods that could overcome the aforementioned limitations. There are several very promising works that approach these issues in a systematic way and each one focus on a specific issue of the MPC-based method. In order to address the issues related to MPC, NMPC or mpMPC various approaches are developed such as:

- Reduction of the complexity of mpMPC solutions through approximate solutions or proper merging of resulted critical regions.
- Expansion of the explicit method to nonlinear systems (mpNMPC).
- Use of real-time variants and suboptimal approaches to improve the optimization time of NMPC.
- Exploit the structure of the problem through proper algorithms implemented.

However the efforts to improve the performance and limit the drawbacks of the MPC-based controllers are not limited only to different approaches to the control problem and its structure. Recently numerous methods related to algorithmic developments have been proposed. An indicative only list includes the following:

- Use of a warm-start homotopy path method (Ferreau et al. 2008),
- A PWA approximation for warm-starting (Zeilinger et al. 2011),
- A partially reduced Sequential QP (SQP) method (Shafer et al. 2007),
- An accelerated dual gradient-projection algorithm (Patrinos and Bemporad 2012),
- New developments in interior-point methods (Wang and Boyd 2010; Domahidi et al. 2012),
- Use of advanced preprocessing (Zavala and Biegler 2009; Yand and Biegler, 2012),
- A combination of multiple shooting and direct transcription method (Tamimi and Li 2010),

- An event-driven triggering method (Eqtami et al. 2011),
- Exploiting a list with frequently used active sets (Pannocchia et al. 2007),
- A set-theoretic method based on an adaptive interpolation (Raimondo et al. 2012),

All these approaches and methods indicate that MPC is a method of interest which is gradually evolving as its advantages are widely recognized since it has a strong potential to numerous applications from small-scale low complexity to large-scale highly nonlinear systems and processes.

Based on the above considerations this section presents a novel combination of these two well established method, NMPC and mpMPC, that cooperate in a control framework which exploits their individual characteristics. The scope of the proposed integrated framework is to combine the benefits that each control methodology has, namely the accuracy and full coverage of the system's operation for the NMPC approach and the fast execution time of the mpMPC approach. Furthermore, as the basis of the NMPC is the solution of an NLP problem it is important to reduce the computational effort (Diehl et al. 2002) between successive iterations, which constitutes the primary objective of the proposed synergy. The objective of the proposed framework is mainly the reduction of the computational effort for the solution of the NLP problem between successive iterations. This is achieved by using a newly proposed preprocessing bound related technique and an existing technique related to the optimization problem:

- A Search Space Reduction (SSR) technique of the feasible space,
- A warm-start initialization procedure of the NLP solver.

The term warm start signifies that information from the previously solved optimization problem is used in order to formulate the subsequent problem. More specifically the optimal solution of a problem is provided as the initial solution of a subsequent one. This technique can significantly reduce the number of iteration towards the optimum point (Benson and Shano 2008). Therefore, it is of great importance to define a good starting point using the information gained from the previous iteration. In the proposed framework an active set method is used and thus the use of warm-start is enabled to achieve better performance.

14.4.4.1 Search Space Exploration and Region Reduction/Elimination

In global optimization there are two directions that the respective algorithms can follow in order to improve their performance and decrease to necessary time for the determination of the optimum solution:

- Reduce the number of variables and screen out the unimportant ones
- Reduce the search space by eliminating unpromising regions

Although both of them have interesting features and are useful for global optimization problems, in the case of control problems only the second one can be used, since all variables of the problems are important. The benefit of using space exploration optimization algorithms is that good and acceptable solutions can be reached with fewer resources, less computation time and better accuracy. Motivated by this idea, that the feasible space can be partitioned and only part of is may be of interest, a novel method for its reduction is discussed in this subsection. The novel element of the newly proposed framework is the development of a search space reduction method. The main concept of the proposed framework is to define the region in a variable's feasible space that includes the optimum solution for a given objective function, by applying an SSR technique. One way to divide the feasible space into many subspaces and carrying out search in each one of these subspaces is the use of a PWA function which has been widely used on the mpMPC control method. Thus, in our case we use an mpQP formulation to define at every iteration a properly adjusted search space for the NLP solver to explore (Ziogou et al. 2013a). In this context an mpMPC controller is used prior to the solution of the NLP problem in order to provide a suggested solution (u_{mp}) which is transformed into upper and lower bounds ($bu_{act,low}, bu_{act,up}$) augmented by a deviation term (e_{bu}):

$$e_{bu} = \frac{bu_{f,up} - bu_{f,low}}{by_{f,up} - by_{f,low}} e_{y,\max} \tag{14.20}$$

where $bu_{f,up}, bu_{f,low}$ are the feasible upper and lower bounds of variable u, and $by_{f,up}, by_{f,low}$ are the respective bounds for variable y.

The term $e_{y,\ max}$ is the maximum model mismatch between the linearized and the nonlinear model and it is determined by an offline simulation study that involves the whole operating range of y.

The space reduction methods starts by determining the upper and lower bounds utilizing information acquired by a PWA function which explores the entire feasible space. This PWA formulation yields the necessary information to be subsequently used to update the active bounds for the selected variables. The bounds are modified at every iteration and as a consequence the search space of u is reduced to a smaller subset based on:

$$\begin{aligned} bu_{act,low} &= \begin{cases} u_{mp} - e_{bu}, & (u_{mp} - e_{bu}) \quad \geq \quad bu_{f,low} \\ bu_{f,low}, & (u_{mp} - e_{bu}) < \quad bu_{f,low} \end{cases} \\ bu_{act,up} &= \begin{cases} u_{mp} + e_{bu}, & (u_{mp} + e_{bu}) \quad \leq \quad bu_{f,up} \\ bu_{f,up}, & (u_{mp} + e_{bu}) > \quad bu_{f,up} \end{cases} \end{aligned} \tag{14.21}$$

where $bu_{act,low}, bu_{act,up}$ are the active bounds for u. Therefore, the optimizer has a set of updated bounds for the respective manipulated variable u. Apart from the bounds modification the rest of the NLP problem formulation remains the same. The proposed strategy at sampling interval k is summarized in Algorithm 1 (Table 14.1).

Table 14.1 Algorithm for search space reduction technique

Algorithm 1 SSR based on PWA and NLP problem
Input: Warm-start solution (x_k, u_k, y_k, Hessian H), measured variables (y_k^{meas}), parameters (p_k), set-points ($y_{sp,k}$)
Output: Vector of manipulated variables u_{k+1}
1: Calculate error e_k and $\hat{y}_k$
2: Locate CR_i for parameter vector ϑ_k and obtain u_{mp}
3: Calculate $bu_{act,low}, bu_{act,up}$
4: Modify bounds $u_l = bu_{act,low}, u_u = bu_{act,up}$
5: Solve NLP problem (14.13)
6: Obtain u_{k+1}^1 from $u_{k+1} = \left[u_{k+1}^1, \ldots, u_{k+1}^{NE}\right]$

Based on the above algorithm the explicit solution can direct the warm-start procedure for the solution of the NLP problem and thus improve its performance. One can see that the bound is constructed around the mpQP solution by considering the ratio in feasible magnitudes of the input and output variables and the model approximation error. The bound reduction is thus ensured by considering the deviation term.

All these approaches and methods indicate that MPC is a method of interest which is gradually evolving as its advantages are widely recognized since it has a strong potential to numerous applications from small-scale low complexity to large-scale highly nonlinear systems and processes. The interconnection between the advanced model-based controllers and the automation system is facilitated through a custom developed software platform based on state-of-the-art industrial protocols in conjunction with an intuitive and user friendly Human Machine Interface (HMI). The development of this infrastructure addresses the challenges related to the interface of control, computing and communication issues between the MPC and the integrated PEMFC unit.

14.5 Application of Advanced MPC in PEM Fuel Cell Systems

The aim of this Section chapter is to implement the aforementioned developments to the online control of the PEM fuel cell described. More specifically, the dynamic model presented in conjunction with the control objectives for the PEM fuel cell are used to develop various controllers based on MPC methods. These methods are eventually deployed to the small-scale fully automated PEM fuel cell unit od CERTH/CPERI and their efficiency is online monitored by the automation system. The outcome is an integrated control framework which is used to evaluate the behavior of the process using different controllers under varying operating conditions, at both nominal operation and in the presence of disturbances. Based on that, the scope of Sect. 14.5 is to:

Table 14.2 Operating constraints of the PEMFC's variables

Power: 0..5.3 W
Current: 2..10 A
Voltage: 0.3..0.9 V
Air flow: 180..900 cc/min
Hydrogen flow: 180..900 cc/min

- Explore the computer-aided framework that was developed at the SCADA system of the small-scale PEM fuel cell unit.
- Provide a thorough analysis of the design and implementation of each MPC method for the control of a PEM fuel cell system,
- Discuss the effect of each controller to the system's behavior and to evaluate the performance of each strategy with respect to the operational objectives.

Based on this scope, the section is divided into three parts, one for each control strategy, which is deployed to the PEMFC unit. The upper and lower bounds of the variables resulting from the operating constraints of the process guide the algorithm to avoid inappropriate and/or unsafe areas. These bounds were determined by the PEMFC system (*I, V, P*) in conjunction with the operating range of the mass flow controllers (MFC) ($\dot{m}_{air,\,in}, \dot{m}_{H2,\,in}$). Imposing minimum and maximum values to the aforementioned variables, the bounds of all states were determined. The upper and lower bounds of the system operating variables are summarized in Table 14.2.

The MPC framework is deployed online to the fuel cell unit using a custom made OPC-based interface that was developed for the communication between the optimizer and the SCADA system (Ziogou et al. 2013b). The prediction horizon is selected to be sufficiently long (Reble and Allgöwer 2012) with respect to the dynamics of the fuel cell ($T_p = 5$ s) divided into N_p intervals. The control horizon (T_c) was set to be equal to the sampling time of the SCADA system (500 ms) and it is divided into N_c intervals whereas the performance index to be minimized is:

$$\min_{u} J = \sum_{j=1}^{Np} \left(\hat{y}_{k+j} - y_{sp,k+j}\right)^T Q \left(\hat{y}_{k+j} - y_{sp,k+j}\right) + \sum_{l=0}^{Nc-1} \Delta u_{k+l}^T R_1 \Delta u_{k+l} \quad (14.22)$$

$$\text{s.t.:} \quad \begin{array}{l} \dot{x} = f_d(x,u), \quad y = g(x,u) \\ x_l \le x \le x_u, u_l \le u \le u_u, y_l \le y \le y_u \end{array} \quad (14.22a)$$

where Q and R_1 are the weighting matrices that will be fine tuned to accomplish the desired behavior. Finally the selected optimization method is the direct transcription using a reduced gradient NLP solver. The nonlinear fuel cell model is discretized based on OCFE. More specifically there are 10 finite elements (*NE*) with 4 collocation points (N_{cop}) each. Apart form the specification of the region of the optimization, it was necessary to provide a well scaled model during the initial problem formulation.

14.5.1 Implementation of the NMPC Strategy to the PEMFC Unit

Initially the NMPC strategy will be applied to the PEMFC unit. A prerequisite for the online application of the controller is to exhibit both fast response and minimize the error towards the set-point. Therefore, the analysis of the behavior focuses on these two metrics: fast and accurate set-point tracking. Verification and validation of the NMPC framework that deals with the control issues of the fuel cell, was performed by deploying the multivariable controller online to the unit. From the software point of view, the initial model was developed in gPROMS which is transformed into Fortran in order to be applied online to the unit. The selection of the specific programming language was based on the solver MINOS (Murtagh and Saunders 1998). A series of experiments were performed and the response of the controller at power demand changes and at variable operating conditions was explored. The behavior of the proposed scheme is exemplified by an indicative case study.

14.5.1.1 Design of the NMPC Controller

From the control objectives that were analyzed in Sect. 14.3, we selected three of them ($y_{SP} = [P_{SP}, \lambda_{O2,SP}, \lambda_{H2,SP}]$) to be controlled by the NMPC whereas the heat management ($y_{SP,PID} = T_{fc,SP}$) is assigned to two PIDs, one for the heat up and one for the cooling. Thus, for the NMPC problem there are three manipulated variables, the current, the air and the hydrogen flow rates ($u = [I, \dot{m}_{air}, \dot{m}_{H2}]$) and three controlled variables $y = [P_{SP}, \lambda_{O2}, \lambda_{H2}]$, one measured, the power and two unmeasured, the excess ratios of oxygen and hydrogen.

Figure 14.10 illustrates the control configuration which is implemented during the current case study. The nonlinear fuel cell model is comprised out of eight differential equations and one algebraic which are discretized into 10 finite elements (u is piecewise constant) with 4 collocation points each. At this control configuration the energy balance is not included. The resulting discretized system has 381

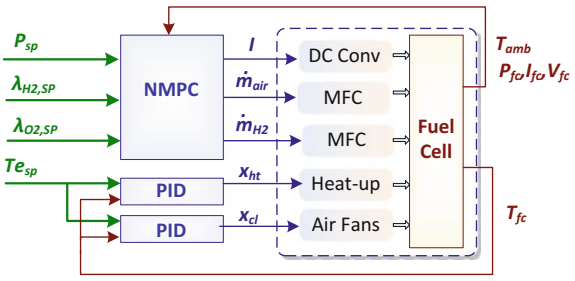

Fig. 14.10 Figure control configuration (NMPC)

variables and 342 constraints. Furthermore, we have one more variable, corresponding to time length (t_f), which is fixed in our problem formulation. The optimization problem has 30 degrees of freedom since the control profiles are discretized into 10 finite elements (u is piecewise constant).

At the performance index Q and R are output and input weighting matrices, respectively. Specifically $[Q_P, Q_{\lambda o2}, Q_{\lambda H2}] = diag(Q)$ and $R = R_I$, where $Q_P, Q_{\lambda o2}, Q_{\lambda H2}$ are penalties on output power ($Q_P = 1.3$), oxygen and hydrogen excess ratio ($Q_{\lambda o2} = 0.23, Q_{\lambda H2} = 0.21$) while R_I is the penalty on the change of the input current ($R_I = 0.04$).

14.5.1.2 NMPC Deployment and PEM Fuel Cell Behavior

The response of the NMPC strategy to various power demands was studied at specific operating conditions of temperature and pressure (T = 338 K, Pt = 1 bar) is shown in this section. The power modifications were within a range of 1 to 5 W, which covers the operational range of the system. The power response to the corresponding control actions of the manipulated variables (Figs. 14.11, and 14.12a, b) as the power demand changes while the oxygen and hydrogen excess ratio remains at a constant set-point ($\lambda_{O2,\ sp}$ = 3, $\lambda_{H2,\ sp}$ = 2.5). The NMPC controller is able to control the fuel cell power to any admissible set-point while maintaining the required excess ratio level.

The mean squared error (MSE) for each objective and some performance metrics are presented at Table 14.3.

The NMPC controller that was deployed to the PEMFC system can efficiently address the issues of power generation in a safe and controlled manner. Overall, it has been clearly illustrated that the NMPC framework is able to deal with uncertainties and achieve trajectory tracking in a satisfactory manner. Also, the NMPC controller has computational requirements that satisfy the time sampling interval without convergence failures while satisfying realistic constraints imposed by the nature of the PEM fuel cell system.

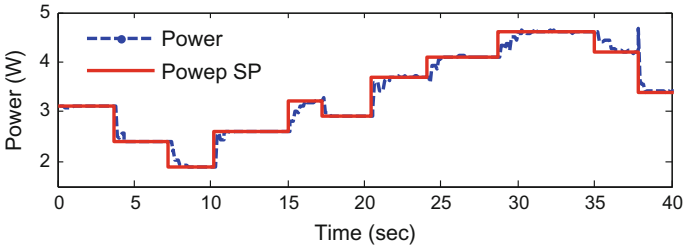

Fig. 14.11 Power demand changes and delivered power by the NMPC controller

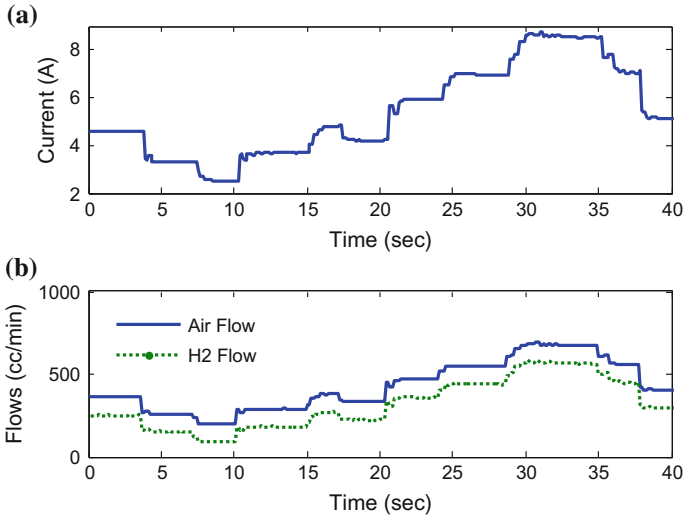

Fig. 14.12 Control actions for the manipulated variables of **a** Current and **b** air and hydrogen mass flows derived by the NMPC controller

Table 14.3 Performance of the online NMPC framework

MSE from set-point	Performance metrics
Power: 20 mW	Max. opt. time: 419 ms
Temperature: 0.6 °C	Average opt. time (SP change): 320 ms
λ_{O2}: 1.5 * 10^{-3}	Average opt. time (steady state): 275 ms
λ_{H2}: 2.1 * 10^{-3}	

14.5.2 mpMPC Strategy for the PEMFC Unit

The second MPC method that was deployed to be evaluated at the PEM Fuel Cell unit is the mpMPC approach. The primary objective of the subsequent analysis is to derive a number of decentralized mpMPC controllers one for each control objective and evaluate the response of the system in simulation mode. The decentralized control configuration which is used is illustrated at Fig. 14.13 including the input/output variables of each controller and the interactions between them.

The four mpMPC controllers are designed based on an equal number of linear models that are derived using a model identification technique. From the software point of view, the initial model was developed in gPROMS (Ziogou et al. 2011b) which is transformed into Matlab code with Simulink for the solution of the mpQP problem and the derivation of the PWA function for the variables into consideration. A discrete reduced order state space (SS) model for each control objective is obtained using a model identification technique that reconstructs adequately the dynamic behavior of the system (Ziogou et al. 2011a). The input-output data are obtained from

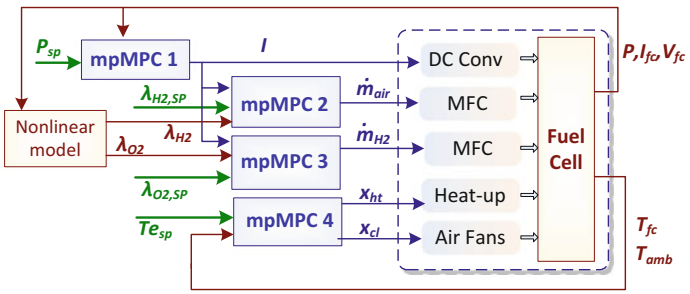

Fig. 14.13 Distributed control configuration (mpMPC)

simulations of the nonlinear model for various operating conditions and the parameters of the SS models are determined from the Identification Toolbox of Matlab. The sampling time for the data is 100 ms for the power, the current and the mass flow rates, and 1 s for the temperature, since temperature has slower response.

14.5.2.1 Design of the mpMPC Controllers

The next step involves the design of mpMPC controllers for the PEM fuel cell system based on the derived SS models. Overall there are four MPC problems which are subject to constraints and they are formulated based on problem (6.1). Thus, for each control variable (represented by an SS model) an MPC problem is formulated. Subsequently, each MPC problem is transformed into an mpQP problem which involves a systematic exploration of the parameter (ϑ) space and results in a set of convex non-overlapping polyhedra (critical regions) of this space, each corresponding to a unique set of active constraints (Sect. 14.4). The parameters for each mpQP problem are presented at Table 14.4.

The control horizon in each problem is 2, therefore there are two optimization variables (u_{t+0}, u_{t+1}). The corresponding parameters of each problem are shown in Table 14.5 along with the respective number each mpMPC controller's critical regions.

The result from the aforementioned actions is four mpMPC controllers and their response will be explored by the following study. Four controllers have been derived in order to fulfill the power demand, while avoiding starvation and maintain the fuel temperature at the desired set point.

Table 14.4 mpQP problem parameters

Objective	Optimization var. (u)	Parameters (θ)
Power	Current	$\theta_1 = [x1\ x2\ I\ P\ P_{sp}]$
O2 Excess ratio	Air flow	$\theta_2 = [x1\ I\ \lambda_{O2}\ \lambda_{O2,sp}]$
H2 Excess ratio	Hydrogen flow	$\theta_3 = [x1\ I\ \lambda_{H2}\ \lambda_{H2,sp}]$
Temperature	Resistance %, Fans %	$\theta_4 = [x1\ T_{amb}\ T_{fc}\ T_{fc,sp}]$

Table 14.5 mpMPC settings and resulting regions

Objective	Optimization variables (u)	Pred. Hor. (Np)	Weight (Q)	Weight (R)	CR
P	$I_{(t+0)}$, $I_{(t+1)}$	10	3	0.01	57
λ_{O2}	$m_{O2(t+0)}$, $m_{O2(t+1)}$	20	1	0.1	13
λ_{H2}	$m_{H2(t+0)}$, $m_{H2(t+1)}$	40	100	0.1	13
T_{fc}	$W_{R(t+0)}$, $W_{R(t+1)}$, $W_{cl(t+0)}$, $W_{cl(t+1,)}$	100	1000	0.001	17

14.5.2.2 mpMPC Deployment and PEM Fuel Cell Behavior

The following case study presents the behavior of the fuel cell system using the aforementioned mpMPC controllers. The response of the decentralized mpMPC controllers to various power demands at specific operating conditions of temperature and pressure ($T = 338$ K, $Pt = 1$ bar) is shown in this section. The power modifications were within a range of 2−6 W, which covers the operational range of the system. The power response to the corresponding control actions of the manipulated variables (Figs. 14.13 and 14.14a, b) as the power demand changes while the hydrogen excess ratio remains at a constant set-point ($\lambda_{H2,\ sp} = 2.5$). On the other hand the oxygen excess ration is modified in order to demonstrate the response of the mpMPC controllers when the requested conditions change.

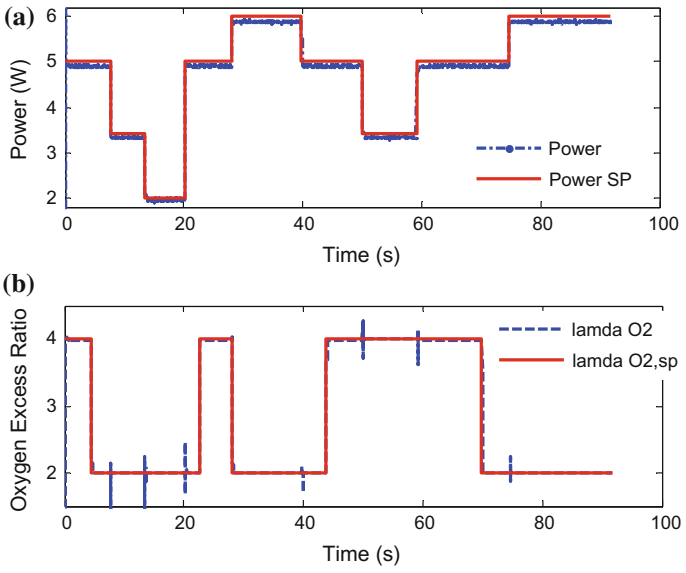

Fig. 14.14 **a** Power demand changes and delivered power, **b** and excess ratio of oxygen by the mpMPC controllers

Table 14.6 Performance of the online NMPC framework

MSE from set-point
Power: 50 mW (max: 150 mW)
Temperature: 1.2°C
λ_{O2}: 1.7 * 10^{-3}
λ_{H2}: 2.5 * 10^{-3}

The mpMPC power controller showed excellent response to load changes and the excess ratio controller demonstrated fast settling time (less than 2 s) after current changes. The current (Fig. 14.14a) and the mass flow rates (Fig. 14.14b) are properly adjusted to fulfill the starvation avoidance constraint by keeping the excess ratio at constant level. The mean squared error (MSE) for each objective and some performance metrics are presented at Table 14.6.

Figures 14.15 and 14.16 illustrate the results of the mpMPC implementation for various temperature set-points as the ambient temperature was constant at 298 K. The performance of the temperature controller is presented at Fig. 14.15, where three temperature set points changes (333, 338, and 343 K) where applied while the power controller's set point is set at constant level (5 W).

It is observed that the controller follows rapidly the set point changes on the temperature without offset. Due to the small size of the PEMFC the system needs to be heated during steady state operation in order to follow the set point (the resistance is working at 7−20 %).

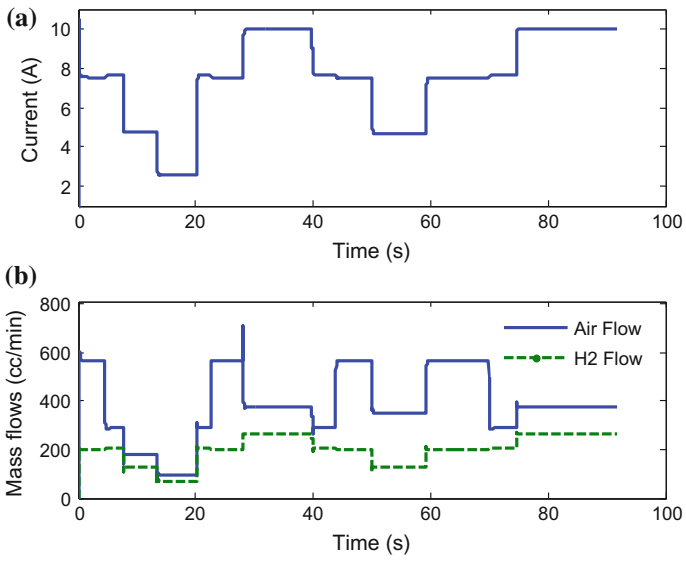

Fig. 14.15 Control actions for the manipulated variables of **a** Current and **b** air and hydrogen mass flows derived by the mpMPC controller

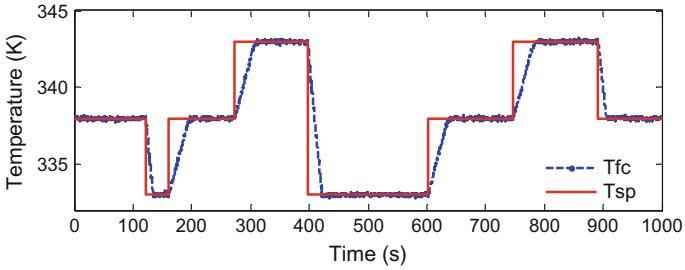

Fig. 14.16 PEM fuel cell temperature profile at temperature set-point changes controlled by the mpMPC

Overall the mpMPC controllers are able to track the desired reference points regardless the fluctuations of the interacting variables. Finally the system response was within the feasible area of operation since the output of the controllers was bounded by the operating constraints and the stability was guaranteed.

14.5.3 Implementation of the exNMPC Strategy to the PEMFC Unit

The third MPC method developed and applied at the PEM fuel cell system, is the combined exNMPC method where the use of mpQP enhances the performance of a typical NMPC formulation through a search space reduction (SSR) algorithm. Based on the exNMPC method a number of controllers are developed which are used in simulation and experimental case studies, in order to exemplify the effectiveness and prove the agility of the newly proposed approach. Figure 14.17 illustrates the control configuration which is implemented during the current case study.

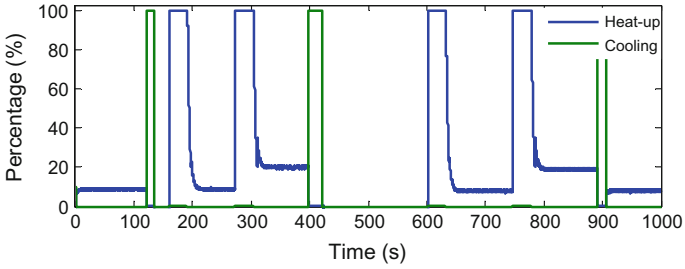

Fig. 14.17 Control actions for the manipulated variables of the temperature cooling/heat-up of the PEM fuel cell by the mpMPC temperature controller (mpMPC)

The development of an exNMPC method is based on two main set of actions, the offline reparatory actions and the online deployment and parameter tuning of the controller. The offline actions are responsible for the development of the exNMPC controller that would satisfy the control and operation objectives of the system and include the following:

- Formulate the NMPC controller (direct transcription and NLP problem).
- Select the manipulated variables to be adjusted online.
- Approximate nonlinear model with linear state space model.
- Calculate the linearization error.
- Determine critical region for each objective by the solution of the mpQP problems.
- Derive PWA functions for each selected control objective.
- Define the SSR parameters.

These actions are implemented for the PEMFC system.

14.5.3.1 exNMPC Problem Formulation

The development of the exNMPC controller involves a typical NMPC controller which is augmented by a preprocessing algorithm (SSR) which is called prior to the solution of the NLP problem. This NMPC controller is based on the direct transcription method and it is formulated according to Sect. 14.5. Besides the NMPC problem formulation, the rest of the preparatory actions involve the SSR algorithm.

NMPC Problem Formulation - Direct Transcription Method

According to the control objectives ($y_{SP} = [P_{SP}, T_{fc,SP}, \lambda_{O2,SP}, \lambda_{H2,SP}]$) defined in Sect. 14.3, there are five manipulated variables ($u = [I, \dot{m}_{air,in}, \dot{m}_{H2,in}, x_{ht}, x_{cl}]$) and four controlled variables ($y = [P_{fc}, T_{fc}, \lambda_{O2}, \lambda_{H2}]$). Two of the manipulated variables (x_{ht}, x_{cl}) are mutually exclusive and they mainly affect one of the controlled variables (T_{fc}). The nonlinear dynamic model of the PEMFC, presented in Sect. 14.3, is comprised of nine differential equations and one algebraic, discretized at 10 finite elements (NE) having 4 collocation points (N_{cop}) each. The analysis of the model discretization based on OCFE is summarized at Table 14.7.

Table 14.7 PEM fuel cell model discretization based on OCFE (NMPC)

Variables
State variables at col. points: 360/Algebraic variables at col. points: 30
Inputs at each finite element: 5
Constraints
States: 270/Algebraic: 30/Continuity at elem. boundaries: 81
Jacobian Matrix: Elements (Total/Non-zero): 168021/3375/Density: 2.009 %

There is one more variable, corresponding to time length (t_f), which is fixed in our problem formulation. Overall the resulting discretized system has 381 variables and 441 equations. The Jacobian has been analytically evaluated to eliminate a costly computational step of the algorithm and to enhance the overall accuracy.

SSR and PWA Functions

Three of the manipulated variables (I, x_{ht}, x_{cl}) are selected to have varying bounds during the operation of the system, that mainly affect two of the controlled variables (P_{fc}, T_{fc}). The development of the SSR technique as analyzed in Sect. 14.4.3, requires a PWA function to approximate the power and the temperature behavior of the fuel cell system using the aforementioned mpMPC approach. Prior to the solution of the mpQP problem two linear discrete state space (ss) models are derived, one for each control objective. The first one (ss_P) approximates the behavior of power and has one input variable (I) and two states ($x_{P,1}, x_{P,2}$) whereas the second ss (ss_{Tfc}) approximates the temperature behavior and has two input variables (x_{ht}, x_{cl}), one disturbance (T_{amb}) and two states ($x_{Tfc,1}, x_{Tfc,2}$).

For each linear model (ss_P, ss_{Tfc}) an mpQP problem is formulated. The first mpQP problem involves five parameters $\vartheta_P = [x_{P,1}\, x_{P,2}\, I\, P\, P_{sp}]$ and the resulting optimal map consists of $N_{CR,P} = 57$ critical regions. The second mpQP problem involves six parameters $\vartheta_{Tfc} = [x_{Tfc,1}\, x_{Tfc,2}\, T_{amb}\, T_{fc}\, T_{fc,sp}]$ while the resulting feasible space, defined by ϑ_{Tfc}, is partitioned into $N_{CR,Tfc} = 23$ critical regions. Based on a simulation analysis the linearization error ($e_{y,\max}$) and the change of bounds are determined (e_{bu}) that will by used for the online adjustment ($bu_{act,low}, bu_{act,up}$) of the active boundaries for each variable. The values of the parameters for the SSR are outlined at Table 14.8.

These parameters are used so that the bounds of I, x_{ht}, x_{cl} are adjusted while the bounds of the other two variables ($\dot{m}_{air}, \dot{m}_{H2}$) are fixed at their feasible bounds.

14.5.3.2 exNMPC Deployment and PEM Fuel Cell Behavior

An experimental scenario is presented that involves various changes in power demand and a step change at temperature. The duration of the scenario is 9 min and the sampling time is 500 ms. The scope of the centralized controller is to concurrently fulfill the four objectives ($y_{SP} = [P_{SP}, T_{fc,SP}, \lambda_{O2,SP}, \lambda_{H2,SP}]$) within the

Table 14.8 Parameters for the SSR algorithm of exNMPC

Manipulated variable	Linearization error ($e_{y,\max}$)	Max deviation from set-point	Max change of bounds (e_{bu})
x_{ht}, x_{cl}	8.60 %	4.3 °C	±8 %
I	17.10 %	0.684 W	±2A

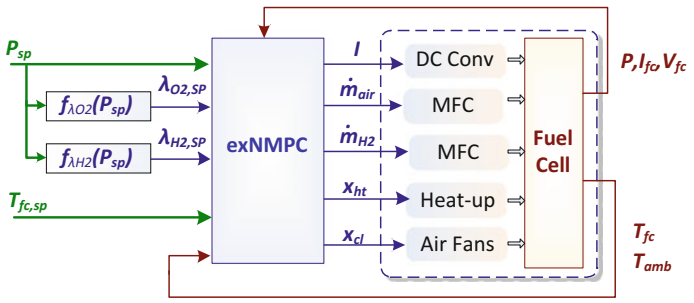

Fig. 14.18 Control configuration of the exNMPC strategy

predefined time constraints of the system. Figure 14.18 illustrates the temperature set-point profile and the fuel cell temperature where the temperature increase from 48−61 °C.

When the set-point is reached the system settles to the desired value with a negligible deviation of ±0.1 °C after a few oscillations. The maximum overshoot and undershoot is 0.7 °C and −0.8 °C respectively. A number of power demand changes were also applied to the system during the 9 min of the experiment. These step changes are randomly generated and cover the full operating range of the fuel cell (Figs. 14.19 and 14.20).

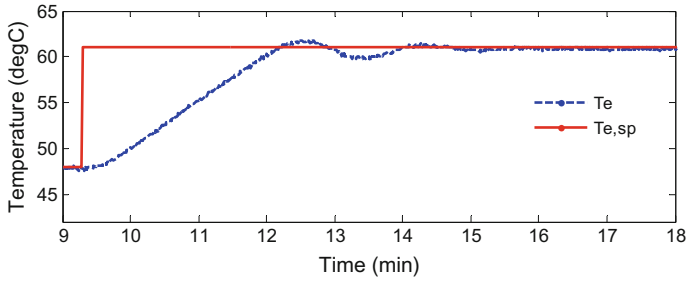

Fig. 14.19 Step change at the fuel cell temperature controlled by the exNMPC

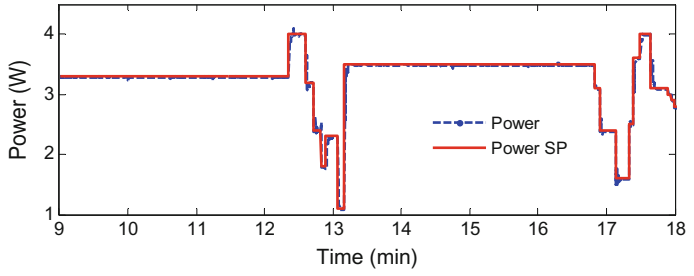

Fig. 14.20 Step change at the fuel cell power controlled by the exNMPC

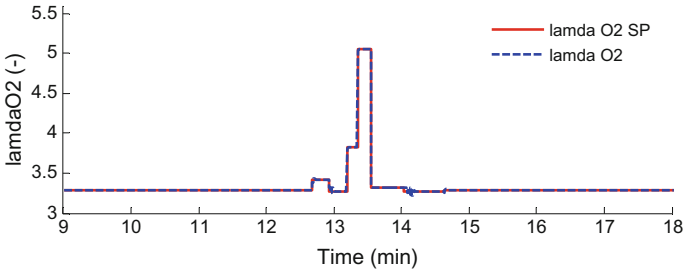

Fig. 14.21 Oxygen excess ratio profile controlled by the exNMPC

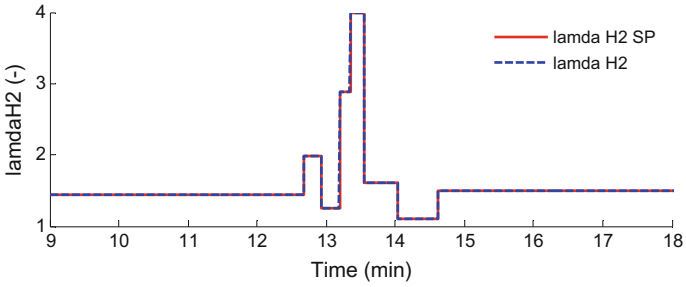

Fig. 14.22 Hydrogen excess ratio profile controlled by the exNMPC

From the power delivery point of view the exNMPC framework exhibits a very good performance as it can respond to frequent and abrupt changes of the power demand. Figures 14.21 and 14.22 illustrate the corresponding oxygen and hydrogen excess ratio profiles that are adjusted according to the power demand, based on the minimum air and hydrogen considerations.

As illustrated at Figs. 14.20 and 14.21 the excess ratios of oxygen and hydrogen range from 3.2 to 5 and 1.0 to 4.1 respectively and the controller is able to adjust the hydrogen flow rate in order to reach the set-point which is modified in a feed forward manner.

Finally the mean squared error (MSE) for each objective and some performance metrics are presented at Table 14.9.

Table 14.9 Performance of the Online exNMPC framework

MSE from set-point		Performance metrics	
Power	14 mW	Max. opt. time	290 ms
Temperature	0.14 °C	Average opt. time	186 ms
λ_{O2}	1.7×10^{-3}	Max. iterations	172
λ_{H2}	4.6×10^{-3}	Max. func. calls	433

The MSE shows that the exNMPC can accurately fulfill the objective for power delivery (P_{SP}) in a safe operating region while minimizing the gas consumption ($\lambda_{O2,SP}, \lambda_{H2,SP}$) and concurrently provide a stable environment with respect to the temperature ($T_{fc,SP}$) condition. From the performance metrics of Table 14.2 it is evident that the computational constraints are satisfied and the results from the simulation study are verified by the online application of the controller to the fuel cell unit. Hence, the optimum operation of the fuel cell is achieved at varying operating conditions and rapid power changes.

Based on the aforementioned analysis it is evident that the centralized controller that relies on the exNMPC strategy is able to guide the system to the desired condition based on the power demands and by respecting the constraints posed by the system operation related to safety and long-term stable operation.

In this Sect. (14.4.3) the behavior of the newly proposed synergetic exNMPC controller is presented. A warm-start method is complemented by an SSR technique, relying on a PWA function that sets the basis for the improved behavior of the optimizer. By this cooperation the computational requirements for the solution of an NLP problem are reduced. The importance of this synergy is illustrated by a challenging multivariable nonlinear control problem with measured and unmeasured variables that involves concurrently four operation objectives for the PEM fuel cell system. The response of the proposed strategy is demonstrated through a experimental case study.

By this case the evaluation of the developed advanced control framework concludes. Based on the response analysis of the NMPC, mpMPC and exNMPC it has been demonstrated that the developed framework can used for the testing of various control configuration ranging from conventional PID controllers, decentralized mpMPC schemes and centralized NMPC and exNMPC approaches. The selection of the appropriate method/strategy of controller depends highly on the application or the focus of the study, as all type of controllers have advantages and limitations.

14.6 Concluding Remarks and Main Contributions

As the fuel cell system is an essential element of a promising, benign and environmental friendly technology that could be part of a decarbonised and sustainable future economy, a continuous and persistent effort in technological innovation is needed. Driven by this incentive, the multidisciplinary research effort of this work was built around a fuel cell system, which was supervised by an automation system, complimented by predictive control algorithms that act as a catalyst towards the improvement of the response and protection of the fuel cell's operation and safety. This work, by careful consideration of the current status of the fuel cell technology and the importance of advanced control, has contributed to the following:

- A control-oriented dynamic nonlinear dynamic semi-empirical PEM fuel cell model is utilized and a formal parameter estimation procedure is employed that can be applied to other fuel cell systems.
- The online multivariable nonlinear controllers (NMPC, exNMPC) utilize the full dynamic nonlinear model of the fuel cell. This is greatly enhanced by the use of the direct transcription method that was part of the simultaneous optimization approach based on the reduced gradient projection NLP solver that was utilized.
- A novel synergetic framework (exNMPC) between two well established control methods (mpMPC and NMPC) is discussed based on a newly proposed pre-processing bound related technique. A Search Space Reduction (SSR) algorithm of the feasible space is developed combined with an existing technique related to the optimization problem, a warm-start initialization procedure of the NLP solver. The algorithm bounds the active area of the variables so that the exploration of the search space by the NLP solver is reduced at every iteration during online control. Besides the MPC framework the SSR algorithm can be applied for the enhancement of the solution of dynamic optimization problems.

The modular model-based control framework was developed, deployed online and systematically evaluated. The salient capabilities of the proposed synergetic formulation were revealed through the solution of the multivariable nonlinear control problem involving the optimal operation of the PEM fuel cell system. More specifically, for the motivating system into consideration, all the model-based algorithmic developments have been experimentally tested at nominal conditions, in the presence of disturbances and during start-up. A comparative analysis between the mpMPC, NMPC and rssNMPC controllers reveals the merits and limitations of each approach. As a result the fuel cell system operates economically and at a stable environment regardless of the varying operating conditions. Also the computational time is significantly improved (>50 %) by the deployment of the proposed rssNMPC framework without compromising the accuracy of the obtained solution.

References

A European strategic energy technology plan (SET Plan). (2007). Towards a low carbon future COM(2007) 723.

Ahn, J., & Choe, S. (2008). Coolant controls of a PEM fuel cell system. *Journal of Power Sources, 179*, 252–264.

Bauer, M., & Craig, I. K. (2008). Economic assessment of advanced process control—A survey and framework. *Journal of Process Control, 18*(1), 2–18.

Bavarian, M., Soroush, M., Kevrekidis, I., & Benziger, J. (2010). Mathematical modeling, steady-state and dynamic behavior, and control of fuel cells: A review. *Industrial and Engineering Chemistry Research, 49*, 7922–7950.

Bemporad, A., Morari, M., Dua, V., & Pistikopoulos, E. N. (2002). The explicit linear quadratic regulator for constrained systems. *Automatica, 38*, 3–20.

Benson, H. Y., & Shanno, D. F. (2008). Interior-point methods for nonconvex nonlinear programming: regularization and warmstarts. *Computational Optimization and Applications, 40*, 143–189.

Betts, J. T. (2001). Practical methods for optimal control and estimation using nonlinear programming. *Advances in design and control* (Vol. 19). Philadelphia: SIAM.

Biegler, L. T., & Grossmann, I. E. (2004). Part I: Retrospective on optimization. *Computers & Chemical Engineering, 28*(8), 1169–1192.

Biegler, L. T., Cervantes, A. M., & Wachter, A. (2002). Advances in simultaneous strategies for dynamic process optimization. *Chemical Engineering Science, 57*, 575–593.

CSS. (2011). The impact of control technology. In T. Samad., & A.M. Annaswamy (eds.) *IEEE control systems society*. Retrieved December 2015 from www.ieeecss.org.

Diehl, M., Bock, H., Schloder, J., Findeisen, R., Nagy, Z., & Allgöwer, F. (2002). Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *Journal of Process Control, 12*(4), 577–585.

Diehl, M., Ferreau, H. J., & Haverbeke, N. (2009). Efficient numerical methods for nonlinear MPC and moving horizon estimation. In *Nonlinear model predictive control* (pp. 391–417). Berlin Heidelberg: Springer.

Domahidi, A., Zgraggen, A. U., Zeilinger, M. N., Morari, M., Jones, C. N. (2012). Efficient interior point methods for multistage problems arising in receding horizon control. In *IEEE 51st Annual Conference on Decision and Control (CDC)* 10–13 December 2012 (pp. 668, 674).

Dua, V., Bozinis, N. A., & Pistikopoulos, E. N. (2002). A multiparametric programming approach for mixed-integer quadratic engineering problems. *Computers and Chemical Engineering, 26*, 715–733.

Engell, S. (2007). Feedback control for optimal process operation. *Journal of Process Control, 17*, 203–219.

Eqtami, A., Dimarogonas, D. V., & Kyriakopoulos, K. J. (2011). Novel event-triggered strategies for model predictive controllers. In *50th IEEE conference on decision and control and european control conference (CDC-ECC)*, 12−15 December 2011 (pp. 3392–3397).

European Commission (2011). A roadmap for moving to a competitive low carbon economy in 2050. In *COM 112*, Brussels.

Ferreau, H. J., Bock, H. G., & Diehl, M. (2008). An online active set strategy to overcome the limitations of explicit MPC. *International Journal of Robust and Nonlinear Control, 18*(8), 816–830.

Findeisen, R., Allgöwer, F., & Biegler, L. T. (2007). Assessment and future directions of nonlinear model predictive control. In *Lecture Notes in Control and Information Sciences* (Vol 358). Verlag, Berlin: Springer.

Finlayson, B. (1992). Numerical *methods for problems with moving fronts*. Ravenna Park Publishing.

Fuel Cell Today. (2012). *The Fuel Cell Industry Review*. Retrieved December 2015 from http://www.fuelcelltoday.com.

Kameswaran, S., & Biegler, L. T. (2008). Convergence rates for direct transcription of optimal control problems using collocation at Radau points. *Computational Optimization and Applications, 41*(1), 81–126.

Kvasnica, M., Grieder, P., Baotic, M., & Morari, M. (2004). *Multi-Parametric Toolbox (MPT)* (pp. 448–462).

Magni, L., Raimondo, D., & Allgöwer, F. (2009). Nonlinear model predictive control: Towards new challenging applications. In *Lecture Notes in Control and Information Sciences* (Vol. 384). Berlin: Springer Verlag.

Mayne, D. Q., Rawlings, J. B., Rao, C. V., & Scokaert, P. O. M. (2000). Constrained model predictive control: Stability and optimality. *Automatica, 36*, 789–814.

Murtagh, B. A., Saunders, M. A. (1998). MINOS 5.5 User's Guide. *Technical report SOL 83–20R*, Stanford University.

Pannocchia, G., Rawlings, J. B., & Wright, S. J. (2007). Fast, large-scale model predictive control by partial enumeration. *Automatica, 43*, 852–860.

Patrinos, P., & Bemporad, A. (2012). An accelerated dual gradient-projection algorithm for linear model predictive control. In *IEEE 51st Annual Conference on Decision and Control (CDC)* 10–13 December 2012 (pp. 662, 667).

Pistikopoulos, E. N. (2012). From multi-parametric programming theory to MPC-on-a-chip multi-scale systems applications. *Computers & Chemical Engineering, 47*, 57–66.

Pistikopoulos, E. N., Dua, V., Bozinis, N. A., Bemporad, A., & Morari, M. (2002). On-line optimization via off-line parametric optimization tools. *Computers & Chemical Engineering, 26*(2), 175–185.

Pistikopoulos, E. N., Georgiadis, M., & Dua, V. (2007). *Multi-parametric model-based control: theory and applications*. Weinheim: Wiley-VCH.

POP. (2007). Parametric optimization solutions (ParOS) Ltd.

Pukrushpan, J. T., Peng, H., & Stefanopoulou, A. G. (2004). Control-oriented modeling and analysis for automotive fuel cell systems. *Journal of Dynamic Systems Measurement and Control-Transactions of the ASME, 126*(1), 14–25.

Qin, S., & Badgwell, T. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice, 11*(7), 733–764.

Raimondo, D. M., Riverso, S., Summers, S., Jones, C., Lygeros, J., & Morari, M. (2012). A set-theoretic method for verifying feasibility of a fast explicit nonlinear model predictive controller. *Distributed Decision Making and Control, LNCIS, 417*, 289–311.

Rawlings, J. B., & Mayne, D. Q. (2009). *Model predictive control: Theory and design*. Madison: Nob Hill Publishing.

Reble, M., & Allgöwer, F. (2012). Unconstrained model predictive control and suboptimality estimates for nonlinear continuous-time systems. *Automatica, 48*(8), 1812–1817.

Schafer, A., Kuehl, P. Baschuk, N., Diehl, M., Schloder, J., & Bock, H. (2007). Fast reduced multiple shooting methods for nonlinear model predictive control. *Chemical Engineering and Processing, 46*(11), 1200–1214.

Schmittinger, W., & Vahidi, A. (2008). A review of the main parameters influencing long-term performance and durability of PEM fuel cells. *Journal of Power Sources, 180*, 1–14.

Stefanopoulou, A. G., & Suh, K. W. (2007). Mechatronics in fuel cell systems. *Control Engineering Practice, 15*, 277–289.

Tamimi, J., & Li, P. (2010). A combined approach to nonlinear model predictive control of fast systems. *Journal of Process Control, 20*(9), 1092–1102.

Wang, Y., & Boyd, S. (2010). Fast model predictive control using online optimization. *IEEE Transactions on Control Systems Technology, 18*(2), 267–278.

Yang, X., & Biegler, L. T. (2012). Advanced-multi-step Nonlinear Model Predictive Control. In *8th IFAC Symposium on Advanced Control of Chemical Process*es, Singapore, July 10−13, (pp. 426–431).

Zavala, V. M., & Biegler, L. T. (2009). The advanced step NMPC controller: Optimality. *Stability and Robustness. Automatica, 45*(1), 86–93.

Zeilinger, M. N., Jones, C. N., & Morari, M. (2011). Real-Time suboptimal model predictive control using a combination of explicit MPC and online optimization. *IEEE Transactions on Control Systems Technology, 56*(7), 1524–1534.

Ziogou, C., Pistikopoulos, E. N., Georgiadis, M. C., Voutetakis, S., & Papadopoulou, S. (2013a). Empowering the performance of advanced NMPC by multi-parametric programming—An application to a PEM fuel cell system. *Industrial Engineering and Chemistry Research, 52*(13), 4863–4873.

Ziogou, C., Papadopoulou, S., Georgiadis, M. C., & Voutetakis, S. (2013b). On-line nonlinear model predictive control of a PEM fuel cell system. *Journal of Process Control, 23*(4), 483–492.

Ziogou C., Panos C., Kouramas K., Papadopoulou S., Georgiadis M. C., Voutetakis S., et al. (2011a). Multi-parametric model predictive control of an automated integrated fuel cell testing unit. *Computer Aided Chemical Engineering, 29*, 744–747.

Ziogou, C., Voutetakis, S., Papadopoulou, S., & Georgiadis, M. C. (2011b). Modeling simulation and validation of a PEM fuel cell system. *Computers and Chemical Engineering, 35*, 1886–1900.

Part III
Planning and Operation of Energy Systems

Chapter 15
An Optimization Framework for Power Systems Planning Considering Unit Commitment Constraints

Nikolaos E. Koltsaklis, Georgios M. Kopanos and Michael C. Georgiadis

Abstract This chapter presents a generic mixed integer linear programming (MILP) model that integrates the unit commitment problem (UCP), i.e., daily energy planning with the long-term generation expansion planning (GEP) framework. Typical daily constraints at an hourly level such as start-up and shut-down related decisions (start-up type, minimum up and down time, synchronization, soak and desynchronization time constraints), ramping limits, system reserve requirements are combined with representative yearly constraints such as power capacity additions, power generation bounds of each unit, peak reserve requirements, and energy policy issues (renewables penetration limits, CO_2 emissions cap and pricing). For modelling purposes, a representative day (24 h) of each month over a number of years has been employed in order to determine the optimal capacity additions, electricity market clearing prices, and daily operational planning of the studied power system. The model has been tested on an illustrative case study of the Greek power system. Our approach aims to provide useful insight into strategic and challenging decisions to be determined by investors and/or policy makers at a national and/or regional level by providing the optimal energy roadmap under real operating and design constraints.

Nomenclature

Acronyms

EU	European Union
GEP	Generation Expansion Planning
MILP	Mixed Integer Linear Programming
NGCC	Natural Gas Combined Cycle

N.E. Koltsaklis · M.C. Georgiadis (✉)
Department of Chemical Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
e-mail: mgeorg@auth.gr

G.M. Kopanos
School of Water, Energy and Environment, Cranfield University, Bedfordshire MK43 0AL, UK

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_15

NGGT	Natural Gas Turbine (or Natural Gas Open Cycle)
O&M	Operational and Maintenance
RET	Renewable Energy Technologies
SMP	System Marginal Price
UCP	Unit Commitment Problem

Sets

$a \in A$	Set of start-up types {hot, warm, cold}
$bl \in BL$	Set of blocks of the energy offer function of each hydrothermal unit (or energy bids for load representatives)
$i \in I^{EX}$	Set of existing units
$i \in I^{HT}$	Set of hydrothermal units
$i \in I^{NEW}$	Set of new candidate units
$i \in I^{RES}$	Set of renewable units (including hydro units)
$i \in I^{RES-}$	Set of renewable units (not including hydro units)
$i \in I^{S}$	Set of units $i \in I$ that are (or can be) installed in sector $s \in S$
$i \in I^{TH}$	Set of thermal units
$i \in I^{z}$	Set of units $i \in I$ that are (or can be) installed in zone $z \in Z$
$i \in I$	Set of all units
$m \in M$	Set of months
$nc \in NC^{s}$	Set of neighbouring countries $nc \in NC$ interconnected with sector $s \in S$
$nc \in NC^{z}$	Set of neighbouring countries $nc \in NC$ interconnected with zone $z \in Z$
$nc \in NC$	Set of interconnections (neighbouring countries)
$s \in S^{s'}$	Set of sectors $s \in S$ interconnected with sector $s' \neq s \in S$
$(s, s') \in S$	Set of sectors
$(t, t') \in T$	Set of hours
$(y, y') \in Y$	Set of years
$z \in Z$	Set of zones

Parameters

$AV_{i,z,m,t}$	Availability factor of each unit $i \in I^{RES}$ in zone $z \in Z$, month $m \in M$, and hour $t \in T$ (p.u.)
$CEXPB_{nc,bl,y,m,t}$	Export (load) revenues of block $bl \in BL$ of interconnection $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (€/MW)
$CFL_{s,s',y}$	Maximum corridor flow from sector $s \in S$ to sector $s' \neq s \in S$ in year $y \in Y$ (MW)
$CIMPB_{nc,bl,y,m,t}$	Marginal cost of block $bl \in BL$ of the imported energy offer function from interconnection $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (€/MW)
$CIM_{nc,y}$	Power capacity of interconnection $nc \in NC$ in each year $y \in Y$ (MW)

$CO_2\_CAP_y$ Maximum allowable CO_2 emissions produced in year $y \in Y$ (t CO_2)

$CO_2\_EF_{i,bl}$ CO_2 emission factor of each unit $i \in I^{TH}$, in power capacity block $bl \in BL$ (tCO_2/MWh)

$CPB_{i,bl,y,m,t}$ Marginal cost of block $bl \in BL$ of the energy offer function of each unit $i \in I^{HT}$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (€/MW)

$CR1_{i,y,m,t}$ Price of the primary energy offer of each unit $i \in I^{HT}$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (€/MW)

$CR2_{i,y,m,t}$ Price of the secondary range energy offer of each unit $i \in I^{HT}$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (€/MW)

CRF_i Capital recovery factor of each unit $i \in I^{NEW}$ (p.u.)

CSD_i Shut-down cost of each thermal unit $i \in I^{TH}$ (€)

DUR_m Duration of each month (in days)

$Dem_{s,y,m,t}$ Power load of sector $s \in S$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$EFORIM_{nc,y}$ Unavailability factor of each interconnection $nc \in NC$ in each year $y \in Y$ (p.u.)

$EFOR_{i,y}$ Unavailability factor of each unit $i \in I^{TH}$ in each year $y \in Y$ (p.u.)

$Exblock_{nc,bl,y,m,t}$ Quantity of each power capacity block $bl \in BL$ of each interconnection $nc \in NC$ (exports), in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

FOM_i Fixed operational and maintenance cost of each unit $i \in I^{RES-}$ (€/MW)

$FastR2Req^{down}_{y,m,t}$ System requirements in fast secondary-down reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$FastR2Req^{up}_{y,m,t}$ System requirements in fast secondary-up reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$IC_{i,y}$ Installed capacity of unit $i \in I^{EX}$ in year $y \in Y$ (MW)

$INL_{z,y,m,t}$ Injection losses coefficient in zone $z \in Z$, year $y \in Y$ month $m \in M$, and hour $t \in T$ (p.u.)

$INVC_{i,y}$ Investment cost of unit $i \in I^{NEW}$ in year $y \in Y$ (€/MW)

$Imblock_{nc,bl,y,m,t}$ Quantity of each power capacity block $bl \in BL$ of each interconnection $nc \in NC$ (imports), in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$MRES_y$ Maximum RES penetration in year $y \in Y$ (p.u.)

$PBL_{i,bl,y,m,t}$ Quantity of each power capacity block $bl \in BL$ of the energy offer function of unit $i \in I^{HT}$ in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$PFX_{i,y,m,t}$ Fixed (non-priced) component of the energy offer function of each unit $i \in I$ in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$PMAX^{dp}_i$ Maximum power output (dispatchable phase) of each unit $i \in I^{HT}$ (MW)

$PMAX_i^{sc}$ Maximum power output (when providing secondary reserve) of each unit $i \in I^{HT}$ (MW)

$PMIN_i^{dp}$ Minimum power output (dispatchable phase) of each unit $i \in I^{HT}$(MW)

$PMIN_i^{sc}$ Minimum power output (when providing secondary reserve) of each unit $i \in I^{HT}$(MW)

PSK_i Power output of each thermal unit $i \in I^{TH}$ when operating in soak phase (MW)

$R1Req_{y,m,t}$ System requirements in primary-up reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$R1_i$ Maximum contribution of unit $i \in I^{HT}$ in primary reserve (MW)

$R2Req_{y,m,t}^{down}$ System requirements in secondary-down reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$R2Req_{y,m,t}^{up}$ System requirements in secondary-up reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$R2_i$ Maximum contribution of unit $i \in I^{HT}$ in secondary reserve (MW)

$R3Req_{y,m,t}$ System requirements in tertiary reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$R3_i^{nsp}$ Maximum contribution of unit $i \in I^{HT}$ in non-spinning tertiary reserve (MW)

$R3_i^{sp}$ Maximum contribution of unit $i \in I^{HT}$ in spinning tertiary reserve (MW)

REN_y RES penetration target in the power mix in year $y \in Y$ (p.u.)

$RES\_CAP_{i,z,y}$ Maximum allowable capacity of each unit $i \in I^{RES-}$ in zone $z \in Z$, year $y \in Y$ (p.u.)

RR_down_i Ramp-down rate of unit $i \in I^{HT}$ (MW)

RR_sec_i Ramp rate of unit $i \in I^{HT}$ when providing secondary reserve (MW)

RR_up_i Ramp-up rate of unit $i \in I^{HT}$ (MW)

T_con_i Construction time of unit $i \in I^{NEW}$ (y)

T_dn_i Desynchronization time of thermal unit $i \in I^{TH}$ (h)

T_down_i Minimum down time of thermal unit $i \in I^{TH}$ (h)

T_sd_i Non-operational time (after being shut-down) of thermal unit $i \in I^{TH}$ (h)

$T\_sd_i^{cold}$ Non-operational time of thermal unit $i \in I^{TH}$ before going from warm to cold standby condition (h)

T_sd^{prior} Extended time period in the past (greater than the higher cold reservation time of all thermal units) (h)

$T\_sd_i^{warm}$ Non-operational time of thermal unit $i \in I^{TH}$ before going from hot to warm standby condition (h)

$T\_sk_i^a$ Type-a soak time of thermal unit $i \in I^{TH}$ (h)

$T\_sn_i^a$ Type-a synchronization time of thermal unit $i \in I^{TH}$ (h)

$T\_sd_i^{1,2}$	Non-operational time of thermal unit $i \in I^{TH}$ before changing standby condition (h) (1: from hot to warm condition, 2: from warm to cold condition)
T_up_i	Minimum up time of thermal unit $i \in I^{TH}$ (h)
RSV	Minimum peak reserve requirements (p.u.)

Variables

$cf_{s,s',y,m,t}$	Corridor power flow from sector $s \in S$ to $s \neq s' \in S$ in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$c_{i,y}$	Power capacity of unit $i \in I$ in year $y \in Y$ (MW)
$expb_{nc,bl,y,m,t}$	Quantity of power capacity block $bl \in BL$ exported to neighbouring country $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$exp_{nc,y,m,t}$	Total energy withdrawal (exports) to neighbouring country $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$fastr2_{i,y,m,t}^{down}$	Contribution of unit $i \in I^{HT}$ in fast secondary-down reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$fastr2_{i,y,m,t}^{up}$	Contribution of unit $i \in I^{HT}$ in fast secondary-up reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$im\_inj_{nc,y,m,t}$	Net (taking into account energy losses) energy injection (imports) to neighbouring country $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$impb_{nc,bl,y,m,t}$	Quantity of power capacity block $bl \in BL$ imported from neighbouring country $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$imp_{nc,y,m,t}$	Energy injection (imports) from neighbouring country $nc \in NC$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$nc_{i,y}$	Newly-built capacity of unit $i \in I^{NEW}$ available for the first time in year $y \in Y$ (MW)
$p\_dn_{i,y,m,t}$	Power output of thermal unit $i \in I^{TH}$ when operating in the desynchronization phase in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$p\_inj_{i,y,m,t}$	Net (taking into account energy losses) energy injection (generation) from unit $i \in I^{HT}$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$p\_sk_{i,y,m,t}$	Power output of thermal unit $i \in I^{TH}$ when operating in the soak phase in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$pb_{i,bl,y,m,t}$	Quantity of power capacity block $bl \in BL$ of unit $i \in I^{HT}$, dispatched in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$p_{i,y,m,t}$	Energy injection (generation) from unit $i \in I^{HT}$, in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)
$r1_{i,y,m,t}^{up}$	Contribution of unit $i \in I^{HT}$ in primary-up reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$r2^{down}_{i,y,m,t}$ Contribution of unit $i \in I^{HT}$ in secondary-down reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$r2^{up}_{i,y,m,t}$ Contribution of unit $i \in I^{HT}$ in secondary-up reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$r3_{i,y,m,t}$ Total contribution of unit $i \in I^{HT}$ in tertiary reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$r3^{nsp}_{i,y,m,t}$ Contribution of unit $i \in I^{HT}$ in non-spinning tertiary reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

$r3^{sp}_{i,y,m,t}$ Contribution of unit $i \in I^{HT}$ in spinning tertiary reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$ (MW)

Binary Variables

$cn_{i,y}$ Equals 1 if the decision for the construction of unit $i \in I^{NEW}$ is to be taken in year $y \in Y$

$w_{i,y,m,t}$ Equals 1 if unit $i \in I^{HT}$ is shut-down in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_3ns_{i,y,m,t}$ Equals 1 if unit $i \in I^{HT}$ contributes to non-spinning tertiary reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_dn_{i,y,m,t}$ Equals 1 if thermal unit $i \in I^{TH}$ operates in the desynchronization phase in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_dp_{i,y,m,t}$ Equals 1 if unit $i \in I^{HT}$ operates in the dispatchable phase in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_sc_{i,y,m,t}$ Equals 1 if unit $i \in I^{HT}$ contributes to secondary reserve in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_sk_{i,y,m,t}$ Equals 1 if thermal unit $i \in I^{TH}$ operates in the soak phase in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_sk^{a}_{i,y,m,t}$ Equals 1 if thermal unit $i \in I^{TH}$ operates in the type-a soak phase in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_sn_{i,y,m,t}$ Equals 1 if thermal unit $i \in I^{TH}$ operates in the synchronization phase in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_sn^{a}_{i,y,m,t}$ Equals 1 if thermal unit $i \in I^{TH}$ operates in the type-a synchronization phase in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_st_{i,y,m,t}$ Equals 1 if unit $i \in I^{HT}$ starts-up in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x\_st^{a}_{i,y,m,t}$ Equals 1 if a type-a start-up decision is taken for thermal unit $i \in I^{TH}$ in year $y \in Y$, month $m \in M$, and hour $t \in T$

$x_{i,y,m,t}$ Equals 1 if unit $i \in I^{HT}$ is committed (operational) in year $y \in Y$, month $m \in M$, and hour $t \in T$

15.1 Introduction

Climate change, environmental pollution and shrinking resource availability are issues that have to be addressed in global sustainability context. For decades, many conflict-laden issues such as the growing population, the boosting economic development, the increasing energy demand, the deteriorating environmental quality, the changing climate, and the resource depletion have called for more effective approaches for planning energy systems. However, such planning efforts are further complicated with a variety of processes that should be taken into consideration by decision makers, including the diversity of supply and mitigation technology options available (influencing model size and complexity), the temporal and/or spatial evolutions of parameters over medium- to long-term time horizons, the dynamic variation of system's conditions, the environmental and social arguments (Koltsaklis et al. 2014; Manfren et al. 2011; Li et al. 2011).

With regard to energy systems, there exists high uncertainty for the projection of their trends and evolution, as many parameters can influence them to a significant extent. These parameters include market dynamics (e.g., sharp cost reductions in the investment cost of renewable energy technologies, shale gas boom in the United States), political decisions (e.g., nuclear phase-out in several countries), financial decisions (e.g. credit risk) or even geopolitical tensions affecting fuel price evolution (e.g., oil price) (Koltsaklis and Georgiadis 2015).

As a consequence, there is a need to implement a comprehensive, analytical and detailed energy planning in order to provide a roadmap towards an affordable, sustainable and secure energy future. In this context, a modelling framework is necessary to optimally determine the energy technologies to be utilized in the power sector and/or in other energy sectors, according to specific projections and policy targets. Several modelling approaches of energy systems have been developed including top-down macroeconomic energy models, bottom-up energy systems models covering several end-use sectors and sector-specific energy/power systems planning models. Each methodological framework has its own advantages and disadvantages based on technology details, temporal resolution and interaction among energy (one and/or other sectors) and economy (Kannan and Turton 2013).

The long-term Generation (or capacity) Expansion Planning (GEP) problem determines the optimal type of energy technologies, the capacity expansion, location, and time construction of new power generation units while minimizing total cost over a long planning horizon. The most important factors such as the projected electricity demand as well as any economic (e.g., investment, fuel cost), technical (e.g., operational availability, technical operational minimum) and environmental (CO_2 emissions targets and pricing) parameters are taken into consideration and the corresponding constraints must be satisfied. The optimization of the GEP problem is a very crucial and challenging task that should consider multiple aspects and decision criteria. There are various available technologies that can be deployed to meet the electricity demand. These options can be identified based on several factors, including operational technical characteristics, environmental impact,

variations on the fuel prices as well as construction lead times and life-times of the power plants. The traditional least cost electricity planning attempts to answer these questions and to determine appropriate strategies for the operation and expansion of the electricity sector. In the past, cost minimization was the only criterion in the implementation of energy planning models. However, after 1980s, the environmental dimension has been taken into consideration due to the growing concern about global warming and its effects (Koltsaklis et al. 2015).

Focusing on the power generation sector, for decades, traditional long-term generation expansion planning (GEP) models are often based on simplifications with regard to operational costs and issues ignoring realistic operating details such as start-up and shut-down related decisions, ramp rates and operating reserves. These issues are typically addressed by models developed for the solution of the unit commitment problem (UCP), i.e., the determination of an optimal strategy so as to satisfy the electricity load at minimum total cost by making use of an optimal combination of distinct power plants (Ming et al. 2014). This approach has been successfully implemented on the grounds that load profile is highly predictable and with quite slow time dynamics. However, due to the intermittency and the high fluctuations of the rapidly penetrated renewable energy technologies (RET), i.e., wind and solar, the net load (total load minus load met by renewables) required to be covered by the traditional hydrothermal power plants is characterized by higher uncertainty and exhibits faster time dynamics. As a result, it is highly debatable how much operational detail has to be introduced into the long-term planning models in order to adequately secure the validity and the robustness of the investment decisions to be implemented. For these reasons, the incorporation of short-term decisions (UCP) into the long-term planning framework (GEP problem) is able to enhance and strengthen the accuracy of the decisions to be made and guarantees the stability of power networks (Palmintier and Webster 2011).

The existing literature is rich in works that have separately addressed the GEP problem and the UCP of the power systems of multiple countries. A study provided a survey of GEP modelling techniques focusing on detailed lists of previous research models and approaches applied for GEP in centralized monopolistic electricity market (Kagiannas et al. 2004). Another study described how the needs of utility planners for optimization models have changed over the years in response to environmental concerns, increased competition, and growing short and long-run uncertainty (Hobbs 1995).

A single period mixed integer linear programming (MILP) investment planning model under various yearly electricity demand projections and CO_2 emission targets was presented for the solution of the GEP problem (Ahmet et al. 2015). The same authors modified their deterministic model into a two-stage stochastic mathematical model incorporating financial risk management in their decisions (Ahmet et al. 2014). A long-term mathematical model was developed for the implementation of investment planning in Argentina for the period 2010–2030 (Flores et al. 2014). This approach is based on traditional long-term planning approach and does not include: (i) short-term operational constraints, (ii) chronological load duration curve, i.e., total demand in each year has to be satisfied

(one number), and (iii) environmental policy is not considered. An MILP long-term investment planning model was developed taking into consideration life cycle costs and environmental impacts (Barteczko-Hibbert et al. 2014). However, the inherent variability of RET through the introduction of operational constraints and chronological load demand are not considered in the study. An MILP model for the optimal GEP was developed and the authors employed scenario-based analysis to demonstrate the influence of carbon credit prices and target rates of electricity demand (Han and Lee 2011). Their model considers the availability of energy resources. A multi-regional, multi-period power systems planning model was developed to determine the optimal energy roadmap of a power system (Koltsaklis et al. 2014).

Regarding the UCP, an MILP model was developed for the solution of the UCP and applied on the Belgian power system (Delarue et al. 2009), while the same mathematical model was employed for the investigation of the effects of wind power on the power generation costs and the amount of CO_2 emissions (Delarue et al. 2009). A two-stage stochastic programming formulation of the UCP was developed in order to assess the spinning and non-spinning reserve levels of a large scale power system (Californian test system) along with the costs associated in view of increased wind power penetration (Xiao et al. 2011). A security constrained UCP was developed modelling the Greek power market, and highlighting also the most important aspects of the wholesale energy and ancillary services market structure (Andrianesis et al. 2011).

Operational flexibility constitutes an aspect that should be considered in power systems planning due to the fact that RET (e.g., wind and solar) are characterized by inherent intermittency and frequent fluctuations (Palmintier 2013). During the last three decades, there has been a significant increase in the wind power capacity worldwide from 10 MW in 1980 to 282 GW by the end of 2012. Further penetration of wind turbines depends on several factors including wind power resources potential, existing power mix, relative economic competitiveness of each technology, energy policy (e.g., CO_2 emission pricing), spatial installation of wind turbines and other societal institutional parameters (Timilsina et al. 2013). A study presented a review analysis in order to identify the best possible strategies for the effective large-scale integration of intermittent renewables, and particularly wind power, within the existing power systems. These options include additional power capacity reserves, enhancement of cross-border interconnections and distributed energy storage technologies such as electric vehicles (Purvins et al. 2011).

The impact of regional distribution of wind turbines on power variations is also analyzed. As wind power increases its share in the power generation mix, mid and peak load thermal units enhance their competitiveness against base-load power units (e.g., coal and lignite units) due to the fact that the residual load demand (initial load minus RET power generation) is lower and characterized by significant fluctuations (Belderbos and Delarue 2015). An overview of the available computer tools that can be used to analyze the integration of RET into various energy systems was presented covering a wide range of approaches and technical details (Connolly et al. 2010). A study illustrated that the increased penetration of onshore wind

power into the Portuguese power system can enhance the system's flexibility and does not have negative impacts on the security of electricity supply (Gouveia et al. 2014). This is in line with the design of a low-carbon energy economy, being the cornerstone of the European energy policy. The EnergyPLAN tool was employed in order to identify possible moderation strategies employed in the Serbian energy system to achieve least-cost options for increasing penetration of wind power (Batas Bjelić et al. 2013). The impacts that renewable power injections have on the CO_2 emissions allowances price and on the amount of CO_2 emissions in 12 European Union (EU) Member States (Southern and Western Europe) were also studied (Van den Bergh et al. 2013). An analysis was also presented in order for the Irish energy system to achieve a 100 % penetration of RET utilizing the EnergyPLAN tool (Connolly et al. 2011). Instead of focusing only on the power sector, the study incorporates also the heat and transport sectors to further integrate the deployment of RET. However, the results obtained are not optimal since they are based on numerous assumptions regarding the utilization of several RET.

The main contributions and the salient features of our work presented in this chapter include: (i) analytical and detailed representation of the day-ahead energy and reserve markets inside the long-term planning horizon, (ii) short-term-based strategic investment decisions, (iii) ability to capture the inherent variability and fluctuations of RET at an hourly level and utilization of that information in both planning and scheduling decisions, and (iv) integration between energy policy targets (maximum CO_2 emissions, RET penetration target) and daily production scheduling.

The remainder of this chapter is organized as follows: Sect. 15.2 provides the problem statement, followed by a generic mathematical formulation of the studied problem in Sect. 15.3. A description of the case study including data and assumptions is given in Sect. 15.4, while an analytical discussion of the results obtained is provided in Sect. 15.5. Finally, Sect. 15.6 draws upon some concluding remarks.

15.2 Problem Statement

The problem that is considered in this work can be formally stated as follows:

- The total planning horizon is split into sets of years $y \in Y$, months $m \in M$, and hours $t \in T$. A representative day (24 h) per month and year has been selected for the solution of the combined GEP and UCP. The duration of each month (in days) is provided by DUR_m.
- The power system under consideration is divided into a number of sectors $(s, s') \in S$ which are further split into a certain number of zones $z \in Z$, taking into consideration the interconnection structure among them $(s \in S^{s'})$. Maximum corridor flow between each interconnected pair of sectors is given by $CFL_{s,s',y}$.

There is also the option of possible interconnections $nc \in NC$ with the power grids of neighbouring countries subject to their interconnections or not with specific sectors $nc \in NC^S$ (or zones $nc \in NC^Z$).

- A set of power generating units $i \in I$ is (or can be) installed in each sector $i \in I^S$ (or zone $i \in I^Z$). The installed capacity of each unit is given by $IC_{i,y}$. These units are identified as existing (already installed) units $i \in I^{EX}$ and new candidate units to be installed or not $i \in I^{NEW}$. Both new and existing units are further divided into thermal $i \in I^{TH}$ and $i \in I^{RES}$ renewable units. Set $i \in I^{RES-}$ refers to renewable units without including hydroelectric units.
- Three available start-up types $a \in A$ {A: hot, warm, and cold start-up type} are considered for each thermal unit $i \in I^{TH}$ based on its non-operational time (after being shut-down) T_sd_i. There are specific time periods, $T\_sd_i^{1,2}$, after which each thermal unit $i \in I^{TH}$ changes standby condition, including the non-operational time of each unit before going from hot to warm standby condition, $T\_sd_i^{warm}$, and that before going from warm to cold standby condition, $T\_sd_i^{cold}$, as presented in Table 15.1. After the determination of the appropriate start-up type, the synchronization and soak phases follow respectively, with a duration of $T\_sn_i^a$ and $T\_sk_i^a$. Total hourly time periods also include extended time intervals in the past (T_sd^{prior}) in order for the model to determine the system's technical operation during the beginning of the studied time period (representative day of each month in each year), i.e., start-up and shut-down decisions, ramp rate constraints, etc.
- After the completion of the soak phase, each thermal unit $i \in I^{TH}$ enters the dispatchable phase, while the last operational stage is that of desynchronization, having a duration of T_dn_i hours. Synchronization, soak, dispatch, and desynchronization comprise the operational stages of each thermal unit $i \in I^{TH}$. Note that the total operational time of each unit must be greater than or equal to the unit's minimum up time, T_up_i, while unit's non-operational time (after being shut down) must be greater than or equal to minimum down time, T_down_i, to be able to start-up again. A specific cost is also related to the shut-down decision of each thermal unit $i \in I^{TH}, CSD_i$.
- The capacity of each hydrothermal unit $i \in I^{HT}$, $PBL_{i,bl,y,m,t}$, is divided into a number of blocks $bl \in BL$ to represent more accurately and realistically the

Table 15.1 Modelling of thermal units' start-up type

Start-up type	Non-operational time (after being shut-down), $(T\_sd_i^{min} \le T\_sd_i \le T\_sd_i^{max})$	Synchronization time (h)	Soak time (h)
Hot	$0 \le T\_sd_i < T\_sd_i^{warm}$	$T\_sn_i^{hot}$	$T\_sk_i^{hot}$
Warm	$T\_sd_i^{warm} \le T\_sd_i < T\_sd_i^{cold}$	$T\_sn_i^{warm}$	$T\_sk_i^{warm}$
Cold	$T\_sd_i^{cold} \le T\_sd_i < T\_sd_i^{prior}$	$T\_sn_i^{cold}$	$T\_sk_i^{cold}$

technical characteristics of each unit. For each capacity block and during each time period, every unit offers a specific amount of energy at a certain price (marginal cost of the energy offer function), $CPB_{i,bl,y,m,t}$, to contribute to the power demand satisfaction. Apart from the power output of each unit $i \in I$ to be determined by the optimization process, there is also a fixed (non-priced) power generation in each time period, $PFX_{i,y,m,t}$, including mandatory hydro injection, renewables power injection, and power production from units operating under commissioning status.

- With regard to power imports and exports, power transfer capacities per interconnection $nc \in NC$, are split into certain blocks in each time period, $Imblock_{nc,bl,y,m,t}$ and $Exblock_{nc,bl,y,m,t}$ respectively, having a marginal energy offer function (for imports), $CIMPB_{nc,bl,y,m,t}$, and a marginal export bid (for exports), $CEXPB_{nc,bl,y,m,t}$, per block and time period. Fixed operational and maintenance (O&M) cost of each unit $i \in I^{RES-}$, FOM_i, is also taken into account (the corresponding one of each hydrothermal unit $i \in I^{HT}$ is incorporated in the marginal cost of its energy offer function).
- Regarding power injections (electricity generation and imports), there is an injection losses coefficient in order to calculate net power injections at the market point, $INL_{z,y,m,t}$, depending on each zone and the load level of each time period.
- Each hydrothermal unit $i \in I^{HT}$ has specific technical characteristics: maximum (minimum) power output, $PMAX_i^{dp}$ $\left(PMIN_i^{dp}\right)$ when committed, or $PMAX_i^{sc}$ $\left(PMIN_i^{sc}\right)$ when providing secondary reserve, as well as PSK_i when operating in soak phase (only for thermal units $i \in I^{TH}$). For each capacity block $bl \in BL$, every thermal unit $i \in I^{TH}$ is characterized by a certain emission factor, $CO_2\_EF_{i,bl}$. Each unit is also identified based on its ramp-up and -down rates, RR_up_i and RR_down_i, when committed, as well as on its ramp rates, RR_sec_i, when providing secondary reserve.
- Power system's requirements include: (i) power load requirements in each sector and time period, $Dem_{s,y,m,t}$, (ii) primary-up reserve requirements in each time period, $R1Req_{y,m,t}$, (iii) secondary-up, $R2Req_{y,m,t}^{up}$, and secondary-down, $R2Req_{y,m,t}^{down}$, reserve requirements in each time period, (iv) fast secondary-up, $FastR2Req_{y,m,t}^{up}$, and fast secondary-down, $FastR2Req_{y,m,t}^{down}$, reserve requirements in each time period, and (v) tertiary reserve requirements in each time period, $R3Req_{y,m,t}$.
- Regarding reserve provision capabilities, each hydrothermal unit $i \in I^{HT}$ is characterized by: (i) maximum contribution of each unit in primary reserve, $R1_i$, (ii) maximum contribution of each unit in secondary reserve, $R2_i$, (iii) maximum contribution of each unit in tertiary spinning, $R3_i^{sp}$, and non-spinning reserve,

$R3_i^{nsp}$. Each unit's energy reserve offer has a specific price, i.e., $CR1_{i,y,m,t}$ for the primary energy reserve, and $CR2_{i,y,m,t}$ for the secondary range energy offer. Tertiary energy offer is non-priced.

- Each RES unit $i \in I^{RES}$ is characterized by a specific availability factor in each zone and time period, $AV_{i,z,m,t}$. It is assumed that this factor does not change from year to year, since it is only dependent on the meteorological conditions of each representative day and month. There is also a yearly unavailability factor for each thermal unit $i \in I^{TH}$, $EFOR_{i,y}$, electricity interconnections, i.e., imports $EFORIM_{nc,y}$.
- System's stability is also considered by taking into account minimum peak reserve requirements in each time period, RSV, and maximum yearly RES penetration in the power generation mix, $MRES_y$.
- For the new candidate units $i \in I^{NEW}$, there is a specific investment cost for the construction and installation of each unit in each year, $INVC_{i,y}$. A capital recovery factor, CRF_i, is considered to calculate the annualized cost of each unit $i \in I^{NEW}$. The construction time of each unit $i \in I^{NEW}$, T_con_i, is also taken into consideration, as well as maximum allowable yearly capacity bounds of each unit $i \in I^{RES-}$ in each zone, $RES\_CAP_{i,z,y}$.
- Transition towards an environmentally friendly power generation mix is incorporated through the use of a cap on CO_2 emissions in each year, $CO_2\_CAP_y$, a RES penetration target in the power generation mix, REN_y, and CO_2 emission pricing being incorporated in the marginal cost of the energy offer function of each unit, $CPB_{i,bl,y,m,t}$.

15.3 Mathematical Formulation

15.3.1 Objective Function

The objective function to be optimized concerns the minimization of the total discounted cost for the development and operation of the studied power system, including: (i) total variable operating (marginal) cost of the installed power units incorporating fuel cost, variable operating and maintenance (O&M) cost, and CO_2 emissions cost, (ii) electricity imports cost, (iii) electricity exports revenues, (iv) thermal units' shut-down cost, (v) total reserves provision cost, (vi) total new units' investment cost, and (vii) total fixed O&M cost of renewable units (not including hydro units), as provided by Eq. (15.1).

$$
\begin{aligned}
Min\ Cost = & \overbrace{\sum_{i\in(I^{HT}\bigcap I^{z})}\sum_{z\in Z}\sum_{bl\in BL}\sum_{y\in Y}\sum_{m\in M}\sum_{t\in T}\left(pb_{i,bl,y,m,t}\cdot CPB_{i,bl,y,m,t}\cdot INL_{z,y,m,t}\cdot DUR_{m}\right)}^{Variable\ operating(marginal)cost} \\
& + \overbrace{\sum_{nc\in NC^{z}}\sum_{z\in Z}\sum_{bl\in BL}\sum_{y\in Y}\sum_{m\in M}\sum_{t\in T}\left(impb_{nc,bl,y,m,t}\cdot CIMPB_{nc,bl,y,m,t}\cdot INL_{z,y,m,t}\cdot DUR_{m}\right)}^{Electricity\ imports cost} \\
& - \overbrace{\sum_{nc\in NC^{z}}\sum_{z\in Z}\sum_{bl\in BL}\sum_{y\in Y}\sum_{m\in M}\sum_{t\in T}\left(expb_{nc,bl,y,m,t}\cdot CEXPB_{nc,bl,y,m,t}\cdot DUR_{m}\right)}^{Electricity\ exports\ revenues} \\
& + \overbrace{\sum_{i\in I^{TH}}\sum_{y\in Y}\sum_{m\in M}\sum_{t\in T}\left(w_{i,y,m,t}\cdot CSD_{i}\cdot DUR_{m}\right)}^{Shut-down\ cost} \\
& + \overbrace{\sum_{i\in I^{HT}}\sum_{y\in Y}\sum_{m\in M}\sum_{t\in T}\left[\left\{\left(r1^{up}_{i,y,m,t}\cdot CR1_{i,y,m,t}\right)+\left(r2^{up}_{i,y,m,t}+r2^{down}_{i,y,m,t}\right)\cdot CR2_{i,y,m,t}\right\}\cdot DUR_{m}\right]}^{Reserves\ provision\ cost} \\
& + \overbrace{\sum_{y'\le y\in Y}\sum_{i\in I^{NEW}}CRF_{i}\cdot PMIN^{dp}_{i}\cdot cn_{i,y'}\cdot INVC_{i,y'}}^{Investment\ capital\ cost} + \overbrace{\sum_{i\in I^{RES-}}\sum_{y\in Y}c_{i,y}\cdot FOM_{i}}^{Fixed\ O\&M\ cost\ of\ renewable\ units}
\end{aligned}
\tag{15.1}
$$

15.3.2 Model Constraints

All the possible operating modes of a thermal unit $i\in I^{TH}$ in the unit commitment problem (UCP) including start-up, synchronization, soak, desynchronization, as well as minimum up and down time constraints, are described by constraints (15.2)–(15.12). More specifically:

15.3.2.1 Start-up Constraints

$$
x\_st^{a}_{i,y,m,t}\le\sum_{t'=t-T\_sd^{max}_{i}+1}^{t-T\_sd^{min}_{i}}w_{i,y,m,t'}\quad\forall a\in A,i\in I^{TH},y,m,t \tag{15.2}
$$

$$
x\_st_{i,y,m,t}=\sum_{a\in A}x\_st^{a}_{i,y,m,t}\quad\forall i\in I^{TH},y,m,t \tag{15.3}
$$

After being shut-down $\left(x_{i,y,m,t}=0\right)$ for T_sd_i hours (T_sd_i must be greater than or equal to the minimum down time of each unit $i\in I^{TH}$, T_down_i), and depending on the duration of that time, T_sd_i (see Table 15.1), an appropriate start-type decision is selected for each unit $i\in I^{TH}$ in each time period, as described by constraints (15.2). Equation (15.3) ensure that a thermal unit $i\in I^{TH}$ is able to start-up with only one start-up type $a\in A$ at each time period.

15.3.2.2 Synchronization Phase Constraints

$$x\_sn^{a}_{i,y,m,t} \leq \sum_{t'=t-T\_sn^{a}_{i}+1}^{t} x\_st^{a}_{i,y,m,t'} \quad \forall a \in A, i \in I^{TH}, y, m, t \tag{15.4}$$

$$x\_sn_{i,y,m,t} = \sum_{a \in A} x\_sn^{a}_{i,y,m,t} \quad \forall i \in I^{TH}, y, m, t \tag{15.5}$$

Once the appropriate start-up decision is taken ($x\_st^{a}_{i,ym,t} = x\_st_{i,y,m,t} = 1$ for a certain $a \in A$), the thermal unit $i \in I^{TH}$ enters the synchronization operational stage ($x\_sn^{a}_{i,ym,t} = x\_sn_{i,y,m,t} = 1$ for a certain $a \in A$) having a duration of $T\_sn^{a}_{i}$ hours, as represented by constraints (15.4) and (15.5). Note that the power output of each unit $i \in I^{TH}$ during the synchronization phase is zero.

15.3.2.3 Soak Phase Constraints

$$x\_sk^{a}_{i,y,m,t} \leq \sum_{t'=t-T\_sn^{a}_{i}-T\_sk^{a}_{i}+1}^{t-T\_sn^{a}_{i}} x\_st^{a}_{i,y,m,t'} \quad \forall a \in A, i \in I^{TH}, y, m, t \tag{15.6}$$

$$x\_sk_{i,y,m,t} = \sum_{a \in A} x\_sk^{a}_{i,y,m,t} \quad \forall i \in I^{TH}, y, m, t \tag{15.7}$$

$$p\_sk_{i,y,m,t} = PSK_i \cdot x\_sk_{i,y,m,t} \quad \forall i \in I^{TH}, y, m, t \tag{15.8}$$

After completing the synchronization stage, the thermal unit $i \in I^{TH}$ enters the soak operating stage ($x\_sk^{a}_{i,ym,t} = x\_sk_{i,y,m,t} = 1$ for a specific $a \in A$), having a duration of $T\_sk^{a}_{i}$ hours and during which unit's power output equals a fixed value, PSK_i, according to constraints (15.6)–(15.8).

15.3.2.4 Desynchronization Phase Constraints

$$x\_dn_{i,y,m,t} = \sum_{t'=t+1}^{t+T\_dn_i-1} w_{i,y,m,t'} \quad \forall i \in I^{TH}, y, m, t \tag{15.9}$$

$$p\_dn_{i,y,m,t} = \sum_{t'=t}^{t+T\_dn_i-1} x\_dn_{i,y,m,t} \cdot (t' - t) \cdot \frac{PMIN^{dp}_{i}}{T\_dn_i} \quad \forall i \in I^{TH}, y, m, t \tag{15.10}$$

The next operational stage of each unit $i \in I^{TH}$ is that of dispatchable stage $\left(x\_dp_{i,y,m,t} = 1\right)$, with a duration of T_dp_i hours and during which unit's power

output is able to range from its technical minimum, $PMIN_i^{dp}$, to its maximum, $PMAX_i^{dp}$, subject to technical up and down ramp rates (RR_up_i and RR_down_i correspondingly) contributing to the power demand satisfaction, $Dem_{s,y,m,t}$, and to reserve requirements, if selected. Finally, the last possible operational stage of a unit $i \in I^{TH}$ is that of desynchronization, having a duration of T_dn_i hours, as described by constraints (15.9). Equation (15.10) ensure that unit's power output decreases with a particular power sequence during that stage $\left(x\_dn_{i,y,m,t} = 1\right)$.

15.3.2.5 Minimum Up and Down Time Constraints

Constraints (15.11) and (15.12) describe the minimum up and down times of each hydrothermal unit $i \in I^{HT}$ correspondingly, i.e., a thermal unit $i \in I^{TH}$ must remain operational (or non-operational) in each time period if it has started-up (or shut-down) during the previous $(T\_up_i - 1)$ (or $T\_down_i - 1$) hours respectively. In other words, the sum of $T\_sn_i + T\_sk_i + T\_dp_i + T\_dn_i$ must be greater than or equal to the minimum up time of each unit $i \in I^{TH}$, T_up_i, and the unit's non-operational time (after being shut-down), T_sd_i, must be greater than or equal to the minimum down time of each unit $i \in I^{TH}$, T_down_i.

$$\sum_{t' = t - T\_up_i + 1}^{t} x\_st_{i,y,m,t'} \le x_{i,y,m,t} \quad \forall i \in I^{TH}, y, m, t \tag{15.11}$$

$$\sum_{t' = t - T\_down_i + 1}^{t} w_{i,y,m,t'} \le 1 - x_{i,y,m,t} \quad \forall i \in I^{TH}, y, m, t \tag{15.12}$$

15.3.2.6 Logical Status of Unit Commitment

Constraints (15.13) ensure that each hydrothermal unit $i \in I^{HT}$ can only be at one of the possible operating states in each time period when committed. Note that $x\_sn_{i,y,m,t} = x\_sk_{i,y,m,t} = x\_dn_{i,y,m,t} = 0$ for a hydro unit. Constraints (15.14) and (15.15) model the start-up $\left(x\_st_{i,y,m,t}\right)$ and shut-down $\left(w_{i,y,m,t}\right)$ decision of each unit $i \in I^{HT}$ during each time period. According to constraints (15.15), a start-up and a shut-down decision cannot be taken simultaneously in each time period for a hydrothermal unit $i \in I^{HT}$. Constraints (15.16) ensure that a hydrothermal unit $i \in I^{HT}$ can provide secondary reserve if and only if operates in the dispatchable phase.

$$x_{i,y,m,t} = x\_sn_{i,y,m,t} + x\_sk_{i,y,m,t} + x\_dp_{i,y,m,t} + x\_dn_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.13}$$

$$x\_st_{i,y,m,t} - w_{i,y,m,t} = x_{i,y,m,t} - x_{i,y,m,t-1} \quad \forall i \in I^{HT}, y, m, t \tag{15.14}$$

$$x\_st_{i,y,m,t} + w_{i,y,m,t} \leq 1 \quad \forall i \in I^{HT}, y, m, t \tag{15.15}$$

$$x\_sc_{i,y,m,t} \leq x\_dp_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.16}$$

15.3.2.7 Correlation Between Design and Operational Decisions

Constraints (15.17) correlate design and operational decisions, i.e., a unit $i \in I^{NEW}$ is not operationally available unless the decision for its construction has been taken $\left(cn_{i,y} = 1\right)$ and its installation has been completed, requiring T_con_i years.

$$x_{i,y,m,t} \leq \sum_{y' \leq y} cn_{i,y'-T\_con_i} \quad \forall i \in I^{NEW}, y, m, t \tag{15.17}$$

15.3.2.8 Power Output Constraints

Equation (15.18) describe the power output of each hydrothermal unit $i \in I^{HT}$ being divided into two parts: (i) fixed (non-priced) component accounting for power mandatory hydro injection and/or hydrothermal power injection from units being under commissioning status in each time period $\left(PFX_{i,y,m,t}\right)$, and (ii) priced component based on the energy offer per power capacity block of each unit $i \in I^{HT}$ in each time period $\left(pb_{i,bl,y,m,t}\right)$. Fixed (non-priced) renewable energy technologies (without including hydroelectric units) power injection, based on the availability of each renewable energy resource (e.g., wind, solar) in each zone and time period is provided by Eq. (15.19). Constraints (15.20) ensure that the segment of each block $bl \in BL$ of each existing hydrothermal unit's marginal cost function dispatched in each time period, $pb_{i,bl,y,m,t}$, must not exceed the size of the corresponding step of unit's marginal cost function. Constraints (15.21) ensure the above requirement for a new unit $i \in I^{NEW}$, i.e., the constraints are subject to the decision for its construction or not $\left(cn_{i,y}\right)$, and to the time period required for its construction $\left(T\_con_i\right)$.

$$p_{i,y,m,t} = PFX_{i,y,m,t} + \sum_{bl \in BL} pb_{i,bl,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.18}$$

$$p_{i,y,m,t} = AV_{i,z,m,t} \cdot c_{i,y} \quad \forall i \in I^{RES-}, z, y, m, t \tag{15.19}$$

$$pb_{i,bl,y,m,t} \leq PBL_{i,bl,y,m,t} \quad \forall i \in I^{EX}, bl, y, m, t \tag{15.20}$$

$$pb_{i,bl,y,m,t} \leq PBL_{i,bl,y,m,t} \cdot \sum_{y' \leq y} cn_{i,y'-T\_con_i} \quad \forall i \in I^{NEW}, bl, y, m, t \tag{15.21}$$

15.3.2.9 Imports and Exports Constraints

Equations (15.22) and (15.23) describe the power injection (withdrawal) of imports (exports) from (to) neighbouring country $nc \in NC$ during each time period correspondingly. Equations (15.24) and (15.25) ensure that the segment of each block $bl \in BL$ of each interconnected country's energy imports (exports) marginal cost function in each time period, $impb_{nc,bl,y,m,t}$ $\left(expb_{nc,bl,y,m,t}\right)$, must not exceed the size of the corresponding step of each interconnected country's energy imports (exports) marginal cost function, $Imblock_{nc,bl,y,m,t}$ $\left(Exblock_{nc,bl,y,m,t}\right)$.

$$imp_{nc,y,m,t} = \sum_{bl \in BL} impb_{nc,bl,y,m,t} \quad \forall nc, y, m, t \tag{15.22}$$

$$exp_{nc,y,m,t} = \sum_{bl \in BL} expb_{nc,bl,y,m,t} \quad \forall nc, y, m, t \tag{15.23}$$

$$impb_{nc,bl,y,m,t} \le Imblock_{nc,bl,y,m,t} \quad \forall nc, bl, y, m, t \tag{15.24}$$

$$expb_{nc,bl,y,m,t} \le Exblock_{nc,bl,y,m,t} \quad \forall nc, bl, y, m, t \tag{15.25}$$

15.3.2.10 Net Power Injections

Equations (15.26) and (15.27) define net power injections from both units $i \in I^Z$ and interconnected countries $nc \in NC^z$ (imports) by taking into account the injection losses coefficient of each zone $z \in Z$ in each time period. Constraints (15.28)–(15.30) specify the power output limits of each hydrothermal unit $i \in I^{HT}$ in each operating stage. The first three terms on the right-hand side of these constraints define unit's power output during synchronization, soak, and desynchronization phases.

$$p\_inj_{i,y,m,t} = INL_{z,y,m,t} \cdot p_{i,y,m,t} \quad \forall i \in I^Z, z, y, m, t \tag{15.26}$$

$$im\_inj_{nc,y,m,t} = INL_{z,y,m,t} \cdot imp_{nc,y,m,t} \quad \forall nc \in NC^z, z, y, m, t \tag{15.27}$$

$$\begin{aligned} p_{i,y,m,t} - r2^{down}_{i,y,m,t} \ge 0 \cdot x\_sn_{i,y,m,t} + p\_sk_{i,y,m,t} + p\_dn_{i,y,m,t} \\ + PMIN^{dp}_i \cdot \left(x\_dp_{i,y,m,t} - x\_sc_{i,y,m,t}\right) + PMIN^{sc}_i \cdot x\_sc_{i,y,m,t} \end{aligned} \quad \forall i \in I^{HT}, y, m, t \tag{15.28}$$

$$\begin{aligned} p_{i,y,m,t} + r2^{up}_{i,y,m,t} \ge 0 \cdot x\_sn_{i,y,m,t} + p\_sk_{i,y,m,t} + p\_dn_{i,y,m,t} \\ + PMAX^{dp}_i \cdot \left(x\_dp_{i,y,m,t} - x\_sc_{i,y,m,t}\right) + PMAX^{sc}_i \cdot x\_sc_{i,y,m,t} \end{aligned} \quad \forall i \in I^{HT}, y, m, t \tag{15.29}$$

$$\begin{aligned} p_{i,y,m,t} + r1^{up}_{i,y,m,t} + r2^{up}_{i,y,m,t} + r3^{sp}_{i,y,m,t} \ge 0 \cdot x\_sn_{i,y,m,t} \\ + p\_sk_{i,y,m,t} + p\_dn_{i,y,m,t} + PMAX^{dp}_i \cdot x\_dp_{i,y,m,t} \end{aligned} \quad \forall i \in I^{HT}, y, m, t \tag{15.30}$$

15.3.2.11 Reserve Constraints

Constraints (15.31) set the upper bounds of primary-up reserve of hydrothermal unit $i \in I^{HT}$, subject to the decision of its operation (or not) in the dispatchable phase. Constraints (15.32) ensure that the sum of contributions of unit $i \in I^{HT}$ in secondary-up and down reserves, on condition that it provides secondary reserve $\left(x\_sc_{i,y,m,t}\right)$, must be less than or equal to the maximum allowable secondary reserve contribution of unit $i \in I^{HT}$. Constraints (15.33) and (15.34) specify the upper limits of tertiary spinning and non-spinning reserves of hydrothermal unit $i \in I^{HT}$ respectively, subject to the decision of its operation (or not) in the dispatchable phase $\left(x\_dp_{i,y,m,t}\right)$, and to the decision for the provision (or not) of tertiary non-spinning reserve in each time period $(x\_3ns_{i,y,m,t})$ correspondingly. Constraints (15.35) guarantee that the contribution of hydrothermal unit $i \in I^{HT}$ in the tertiary non-spinning reserve is greater than or equal to the unit's minimum power output $\left(PMIN_i^{dp}\right)$. Constraints (15.36) ensure that a hydrothermal unit $i \in I^{HT}$ can only provide tertiary non-spinning reserve, if and only if is non-operational. Equation (15.37) states that the total contribution of each unit $i \in I^{HT}$ in tertiary reserve, equals the sum of its provision in tertiary spinning and non-spinning reserves. Constraints (15.38) ensure that provision of each unit $i \in I^{HT}$ in fast secondary-up reserve within a period of 1 min must be less than or equal to its provision in secondary-up reserve. Moreover, constraints (15.39) set an upper limit of the fast secondary-up reserve. Constraints (15.40) and (15.41) express the same limits for the fast secondary-down reserve of each unit $i \in I^{HT}$ as constraints (15.38) and (15.39) for fast secondary-up reserve respectively.

$$r1_{i,y,m,t}^{up} \le R1_i \cdot x\_dp_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.31}$$

$$r2_{i,y,m,t}^{up} + r2_{i,y,m,t}^{down} \le R2_i \cdot x\_sc_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.32}$$

$$r3_{i,y,m,t}^{sp} \le R3_i^{sp} \cdot x\_dp_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.33}$$

$$r3_{i,y,m,t}^{nsp} \le R3_i^{nsp} \cdot x\_3ns_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.34}$$

$$r3_{i,y,m,t}^{nsp} \ge PMIN_i^{dp} \cdot x\_3ns_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.35}$$

$$x\_3ns_{i,y,m,t} \le 1 - x_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.36}$$

$$r3_{i,y,m,t} = r3_{i,y,m,t}^{sp} + r3_{i,y,m,t}^{nsp} \quad \forall i \in I^{HT}, y, m, t \tag{15.37}$$

$$fastr2_{i,y,m,t}^{up} \le r2_{i,y,m,t}^{up} \quad \forall i \in I^{HT}, y, m, t \quad \forall i \in I^{HT}, y, m, t \tag{15.38}$$

$$fastr2^{up}_{i,y,m,t} \leq RR\_sec_i \cdot 1\,min \quad \forall i \in I^{HT}, y, m, t \tag{15.39}$$

$$fastr2^{down}_{i,y,m,t} \leq r2^{down}_{i,y,m,t} \quad \forall i \in I^{HT}, y, m, t \tag{15.40}$$

$$fastr2^{down}_{i,y,m,t} \leq RR\_sec_i \cdot 1\,min \quad \forall i \in I^{HT}, y, m, t \tag{15.41}$$

15.3.2.12 System Constraints

Constraints (15.42)–(15.47) describe system requirements for all reserve types, i.e., primary-up, secondary-up and down, tertiary spinning and non-spinning, as well as fast secondary-up and down correspondingly.

$$\sum_{i \in I^{HT}} r1^{up}_{i,y,m,t} \geq R1Req_{y,m,t} \quad \forall y, m, t \tag{15.42}$$

$$\sum_{i \in I^{HT}} r2^{up}_{i,y,m,t} \geq R2Req^{up}_{y,m,t} \quad \forall y, m, t \tag{15.43}$$

$$\sum_{i \in I^{HT}} r2^{down}_{i,y,m,t} \geq R2Req^{down}_{y,m,t} \quad \forall y, m, t \tag{15.44}$$

$$\sum_{i \in I^{HT}} r3_{i,y,m,t} \geq R3Req_{y,m,t} \quad \forall y, m, t \tag{15.45}$$

$$\sum_{i \in I^{HT}} fastr2^{up}_{i,y,m,t} \geq FastR2Req^{up}_{y,m,t} \quad \forall y, m, t \tag{15.46}$$

$$\sum_{i \in I^{HT}} fastr2^{down}_{i,y,m,t} \geq FastR2Req^{down}_{y,m,t} \quad \forall y, m, t \tag{15.47}$$

15.3.2.13 Ramp Limits

Each hydrothermal unit $i \in I^{HT}$ has the capability of increasing or decreasing its power output according to specific ramp rates in each time period, the values of which are dependent on the unit's operational mode, i.e., if it provides secondary reserve or not $\left(x\_sc_{i,y,m,t}\right)$. Constraints (15.48) and (15.49) describe the ramp rate limits, up and down respectively, of each hydrothermal unit $i \in I^{HT}$ in each time period.

$$p_{i,y,m,t} - p_{i,y,m,t-1} \leq (1 - x\_sc_{i,y,m,t}) \cdot RR\_up_i + RR\_sec_i \cdot x\_sc_{i,y,m,t} \cdot 60 \quad \forall i \in I^{HT}, y, m, t \tag{15.48}$$

$$p_{i,y,m,t-1} - p_{i,y,m,t} \leq (1 - x\_sc_{i,y,m,t}) \cdot RR\_down_i + RR\_sec_i \cdot x\_sc_{i,y,m,t} \cdot 60 \quad \forall i \in I^{HT}, y, m, t \tag{15.49}$$

15.3.2.14 Corridor Limits

Constraints (15.50) ensure that the corridor flow between two interconnected sectors $s, s^{'} \in S^{s}$, must not exceed the maximum available capacity of the corridor $\left(CFL_{s,s^{'},y}\right)$ in each time period.

$$cf_{s,s^{'},y,m,t} \leq CFL_{s,s^{'},y} \quad \forall\left(s,s^{'}\right) \in S^{s}, y, m, t \tag{15.50}$$

15.3.2.15 Demand Balance

Equation (15.51) describe the demand balance of the power system. More specifically, net power injection from all the installed power units $\left(\sum_{i \in I^{s}} p\_inj_{i,y,m,t}\right)$ plus net electricity flow rates $\left(\sum_{s^{'} \in S^{s}} cf_{s^{'},s,y,m,t} - \sum_{s^{'} \in S^{s}} cf_{s,s^{'},y,m,t}\right)$ and net electricity imports $\left(\sum_{nc \in NC^{s}} im\_inj_{nc,y,m,t} - \sum_{nc \in NC^{s}} exp_{nc,y,m,t}\right)$ to each sector, must be equal to the sector's electricity demand $Dem_{s,y,m,t}$.

$$\begin{aligned} &\sum_{i \in I^{s}} p\_inj_{i,y,m,t} + \sum_{nc \in NC^{s}} im\_inj_{nc,y,m,t} + \sum_{s^{'} \in S^{s}} cf_{s^{'},s,y,m,t} \\ &\quad = \sum_{s^{'} \in S^{s}} cf_{s,s^{'},y,m,t} + \sum_{nc \in NC^{s}} exp_{nc,y,m,t} + Dem_{s,y,m,t} \end{aligned} \quad \forall s, y, m, t \tag{15.51}$$

Design Constraints

15.3.2.16 Additional Capacity Limits

Constraints (15.52)–(15.54) define the additional capacity limits. More specifically, constraints (15.52) state that the capacity of the newly-built unit $i \in I^{NEW}$, $nc_{i,y}$, will be available for the first time (and equal to its nominal capacity, $Pmax_{i}^{dp}$), only after the required unit's time construction (T_con_{i}), and subject to the decision for its construction or not $\left(cn_{i,y}\right)$. Equation (15.53) ensure that once the newly-built unit $i \in I^{NEW}$, $nc_{i,y}$, will be available for the first time, it will remain available until the end of the studied period. Equation (15.54) state that the available capacity of each existing unit $i \in I^{EX}$ amounts to the unit's installed capacity, subject to each unit's decommissioning plan. Constraints (15.55) enforce the construction of each unit candidate unit to be implemented at most once during the whole planning horizon.

$$nc_{i,y+T\_con_i} = PMAX_i^{dp} \cdot cn_{i,y} \quad \forall i \in I^{NEW}, y \tag{15.52}$$

$$c_{i,y} = \sum_{y' \le y} nc_{i,y'} \quad \forall i \in I^{NEW}, y \tag{15.53}$$

$$c_{i,y} = IC_{i,y} \quad \forall i \in I^{EX}, y \tag{15.54}$$

$$\sum_{y \in Y} cn_{i,y} \le 1 \quad \forall i \in I^{NEW} \tag{15.55}$$

15.3.2.17 Reserve Margin

Constraints (15.56) refer to reserve margin constraints, i.e., the installed capacities of all units being available in each time period $\left(\sum_{i \in I^{TH}} c_{i,y} \cdot (1 - EFOR_{i,y}) + \sum_{i \in (I^{RES} \bigcap I^z)} \sum_{z \in Z} c_{i,y} \cdot AV_{i,z,m,t} \right)$, plus the available imports $\sum_{nc \in NC} CIM_{nc,y} \cdot (1 - EFORIM_{nc,y})$ must be adequate enough to meet the total projected load $\left(\sum_{s \in S} Dem_{s,y,m,t} \right)$ plus an allowance (RSV) above that level for security reasons

$$\sum_{i \in I^{TH}} c_{i,y} \cdot (1 - EFOR_{i,y}) + \sum_{i \in (I^{RES} \cap I^z)} \sum_{z \in Z} c_{i,y} \cdot AV_{i,z,m,t} + \sum_{nc \in NC} CIM_{nc,y} \cdot (1 - EFORIM_{nc,y}) \ge (1 + RSV) \cdot \sum_{s \in S} Dem_{s,y,m,t} \quad \forall y, m, t \tag{15.56}$$

15.3.2.18 Annual Generation Limits

Constraints (15.57) and (15.58) refer to annual allowable power generation levels of both thermal and renewable (including hydro) units, based on their annual capacity and availability (dependent on each zone, month, and hour) factors correspondingly.

$$\sum_{m \in M} \sum_{t \in T} p_{i,y,m,t} \cdot DUR_m \le c_{i,y} \cdot \left(1 - EFOR_{i,y}\right) \cdot 8760 \quad \forall i \in I^{TH}, y \tag{15.57}$$

$$\sum_{m \in M} \sum_{t \in T} p_{i,y,m,t} \cdot DUR_m \le \sum_{m \in M} \sum_{t \in T} c_{i,y} \cdot AV_{i,z,m,t} \cdot 8760 \quad \forall i \in \left(I^{RES} \bigcap I^z\right), z, y \tag{15.58}$$

15.3.2.19 Renewable Energy Technologies' Production

Constraints (15.59) state that a specific amount $\left(REN_y\right)$ of the total electricity produced $\left(\sum_{i\in I}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m\right)$ in each year $y\in Y$ should come from renewable energy technologies $\left(\sum_{i\in I^{RES}}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m\right)$.

$$\sum_{i\in I^{RES}}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m \geq REN_y\cdot \sum_{i\in I}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m \quad \forall y\in Y \tag{15.59}$$

15.3.2.20 CO_2 Emissions Cap

According to constraints (15.60), a maximum allowable bound $\left(CO_2\_CAP_y\right)$ is imposed on the amount of CO_2 emissions produced by all power generating units.

$$\sum_{i\in I}\sum_{bl\in BL}\sum_{m\in M}\sum_{t\in T} pb_{i,bl,y,m,t}\cdot CO_2\_EF_{i,bl}\cdot DUR_m \leq CO_2\_CAP_y \quad \forall y\in Y \tag{15.60}$$

15.3.2.21 Maximum Penetration of Renewable Energy Technologies

Constraints (15.61) state that the electricity produced from renewable energy technologies $\left(\sum_{i\in I^{RES}}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m\right)$ in each year $y\in Y$ must not exceed a specific proportion $\left(MRES_y\right)$ of the total electricity produced $\left(\sum_{i\in I}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m\right)$ in the same period.

$$\sum_{i\in I^{RES}}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m \leq MRES_y\cdot \sum_{i\in I}\sum_{m\in M}\sum_{t\in T} p_{i,y,m,t}\cdot DUR_m \quad \forall y \tag{15.61}$$

15.3.2.22 Maximum Capacity of Renewable Energy Technologies

Finally, constraints (15.62) impose an upper limit on the amount of the available capacity $\left(RES\_CAP_{i,y}\right)$ of each renewable energy technology (not including hydroelectric units) that can be installed in each zone $z\in Z$ and year $y\in Y$.

$$c_{i,y} \leq RES\_CAP_{i,z,y} \quad \forall i\in \left(I^{RES-}\bigcap I^z\right), z, y \tag{15.62}$$

15.4 Case Study

The Greek power system is used as a representative case study. The examined planning horizon is that between 2014 and 2030. Twelve months (January to December) have been considered, and a representative day (24 h) per month, each having the duration of the corresponding month in days. The Greek power system has been divided into two sectors (North and South system), as well as into five zones (Zones 1 and 2 are involved in the North system, as well as Zones 3–5 belong to the South system). Maximum corridor flow between North and South sectors is set to 3100 MW throughout the planning horizon. Five neighbouring countries (Albania, Bulgaria, FYROM, Turkey, and Italy) are interconnected with the Greek power system. More specifically, Albania, Bulgaria and FYROM are interconnected with Zone 2 (North system), Turkey with Zone 1 (North system), and Italy with Zone 4 (South system).

The existing capacity per technology type and zone is taken into account and presented in Table 15.2 (Koltsaklis et al. 2015; IPTO 2013, 2014; LAGIE, RES & CHP Monthly Statistics 2012; LAGIE, RES & CHP Monthly Statistics 2013; LAGIE, DAS Monthly Reports 2012; LAGIE, DAS Monthly Reports 2013).

Lignite, natural gas, coal and oil units comprise the thermal units, while RET units consist of large hydroelectric, wind, photovoltaics and other RET units (e.g., biomass, small hydro, and CHP units). The decommissioning plan of the existing power fleet is also provided in Table 15.3. As it can be observed, there is a gradual decommissioning of old lignite power stations, along with some natural gas and oil-fired power plants.

New candidate units are also available for installation, the construction time and technical lifetime of which are provided in Table 15.4 (Koltsaklis et al. 2015).

Detailed data of their unit investment costs can be found in Xiao et al. (2011), while the upper bound for the installed wind power capacity in the interconnected system is set to 5500 MW. Maximum yearly penetration rate is assumed to be 1000 MW for both wind turbines and photovoltaics.

Table 15.2 Existing installed capacity of the Greek interconnected power system at the end of 2013 (MW)

Sector	North system		South system		
Zone	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5
Lignite	0.0	3985.0	0.0	0.0	530.0
Natural gas combined cycle	480.0	390.0	3902.2	0.0	0.0
Natural gas open cycle	0.0	0.0	147.0	0.0	0.0
Hydroelectric	0.0	1747.0	120.0	1433.6	64.0
Oil	0.0	0.0	718.0	0.0	0.0
Wind	260.2	93.9	621.9	214.1	347.6
Photovoltaics	243.3	488.7	869.8	479.6	340.3
Other RES	56.6	115.5	133.0	92.1	3.9
Total	1040.1	6820.0	6511.9	2219.4	1285.7

Table 15.3 Decommissioning plan of the Greek interconnected power system (MW)

Decommissioned plants	2014	2015	2016	2018	2019	2022	2028	**2029**	**2030**
Oil	287	0	0	144	144	143	0	0	0
Natural gas	513	0	0	0	0	0	0	0	0
Lignite	154	116	274	0	1150	270	260	548	283
Total	954	116	274	144	1294	413	260	548	283

Table 15.4 Construction time and lifetime of new candidate units (y)

Unit type	Construction time	Lifetime
New lignite	4	35
New coal	4	35
New NGCC	3	35
New NGGT	2	30
New oil	2	30
New hydro	4	40
New wind	1	20
New photovoltaics	1	20

With regard to the production potential of RET units, they are characterized by a specific availability factor per zone and time period (at a monthly and hourly level). Typical values of the wind and solar availability in the Greek power system are depicted in Figs. 15.1 and 15.2 respectively (Koltsaklis et al. 2015; IPTO 2014; LAGIE, RES & CHP Monthly Statistics 2012; LAGIE, RES & CHP Monthly Statistics 2013). Concerning the yearly availability of the thermal units, typical values can be found in Koltsaklis et al. (2014) and (2015), while the constantly available electricity net imports capacity is assumed to be 600 MW.

Detailed power technical characteristics of representative hydrothermal units are presented in Tables 15.5, 15.6 and 15.7, including ramp-rates (in dispatchable phase and when providing secondary reserve), maximum contribution in all reserve types, power output when operating in soak phase, as well as minimum and maximum power output when committed and when providing secondary reserve. When it comes to reserve provision costs, the price of each unit's primary energy offer is assumed to be 0.001 €/MW, while the price of each unit's secondary range offer is considered to be 9 €/MW throughout the planning horizon.

Note that lignite units cannot provide secondary reserve, while hydroelectric units cannot contribute to primary reserve requirements. Furthermore, lignite (base-load) and natural gas combine cycle units (medium-load) cannot provide tertiary non-spinning reserve requirements. Table 15.8 shows representative CO_2 emission factors per power block and thermal power technology type (Koltsaklis et al. 2015). Note that the CO_2 emission factors adopted for the renewable energy technologies do not include life cycle analysis, since this approach is out the scope of our work.

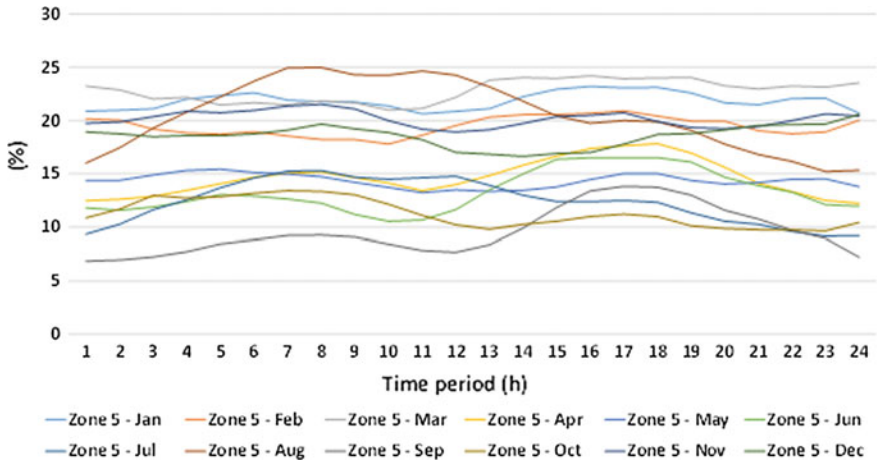

Fig. 15.1 Hourly wind availability factor (%) in each representative day of each month in Zone 5

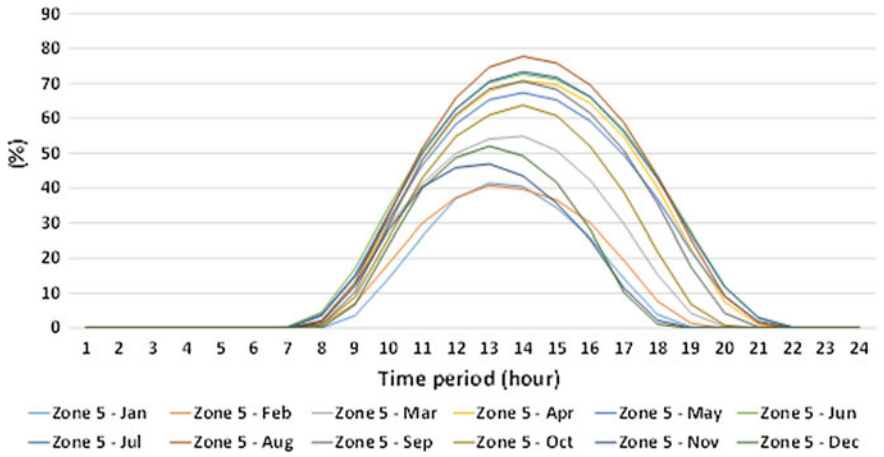

Fig. 15.2 Hourly solar availability factor (%) in each representative day of each month in Zone 5

As described in Sect. 15.2, there are three available start-up types including hot, warm and cold start-up. Table 15.9 presents each representative unit's maximum non-operational time before changing standby condition, i.e., from hot to warm, and from warm to cold standby condition, as well as corresponding shut-down cost per technology type.

According to the selected start-up type, there are different synchronization and soak time periods per technology type. Representative minimum up and down as well as desynchronization time periods (along with the relevant shut-down cost) per technology type can be found in Koltsaklis et al. (2015).

Table 15.5 Representative ramp rates and maximum contribution in primary reserve per technology type (MW)

Unit type	Ramp-down rate (MW/min)	Ramp rate when providing secondary reserve (MW/min)	Ramp-up rate (MW/min)	Maximum contribution in primary reserve (MW)
Lignite-1	4	0	3	24
Oil-1	2	0	2	12
NGCC-1	9	9	9	25
NGCC-2	12	12	12	34
NGCC-3	11	11	11	28
NGCC-4	10	10	10	30
NGGT-1	8	0	3	0
Hydro-1	80	60	80	0
Hydro-2	40	40	40	0

Table 15.6 Representative maximum contribution in secondary, spinning and non-spinning tertiary reserve per technology type (MW)

Unit type	Maximum contribution in secondary reserve (MW)	Maximum contribution in non-spinning tertiary reserve (MW)	Maximum contribution in spinning tertiary reserve (MW)
Lignite-1	0	0	45
Oil-1	0	0	30
NGCC-1	370	0	135
NGCC-2	422	422	180
NGCC-3	415	0	165
NGCC-4	476	0	150
NGGT-1	0	49	45
Hydro-1	240	320	320
Hydro-2	230	230	230

The determination of which power unit is to operate in each time period is based on both technical and economic criteria. Representative marginal cost of the energy offer function (4 power blocks) per technology type, as well as typical capacity blocks per technology type can be found in Koltsaklis et al. (2015). Similar energy offers are submitted to the electricity market operator by both electricity importers and exporters. Note that the marginal cost of RET units is assumed to be zero (non-priced power generation) as well as there is the fixed, non-priced power contribution from hydroelectric units (mandatory hydro injection) and from units operating under commissioning status. As a consequence, these units are given priority when entering the power system. The fixed O&M cost of hydrothermal units is incorporated in their marginal costs, while that of RET (not including hydroelectric units) is given in Koltsaklis et al. (2015).

Table 15.7 Representative power outputs in different operational stages per technology type (MW)

Unit type	Minimum power output when providing secondary reserve (MW)	Maximum power output when providing secondary reserve (MW)	Fixed power output when operating in soak phase (MW)	Maximum power output in dispatchable phase (MW)	Minimum power output in dispatchable phase (MW)
Lignite-1	0	0	80	274	150
Oil-1	0	0	30	144	50
NGCC-1	220	370	80	390	220
NGCC-2	220	422	42	422	220
NGCC-3	130	476	68	480	130
NGCC-4	220	415	44	433.5	220
NGGT-1	0	0	5	49	10
Hydro-1	25	360	0	360	0
Hydro-2	50	300	0	300	0

Table 15.8 Representative CO_2 emission factor per capacity block and technology type (t CO_2/MWh) (Koltsaklis et al. 2015)

Unit type	Block 1	Block 2	Block 3	Block 4
Existing lignite	1.395	1.352	1.334	1.324
Existing NGCC	0.362	0.337	0.322	0.308
Existing NGGT	1.787	0.792	0.652	0.599
Existing oil	0.806	0.746	0.733	0.727
New lignite	1.159	1.035	0.986	0.963
New Coal	0.971	0.848	0.799	0.776
New NGCC	0.425	0.402	0.396	0.378
New NGGT	1.260	0.702	0.625	0.580

Table 15.9 Representative non-operational time intervals before each representative unit's transition to the next standby condition (h) and shut-down cost (€)

Unit type	Non-operational time before going from hot to warm standby condition (h)	Non-operational time before going from warm to cold standby condition (h)	Shut-down cost (€)
Lignite-1	24	144	110000
Oil-1	12	84	45000
NGCC-1	12	72	49900
NGCC-2	16	76	48532
NGCC-3	15	111	45000
NGCC-4	9	57	49000
NGGT-1	1	2	333.3

Table 15.10 Power injection losses coefficients per load block and zone (p.u.)

Load level (MW)	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5
3500	1	0.982	1	1	1
3750	1	0.981	1	1	1
4000	1	0.98	1	1	0.999
4250	0.981	0.979	1	1	0.991
4500	0.98	0.976	1	1	0.992
4750	0.981	0.975	1	1	0.995
5000	0.982	0.974	1	1	0.998
5250	0.984	0.974	0.999	1	0.999
5500	0.95	0.975	0.999	1	1
5750	0.949	0.974	0.999	1	0.996
6000	0.933	0.975	0.998	1	0.999
6250	0.931	0.974	0.997	1	1
6500	0.934	0.974	0.998	1	1
6750	0.933	0.973	0.997	1	1
7000	0.933	0.972	0.996	1	1
7250	0.934	0.973	0.998	1	0.979
7500	0.934	0.972	0.997	1	0.981
7750	0.933	0.971	0.995	1	0.984
8000	0.932	0.969	0.995	1	0.987
8250	0.931	0.968	0.993	1	0.988
8500	0.93	0.966	0.992	1	0.989
8750	0.929	0.965	0.99	1	0.994
9000	0.927	0.963	0.99	0.997	0.995
9250	0.927	0.961	0.991	0.994	0.997
9500	0.926	0.96	0.991	0.995	0.999
9750	0.926	0.958	0.99	0.99	1
10000	0.928	0.957	0.99	0.983	1
10250	0.927	0.956	0.99	0.982	1
10500	0.927	0.954	0.991	0.981	1
10750	0.927	0.953	0.991	0.979	1
11000	0.926	0.952	0.992	0.976	1
11250	0.927	0.951	0.992	0.973	1
11500	0.928	0.951	0.992	0.971	1

In order to calculate net power injections in the electric system (at market point), an injection losses coefficient has been considered based on the total load level and zone, detailed values of which are presented in Table 15.10 (Koltsaklis et al. 2015).

Power system's requirements include power load to be covered by domestic power generation and net imports, as well as all types of reserve requirements (primary-up, secondary-up and -down, fast secondary-up and -down, tertiary spinning and non-spinning). Figure 15.3 depicts the hourly power demand profile

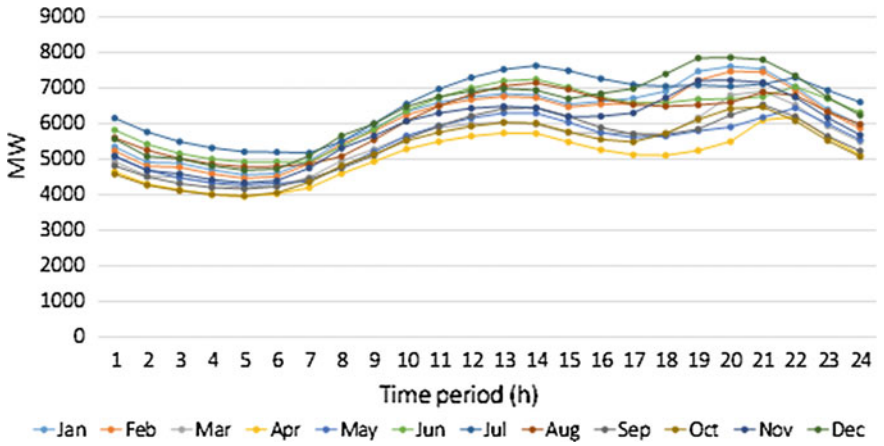

Fig. 15.3 Hourly power demand profile of the interconnected system in each representative day of each month during the first year of the planning horizon (2014)

of the interconnected system in each representative day of each month during the first year of the planning horizon (2014). It has been assumed an annual increase of 1 % in the power demand during the following years. Figures 15.4 and 15.5 portray system's secondary-up and down reserve requirements in selected time intervals (2014, 2020, 2025, and 2030). It is also assumed that they follow the same patterns in each representative day of each month during the same year.

System's stability is guaranteed by incorporating minimum peak reserve requirements (15 % above the projected power demand of each year, representative month-day and hour throughout the studied period), and a maximum yearly RES

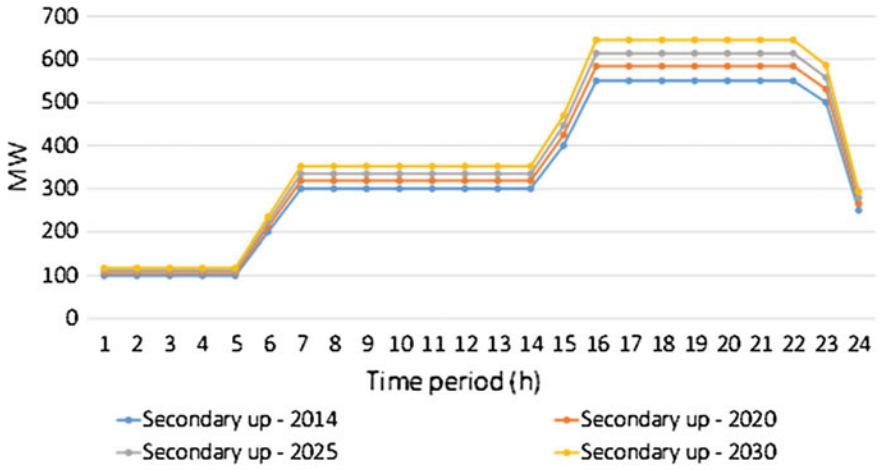

Fig. 15.4 Secondary-up reserve requirements in selected time intervals (MW)

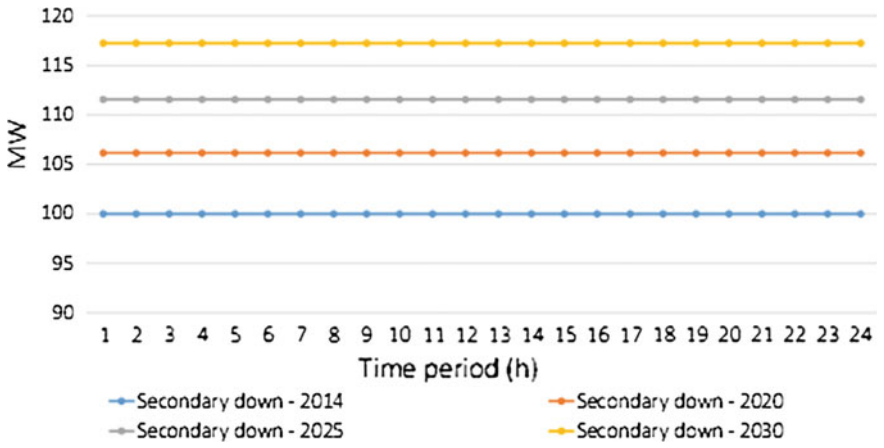

Fig. 15.5 Secondary-down reserve requirements in selected time intervals (MW)

penetration in the power generation mix in our model (60 % of the total electricity generated in each year). An energy policy tool to support RET penetration has been incorporated in our model through the introduction of a RET penetration target in the power mix, according to which 40 % of the electricity demand is to be satisfied from RES technologies from 2020 onwards. Finally, in order to control the amount of CO_2 emissions in the atmosphere produced by the power sector, a "cap and trade" policy tool is considered in our model. The evolution of CO_2 emission pricing follows three possible pathways, starting from 6 €/t in 2014 in all cases and reaching 15, 30, and 50 €/t CO_2, according to the selected CO_2 emission pricing scenario (*low*, *medium*, and *high* CO_2 emission price scenarios). The case of no CO_2 emission pricing is also examined. With regard to CO_2 emissions cap, a specific target has been considered including a 29 % decrease in the amount of CO_2 emissions produced between 2014 and 2030. The impacts of no RET penetration and CO_2 policy on the capacity additions and the power generation mix have been also considered. All these are depicted in Fig. 15.6.

Five distinct scenarios of the Greek interconnected power system have been examined which are diversified based on different policy measures. Table 15.11 presents all the relevant information.

15.5 Results and Discussion

This section provides the results and an analytical discussion of all the model scenarios implemented and solved in our study. The problem has been solved to global optimality making use of the ILOG CPLEX 12.6.0.0 solver incorporated in the General Algebraic Modelling System (GAMS) tool (GAMS 2015). An integrality gap of 1 % has been achieved in all cases (Scenarios 1–5).

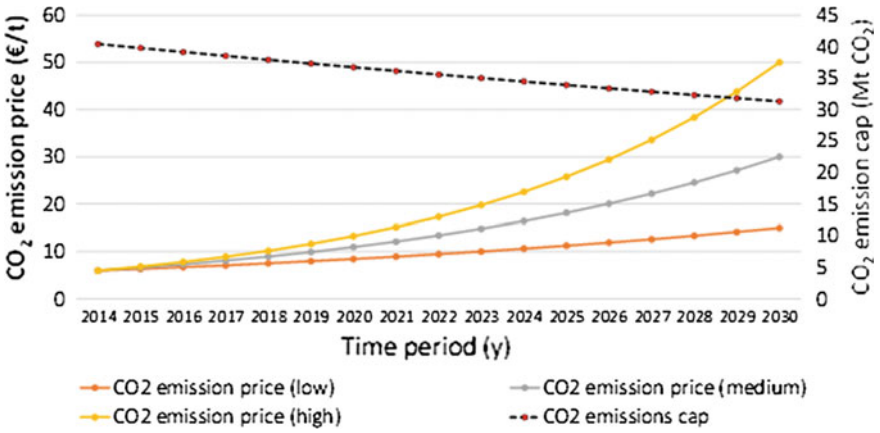

Fig. 15.6 CO_2 emission price (€/t) and cap (Mt CO_2) scenarios

Table 15.11 Assumptions of each examined scenario

Scenarios	CO_2 emission pricing	CO_2 emission cap	Renewables penetration target
Scenario 1	No	Yes	Yes
Scenario 2	Low	Yes	Yes
Scenario 3	Medium	Yes	Yes
Scenario 4	High	Yes	Yes
Scenario 5	No	No	No

15.5.1 Capacity Additions, Production Mix and System's Marginal Price

15.5.1.1 Scenario 1

Scenario 1 includes two of the three available energy policy tools for the promotion and further penetration of RET in the power system: (i) 40 % of the total electricity demand should come from renewable energy technologies (including hydroelectric units), and (ii) a CO_2 emissions cap. However, it is assumed that there is no CO_2 emission pricing, i.e., CO_2 emission price is assumed to be zero throughout the planning horizon.

Figure 15.7 presents the resulting power generation mix during each time period. The results indicate a gradually diminishing share of the lignite power production, from 26.6 TWh in 2014 to 17.3 TWh in 2020, and 10.8 TWh in 2030. A new lignite unit with a capacity of 300 MW is also added in the system. This evolution can be explained by: (i) the mandatory RES power generation, (ii) the existence of a CO_2 emission cap, and (iii) the gradual decommissioning of some old lignite power stations. However, that decrease in the lignite power generation is

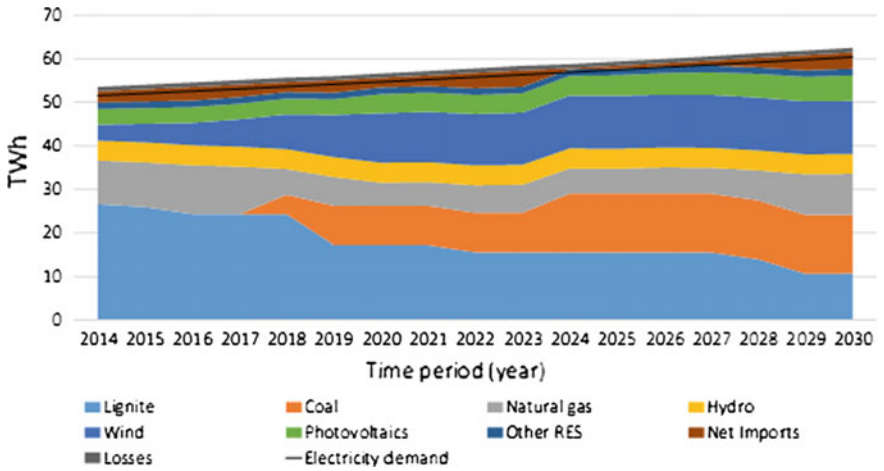

Fig. 15.7 Power generation mix evolution in Scenario 1 (TWh)

Table 15.12 Lowest peak reserve requirements observed in each scenario (%) and system's total cost per scenario (bn €)

Scenarios	Lowest peak reserve (%)	Year	Representative date	System's total cost (2014–2030)
Scenario 1	21.3	2030	Feb, 22th hour	52.3
Scenario 2	15	2026	Feb, 22th hour	57.3
Scenario 3	15.1	2030	Feb, 22th hour	59.2
Scenario 4	15.1	2027	Feb, 22th hour	61.4
Scenario 5	18.2	2030	Jan, 21th hour	50.3

offset to a great extent by the construction of three new coal units having a total capacity of 1800 MW, as depicted in Table 15.12. These units contribute 13.4 TWh to the electricity demand satisfaction in 2030. As a consequence, the sum of lignite and coal power generation in 2030 equals 91 % of the lignite electricity production in 2014. Natural gas units' power generation is characterized by significant fall on the grounds that it starts from 10 TWh in 2014, falls to 5.3 TWh in 2020, remains stable with 5.75 TWh in 2025, and rises again in 2029 when it reaches 9.4 TWh (the same in 2030) due to the decommissioning of some old lignite units.

With regard to RET units, especially to wind turbines and photovoltaics, significant investments are required since an additional amount of 3.9 GW of wind turbines is added to the system (1.4 GW for photovoltaics). This is clearly reflected on the power generation mix because wind turbines' electricity production has skyrocketed, from 3.7 TWh in 2014 to 12.2 TWh in 2030, while photovoltaics increase their market share at a less intense rate, from 3.6 TWh in 2014 to 5.9 TWh in 2030. Water availability per hydroelectric unit is assumed to be the same in each year of the planning horizon and thus, hydroelectric contribution amounts to 4.6 TWh in each year of the studied period, because no new hydropower units are to be constructed.

Electricity trade is highly fluctuating during the whole studied period. More specifically, net electricity imports (difference between electricity imports and exports) begin from 2.6 TWh in 2014, decrease to 2.25 TWh in 2020, drop to 0.27 TWh in 2024, and gradually increase their share towards 2030 (3.76 TWh in 2030) due to the decommissioning of some old lignite power stations. The results highlight that the shrinking share of the existing lignite power capacity during the last 3 years of the study is offset by increased contribution from natural gas units and electricity imports.

With regard to the evolution of the system marginal price (SMP), it is characterized by a slight increase of 8.7 %, from 55.49 €/MWh in 2014 to 60.3 €/MWh in 2030 (weighted annual average price), as presented in Fig. 15.12. This slight increase can be attributed to the penetration of new low-cost thermal units (lignite and mainly coal), as well as to the massive penetration of RES units having a zero marginal cost and thus, the portion of demand to be met by hydrothermal units is quite small and the new low-cost units comprise the marginal units in most hours. No differences are reported between north and south marginal prices on the grounds that their corridor flow constraint is neither activated nor violated.

15.5.1.2 Scenario 2

Scenario 2 includes the same targets with those of Scenario 1 incorporating also the energy policy tool of CO_2 emission pricing, i.e., CO_2 emission allowances price begins from 6 €/tCO_2 and reaches 15 €/t CO_2 in 2030, as can be seen in Table 15.11. Figure 15.8 presents the power generation mix of Scenario 2 throughout the planning horizon. As in Scenario 1, lignite power generation reports a large decrease in its contribution, from 26.6 TWh in 2014 to 8.9 TWh in 2030. No

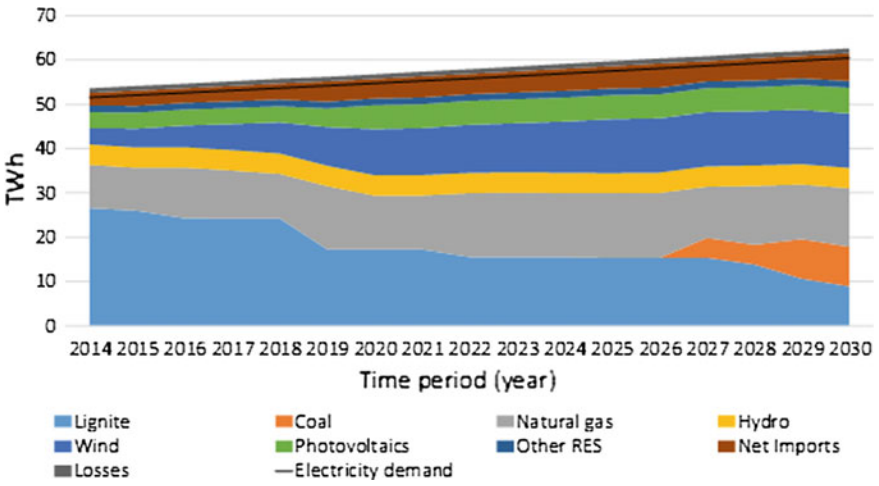

Fig. 15.8 Power generation mix evolution in Scenario 2 (TWh)

new lignite units are added in the system. Due to the introduction of CO_2 emission pricing, only two new coal units (instead of three new coal units of the previous case, i.e., Scenario 1) are to be constructed with a total capacity of 1200 MW. As a result, the sum of coal and lignite electricity generation in 2030 equals 67 % of the lignite electricity generation in 2014 and underlining the impact that CO_2 emission pricing has on the lignite and coal power generation.

The declining share of lignite and coal units is primarily met by natural gas units whose contribution rises by 35.1 %, from 9.8 TWh in 2014 to 13.2 TWh in 2030, redefining the role of natural gas in the country's power generation mix. Note that two new natural gas open cycle units are to be constructed in 2030 with a total capacity of 300 MW. These units have relatively low investment cost but substantially high variable operating cost, approaching the level of 150 €/MWh. These units are typically constructed for covering peak demand periods during a year and guaranteeing peak reserve requirements. As a result, they typically operate a limited amount of time at an annual level. Net electricity imports play also a significant role in covering an increasing part of the power demand, reporting an 87.6 % increase in their contribution, from 2.86 TWh in 2014 to 6.3 TWh in 2030. Wind turbines and photovoltaics follow the same pattern with that of Scenario 1 in order for the RES target to be achieved. When it comes to the evolution of SMP, it is characterized by a higher growth rate when compared to the corresponding one of Scenario 1 (Fig. 15.12). More specifically, it rises by almost 19 % between 2014 and 2030, from 57.7 €/MWh in 2014 to 68.6 €/MWh in 2030. The higher levels of SMP are attributed to the increasing share of natural gas units, having a higher variable cost (70–75 €/MWh) than that of coal and lignite units (45–60 €/MWh).

15.5.1.3 Scenario 3

Scenario 3 includes the same targets with those of Scenario 2. However, a more intense rate for the evolution of CO_2 emission pricing is assumed, i.e., it soars from 6 €/t CO_2 in 2014 to 30 €/t CO_2 in 2030 (see Table 15.11). The impact of the rise in the CO_2 emission price is clearly reflected on the total power generation from coal and lignite units. Only one lignite and one coal unit are to be constructed between 2014 and 2030, having a total installed capacity of 900 MW. Total lignite and coal power generation in 2030 equal 54.3 % of that of lignite electricity production in 2014, i.e., a decrease of 45.7 % (Fig. 15.9).

This fall in lignite- and coal-fired electricity generation is mainly counterbalanced by natural gas units and electricity imports. Natural gas units' production is determined to rise by 64.8 %, from 9.7 TWh in 2014 to 16.01 TWh in 2030. Four new natural gas open cycle units are to be constructed with a total capacity of 600 MW, underscoring the importance of these units in guaranteeing peak reserve requirements when CO_2 emission price constitutes an active energy policy tool. Net electricity imports are characterized by a growth of 140.7 % on the grounds that they begin from 2.91 TWh in 2014 and result in 7 TWh in 2030. Wind turbines,

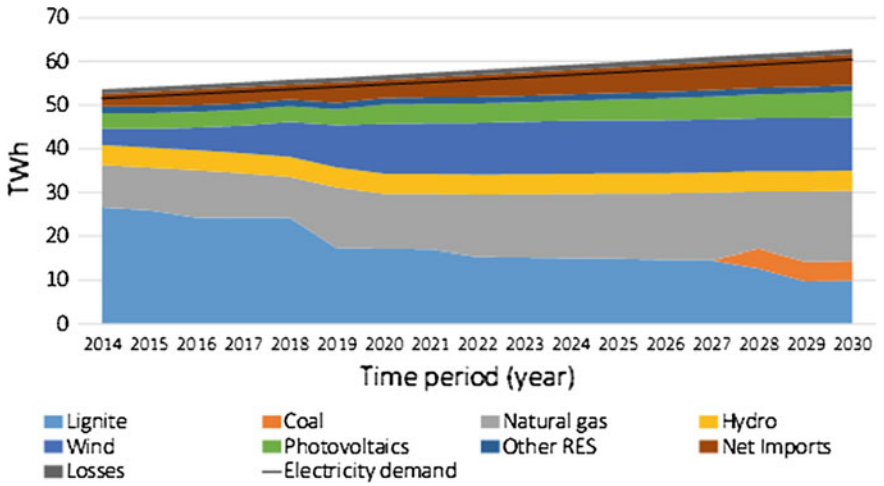

Fig. 15.9 Power generation mix evolution in Scenario 3 (TWh)

photovoltaics and hydroelectric units maintain the same share with the corresponding one of the previous scenarios, due to the constraint for the mandatory penetration of a specific share of RET.

The combined effects of both increased CO_2 emission price and natural gas-fired electricity generation is clearly reflected on the evolution of SMP. More specifically, a rise of 26.8 % of its value is reported, from 57.6 €/MWh in 2014 to 73 €/MWh in 2030 (Fig. 15.12).

15.5.1.4 Scenario 4

Apart from the RET penetration target and the CO_2 emission cap, an even more intense growth rate of the CO_2 emission price is adopted in Scenario 4, i.e., it escalates from 6 €/t CO_2 in 2014 to 50 €/t CO_2 in 2030. Neither coal nor lignite units are to be constructed, while lignite power generation plummets, since it starts from 26.6 TWh in 2014 to result in 7.4 TWh in 2030 (Fig. 15.10).

This is the most favourable scenario for natural gas units in terms of both installed capacity and electricity generation. Six new natural gas units are to be constructed throughout the planning horizon, of which three are natural gas combined cycle units with a total capacity of 1.2 GW and three natural gas open cycle units having a total capacity of 450 MW. This massive penetration of natural gas units is better reflected on their total electricity generation, climbing from 9.7 TWh in 2014 to 15 TWh in 2025, and reaching 22.1 TWh in 2030, i.e., an increase of 127.5 %. As in the previous scenarios, net electricity imports account for a significantly increasing share in the power demand satisfaction, starting from 2.9 TWh in 2014, rising to 5 TWh in 2021, and reaching 8 TWh in 2030, highlighting the

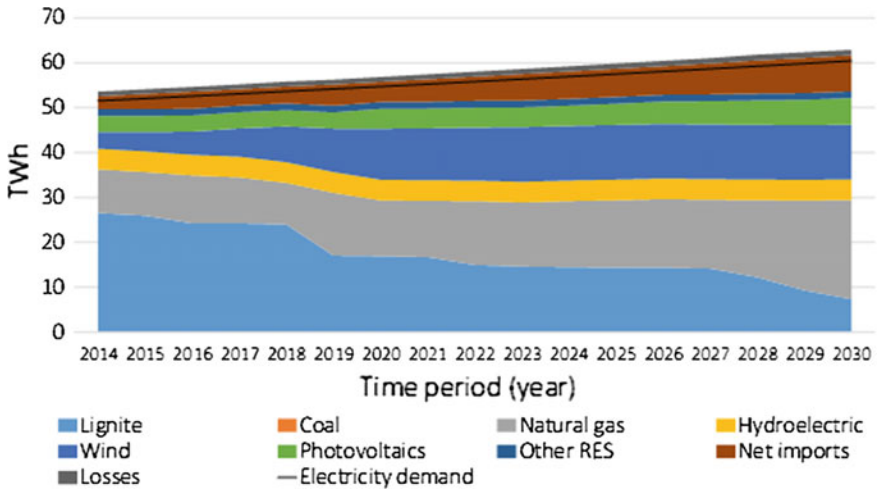

Fig. 15.10 Power generation mix evolution in Scenario 4 (TWh)

flexibility provided in the context of the European power networks integration. The existence of the RET penetration target leads to the same RET growth rates as those presented in the previous three Scenarios (Scenarios 1–3).

Not surprisingly, it can be observed a leap in the market clearing price. An increase of 39.3 % is reported, since it jumps from 57.8 €/MWh in 2014 to 80.5 €/MWh in 2030 (Fig. 15.12). The main reason for that rise can be attributed again to the combined effects of high CO_2 emission price and increased natural gas production.

15.5.1.5 Scenario 5

Scenario 5 differentiates from all previous scenarios (Scenarios 1–4) in terms of the applied energy policy tools. RET penetration target, CO_2 emission cap and pricing are not taken into consideration in that scenario, as presented in Table 15.11. This is the most favourable scenario for coal and lignite units in terms of both lignite and coal capacity. Three new lignite (1.5 GW) and three new coal units (1.8 GW) are to be constructed with a total capacity of 3.3 GW. Lignite electricity generation reports a slight drop of 31.6 %, from 26.6 TWh in 2014 to 18.2 TWh in 2030 (Fig. 15.11).

Total lignite- and coal-fired power generation in 2030 exceeds that of lignite-fired electricity production in 2014 by almost 19 % (5 TWh). A medium rise in the natural gas-fired power generation is also observed, from 10.5 TWh in 2014 to 12.9 TWh in 2030. The most noticeable aspect of this policy scenario is that no new investments on RET units are determined by the model (except for an almost negligible capacity addition of wind turbines in 2030). This is also reflected in their share in the power demand satisfaction. Apart from RET units, net electricity

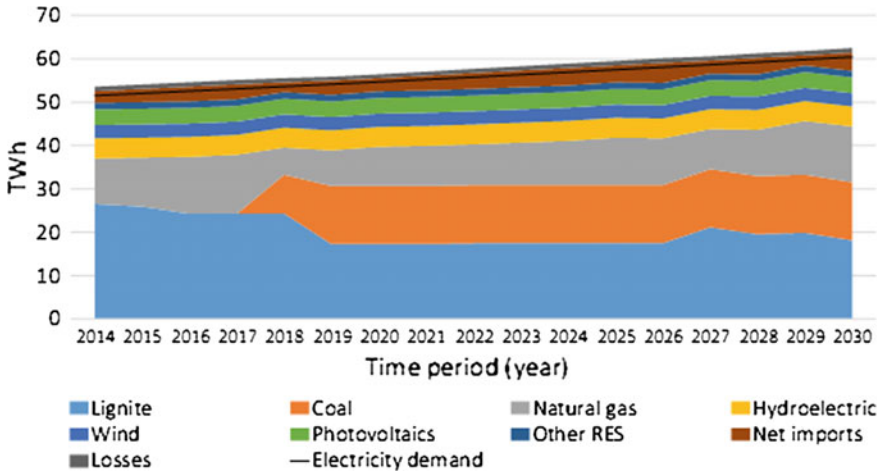

Fig. 15.11 Power generation mix evolution in Scenario 5 (TWh)

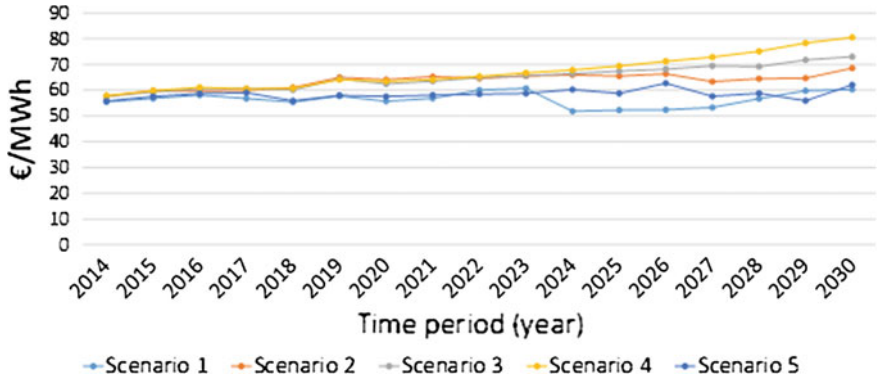

Fig. 15.12 System marginal price evolution (weighted annual average values) in all scenarios (€/MWh)

imports also lose their competitiveness when compared to that of the new lignite and coal units, since they report a small increase from 2.7 TWh in 2014 to 4.2 TWh in 2030.

With regard to SMP evolution, an increase of 11.3 % is reported from 55.8 €/MWh in 2014 to 62.1 €/MWh in 2030 (Fig. 15.12). The SMP of that scenario in 2030 is higher than the corresponding one of Scenario 1, due to the fact that a lower share of electricity demand is met by RET (zero marginal cost), the absence of CO_2 emission pricing, and the relatively higher contribution from natural gas units in Scenario 5 than that of Scenario 1.

15.5.2 CO_2 Emissions

Figure 15.13 illustrates the evolution of CO_2 emissions in each Scenario during the whole planning horizon, as well as the policy target of the desired amount of CO_2 emissions. It can be observed that when the energy policy tool of CO_2 emissions cap is activated (Scenarios 1–4), the amount of determined CO_2 emissions are significantly lower than the maximum allowable level and they are generally characterized by a decreasing trend. Not surprisingly, the most noticeable reduction is reported in Scenario 4, from almost 38 Mt CO_2 in 2014 to 15.7 Mt CO_2 in 2030 (a decrease of 58.6 %), due to the combined effects of both high CO_2 emission price (from 6 €/t CO_2 in 2014 to 50 €/t CO_2 in 2030) and the existence of the cap. With regard to the other three Scenarios where the CO_2 emissions cap is taken into consideration (Scenarios 1, 2 and 3), a fall of 30.6 % (between 2014 and 2030) is observed in Scenario 1, a decrease of 41.7 % in Scenario 2 and a drop of 47.5 % in Scenario 3. The impact of the absence of CO_2 emission pricing is clearly highlighted in Scenario 1, being the scenario with the lowest CO_2 emissions mitigation of those that CO_2 emissions cap has been applied. In the absence of environmental policy, i.e., CO_2 emissions pricing and cap as well as RET penetration target are not considered, the decrease in the amount of CO_2 emissions equals 22.6 % (between 2014 and 2030). Comparing the evolution of CO_2 emissions in Scenario 5 with the maximum allowable ones of the previous 4 scenarios (see Fig. 15.13), it can be easily concluded that market dynamics are unable to meet the specific environmental targets without the introduction of specific policy measures and tools.

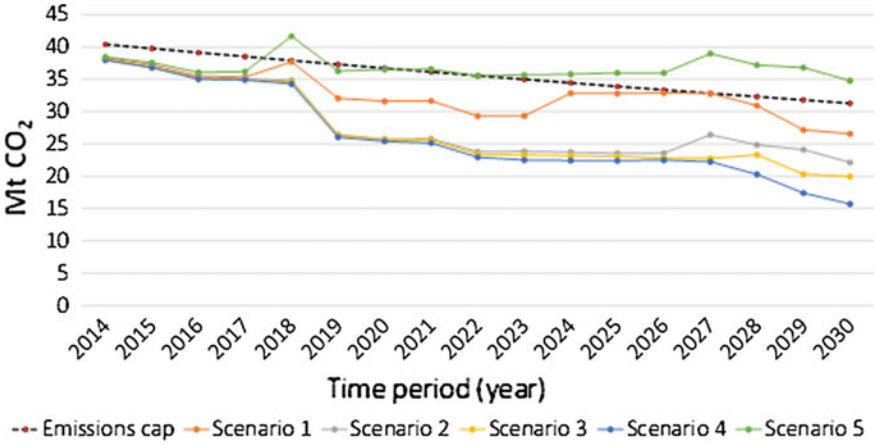

Fig. 15.13 CO_2 emissions evolution during each time period in each scenario (Mt CO_2)

15.5.3 Peak Reserve Requirements and System's Total Cost

Table 15.12 presents the lowest peak reserve requirements observed in each scenario as well as system's total operational and development cost in each scenario. As can be seen, the lowest values in all scenarios are reported during the last hours of February (January in Scenario 5), which can be mainly attributed to the low availability of photovoltaic units during that period. All the reported values are close to the applied peak reserve, i.e., 15 %. When it comes to system's total cost, the lowest cost is achieved in Scenario 5, where there is no environmental policy and target. Furthermore, the progressive increase in the CO_2 emission price lead to a gradual rise in the system's total cost from Scenario 1 to Scenario 4.

15.6 Concluding Remarks

This work presents an integrated generation expansion planning model incorporating key short-term constraints of the daily energy planning (unit commitment problem). Optimal long-term investment decisions are determined by considering the dynamic aspects of the electricity market. The applicability of the proposed framework has been illustrated on the Greek power system. The findings of the study include the optimal power production mix, capacity additions, and evaluation of SMP evolution. Five scenarios have been studied to quantify the impacts of several energy policy tools on the development and planning of the power system. Lignite and coal units are favoured in the absence of environmental policy or in a low CO_2 emission price scenario. Natural gas open cycle units provide flexibility to the system, guarantee peak reserve requirements, and they enter the system (new units) when there is a CO_2 emission cap and/or CO_2 emissions pricing is put into practice. Natural gas combined cycle units are added to the system (new units) when there is a high CO_2 emissions price scenario (50 €/t CO_2). With regard to RET units, there are almost no relevant investments when no environmental policy is applied. A significant penetration of wind turbines is observed in all other cases. Finally, SMP is proved to be highly sensitive to the type of units utilized for power generation and remains at low levels in the absence of environmental policy. The findings of the study highlight a positive correlation of significant RES penetration with high natural gas production and electricity trade, offering more flexibility to the power demand satisfaction.

Research challenges that need to be considered in future works include a detailed assessment of a power system's projected requirements in each reserve type (mainly secondary and tertiary reserve) in parallel with the expected massive penetration of renewables so as to definitely secure the well-organized operation of the network.

Acknowledgments Financial support from the European Commission's Marie Curie IRSES project (Contract No: PIRSES-GA-2011-294987) "Energy Systems Engineering" (ESE) is gratefully acknowledged.

References

Ahmed, S., Elsholkami, M., Elkamel, A., Du, J., Ydstie, E. B., & Douglas, P. L. (2014). Financial risk management for new technology integration in energy planning under uncertainty. *Applied Energy, 128*, 75–81.

Ahmed, S., Elsholkami, M., Elkamel, A., Du, J., Ydstie, E. B., & Douglas, P. L. (2015). New technology integration approach for energy planning with carbon emission considerations. *Energy Conversion and Management, 95*, 170–180.

Andrianesis, P., Biskas, P., & Liberopoulos, G. (2011). An overview of Greece's wholesale electricity market with emphasis on ancillary services. *Electric Power Systems Research, 81*, 1631–1642.

Barteczko-Hibbert, C., Bonis, I., Binns, M., Theodoropoulos, C., & Azapagic, A. (2014). A multi-period mixed-integer linear optimisation of future electricity supply considering life cycle costs and environmental impacts. *Applied Energy, 133*, 317–334.

Batas Bjelić, I., Rajaković, N., Ćosić, B., & Duić, N. (2013). Increasing wind power penetration into the existing Serbian energy system. *Energy, 57*, 30–37.

Belderbos, A., & Delarue, E. (2015). Accounting for flexibility in power system planning with renewables. *International Journal of Electrical Power and Energy Systems, 71*, 33–41.

Connolly, D., Lund, H., Mathiesen, B. V., & Leahy, M. (2010). A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied Energy, 87*(4), 1059–1082.

Connolly, D., Lund, H., Mathiesen, B. V., & Leahy, M. (2011). The first step towards a 100 % renewable energy-system for Ireland. *Applied Energy, 88*(2), 502–507.

Delarue, E., Bekaert, D., Belmans, R., & D'Haeseleer, W. (2009). Development of a comprehensive electricity generation simulation model using a mixed-integer programming approach. *TME working paper–energy and environment. KULeuven Energy Institute*. Retrieved March 23, 2015 from http://www.mech.kuleuven.be/en/tme/research/energy_environment/Pdf/WPEN2007-09.

Delarue, E. D., Luickx, P. J., & D'haeseleer, W. D. (2009b). The actual effect of wind power on overall electricity generation costs and CO_2 emissions. *Energy Conversion and Management, 50*(6), 1450–1456.

Flores, J. R., Montagna, J. M., & Vecchietti, A. (2014). An optimization approach for long term investments planning in energy. *Applied Energy, 122*, 162–178.

GAMS Development Corporation (2015). A user's guide. Washington, DC. Retrieved March 23, 2015 from http://www.gams.com/dd/docs/bigdocs/GAMSUsersGuide.pdf.

Gouveia, J. P., Dias, L., Martins, I., & Seixas, J. (2014). Effects of renewables penetration on the security of Portuguese electricity supply. *Applied Energy, 123*, 438–447.

Han, J. H., & Lee, I. B. (2011). Development of a scalable infrastructure model for planning electricity generation and CO_2 mitigation strategies under mandated reduction of GHG emission. *Applied Energy, 88*(12), 5056–5068.

Hellenic Electricity Market Operator S.A. (LAGIE). (2012). DAS Monthly Reports. Piraeus, Greece. Retrieved February 05, 2015 from http://www.lagie.gr/fileadmin/groups/EDRETH/DAS_Monthly_Reports/DASMonthlyReport_December2012.pdf.

Hellenic Electricity Market Operator S.A. (LAGIE). (2013). DAS Monthly Reports. Piraeus, Greece. Retrieved February 05, 2015 from http://www.lagie.gr/fileadmin/groups/EDRETH/DAS_Monthly_Reports/201312_DAS_Monthly_Report.pdf.

Hellenic Electricity Market Operator S.A. (LAGIE). (2012). RES & CHP Monthly Statistics. Piraeus, Greece. Retrieved February 05, 2015 from http://www.lagie.gr/fileadmin/groups/EDRETH/RES/2012_12_EN_MONTHLY_RES.pdf.

Hellenic Electricity Market Operator S.A. (LAGIE). (2013). RES & CHP Monthly Statistics. Piraeus, Greece. Retrieved February 05, 2015 from http://www.lagie.gr/fileadmin/groups/EDRETH/RES/2013_12_GR_MONTHLY_RES.pdf.

Hellenic Independent Power Transmission Operator S.A. (IPTO). (2014). Operations & Data. Athens, Greece. Retrieved February 01, 2015 from http://www.admie.gr/nc/en/home/.

Hellenic Independent Power Transmission Operator S.A. (IPTO). (2013). A study for the capacity assurance of the Greek power system for the period 2013–2020. Athens, Greece. Retrieved February 02, 2015 from http://www.admie.gr/fileadmin/groups/EDAS_DSS/Meleti_eparkeias_final.pdf.

Hobbs, B. (1995). Optimization methods for electric utility resource planning. *European Journal of Operational Research, 83*(1), 1–20.

Kagiannas, A., Askounis, D., & Psarras, J. (2004). Power generation planning: A survey from monopoly to competition. *International Journal of Electrical Power and Energy Systems, 26*(6), 413–421.

Kannan, R., & Turton, H. (2013). A long-term electricity dispatch model with the TIMES framework. *Environment Modeling and Assessment, 18*(3), 325–343.

Koltsaklis, N. E., & Georgiadis, M. C. (2015). A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. *Applied Energy, 158*, 310–331.

Koltsaklis, N. E., Dagoumas, A. S., Kopanos, G. M., Pistikopoulos, E. N., & Georgiadis, M. C. (2014). A spatial multi-period long-term energy planning model: A case study of the Greek power system. *Applied Energy, 115*, 456–482.

Koltsaklis, N. E., Liu, P., & Georgiadis, M. C. (2015). An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response. *Energy, 82*, 865–888.

Li, Y. P., Huang, G. H., & Chen, X. (2011). Planning regional energy system in association with greenhouse gas mitigation under uncertainty. *Applied Energy, 88*, 599–611.

Manfren, M., Caputo, P., & Costa, G. (2011). Paradigm shift in urban energy systems through distributed generation: Methods and models. *Applied Energy, 88*, 1032–1048.

Ming, Z., Kun, Z., & Liang, W. (2014). Study on unit commitment problem considering wind power and pumped hydro energy storage. *International Journal of Electrical Power and Energy Systems, 63*, 91–96.

Palmintier, B., & Webster, M. (2011). Impact of unit commitment constraints on generation expansion planning with renewables. *IEEE Power and Energy Society General Meeting, 2011*, 1–7.

Palmintier, B. S. (2013). Incorporating operational flexibility into electric generation planning: Impacts and methods for system design and policy analysis. PhD Thesis. Massachusetts Institute of Technology.

Purvins, A., Zubaryeva, A., Llorente, M., Tzimas, E., & Mercier, A. (2011). Challenges and options for a large wind power uptake by the European electricity system. *Applied Energy, 88*(5), 1461–1469.

Timilsina, G. R., Cornelis van Kooten, G., & Narbel, P. A. (2013). Global wind power development: Economics and policies. *Energy Policy, 61*, 642–652.

Van den Bergh, K., Delarue, E., & D'Haeseleer, W. (2013). Impact of renewables deployment on the CO_2 price and the CO_2 emissions in the European electricity sector. *Energy Policy, 63*, 1021–1031.

Xiao, J., Hodge, B.-M. S., Pekny, J. F., & Reklaitis, G. V. (2011). Operating reserve policies with high wind power penetration. *Computers and Chemical Engineering, 35*(9), 1876–1885.

Chapter 16
Modelling, Design and Control Optimization of a Residential Scale CHP System

Nikolaos A. Diangelakis and Efstratios N. Pistikopoulos

Abstract We present an analytical dynamic mathematical model and a simultaneous design and control optimization of a residential scale combined heat and power system (CHP). The mathematical model features a detailed description of the internal combustion engine based on a mean value approach, and simplified sub-models for the throttle valve, the intake and exhaust manifolds, and the external circuit. We treat the CHP unit as the interconnection of two distinct subsystems; the power production subsystem and the heat recovery subsystem. The validated zero-dimensional (0D) dynamic mathematical model of the system is implemented in gPROMS©, and used for optimization studies. A mixed-integer dynamic optimization problem is introduced that simultaneously determines the size of the internal combustion engine and the optimal control scheme of the CHP subsystems.

16.1 Introduction

Combined heat and power (CHP) systems constitute an upcoming technology that has the potential to replace the conventional processes used so far for the production of usable heat and electricity. The cogeneration of heat and power in a single process increases the system efficiency, while it decreases the operational cost. CHP systems utilize the same amount of fuel for the generation of both electrical power and usable heat, emitting consequently lower amount of exhaust gases. Therefore, owing to their

N.A. Diangelakis
Centre for Process Systems Engineering, Department of Chemical Engineering,
Imperial College London, London SW7 2AZ, UK
e-mail: nikolaos.diangelakis11@imperial.ac.uk

N.A. Diangelakis
Artie McFerrin Department of Chemical Engineering,
Texas A&M University, College Station, TX 77843, USA

E.N. Pistikopoulos (✉)
Artie McFerrin Department of Chemical Engineering and Texas A&M Energy Institute,
Texas A&M University, College Station, TX 77843, USA
e-mail: stratos@tamu.edu

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_16

environmentally friendly and cost effective nature, CHP systems can play a dominant role in the emission reduction strategies (OECD/IEA 2008). A review about CHP technologies and methodologies can be found in Diangelakis et al. (2014) and the references therein. Several works have been published on the investigation of the operation and supply related aspects of the CHP operation on a domestic level. In Table 16.1 we present indicative recent works of CHP modelling and CHP supply problems.

Designing economically profitable plants and improving their operational performance are the main goals of process design and optimal control, respectively. In a traditional hierarchical automation framework a two level sequential approach is applied where the process is designed first and then a controller is deployed for its efficient operation. Alternative approaches involve the solution of these two problems simultaneously in one mathematical framework. The integration of design and control has become an active area of systems engineering research and new trends continue to evolve. Over the years a number of attempts have been made to integrate (a) process feasibility and flexibility under uncertainty; (b) system controllability with regulatory control structure selection and its tuning; as well as (c) design and model predictive control with system stability and robustness into a single unified framework. Table 16.2 provides an indicative list of contributions in this area.

Table 16.1 Small scale cogeneration modelling and power supply review—indicative list

Authors (Year)	Contributions
Wu and Wang (2006)	Review on Combined cooling, heat and power cogeneration systems
Onovwiona and Ugursal (2006)	Technology review on heat and power cogeneration systems
Savola and Keppo (2005)	Small scale CHP model development for the investigation of its thermodynamic operation in part loads
Videla and Lie (2006)	Modelling and dynamic simulation of small scale internal combustion based CHP systems
Konstantinidis et al. (2010)	Multi-parametric model predictive control of small scale CHP systems suitable for domestic/residential use
Diangelakis et al. (2014)	First principle modelling and design optimisation of an internal combustion based residential scale CHP system
Savola and Fogelholm (2007)	MINLP formulation for increasing the power production of small scale biomass fueled CHP systems
Menon et al. (2013)	Optimal design of polygeneration systems, including CHP technologies, under optimal control assumptions
Fazlollahi et al. (2012)	Multi-objective optimisation techniques for the optimal design of complex energy systems, including CHP technologies
Mehleri et al. (2011, 2012)	Distributed energy systems optimisation, based on superstructure approach

Table 16.2 Integration of design and control—indicative list

Authors (Year)	Contributions
Lee et al. (1972)	Introduction to design and control
Narraway et al. (1991)	Steady state and dynamic economics
Brengel and Seider (1992)	Co-ordinated optimisation of design and control via NMPC
Luyben and Floudas (1994)	Superstructure of design alt. into MINLP
Mohideen et al. (1996)	Economically optimal design and control
Figueroa et al. (1996), Bahri et al. (1997)	Design & control/back-off optimization and flexibility & controllability analysis in process design
Fraga et al. (2000)	Automated process synthesis/design with system dynamics consideration
Bansal et al. (2000, 2002)	MIDO approach, PI control schemes
Van Schijndel and Pistikopoulos (2000)	Review on process design and operability
Kookos and Perkins (2001)	Bouncing scheme to reduce the relaxed design and control problem size
Lewin et al. (2002)	On the consideration of advanced control in the design optimisation
Sakizlis et al. (2003, 2004)	Simultaneous mpMPC and online design optimisation. Case studies on binary distilation column and evaporator
Seferlis and Georgiadis (2004)	The integration of design and control
Gani (2004), Hamid et al. (2010)	Review on design & control and model-based methodology for the integration of design and control
Olsen et al. (2005)	Control interactions with a vinyl acetate process process design
Malcolm et al. (2007)	Framework for design & control
Flores-Tlacuahuac and Biegler (2007)	MIDO for the simultaneous design and control problem solution
Ricardez-Sandoval et al. (2009)	Literature review on design and control integration
Würth et al. (2011)	Dynamic optimisation and NMPC
Yuan et al. (2012)	Review on design and control
Ricardez-Sandoval (2012), Sánchez-Sánchez and Ricardez-Sandoval (2013)	Probabilistic & robust approach to optimal design and control
Gebreslassie et al. (2012)	Design under uncertainty via multi-objective optimisation
Liu et al. (2013)	Trade offs between design and operation
Washington and Swartz (2014)	Multiperiod, parallel, mixed integer dynamic optimisation
Diangelakis et al. (2014)	Sequential design optimisation and mp-MPC. Application on a residential cogeneration systems
Vega et al. Vega et al. (2014a, b)	Review, classification and example of integrated design and control
Gutierrez et al. (2014)	Control structure selection for an MPC-based approach
Chi et al. (2015)	Design optimisation with data-driven control schemes

The current work focuses on the modeling, design and control optimization of a CHP system for a residential 10-house district. The CHP's mathematical model consists of a detailed model for a natural gas internal combustion engine (ICE) based on a mean value approach (Heywood 1989; Videla and Lie 2006) and simplified dynamic models for the throttle valve, the intake and exhaust manifolds, and the external circuit. The validated zero-dimensional (0D) dynamic mathematical model of the system is implemented in gPROMS© (Process Systems Enterprise 1997–2016). The objective of the design optimization is to estimate the optimum displacement volume of the internal combustion engine that minimizes the operational costs while satisfying the maximum electrical demand of a residential district. The optimization is performed in gOPT© with the implementation of a control vector parameterization algorithm via single-shooting.

16.2 CHP Process and Mathematical Model

The CHP system described in this work is illustrated in Fig. 16.1 and consists of the throttle valve, the intake and exhaust manifolds, the natural gas powered internal combustion engine and the cooling system. More specifically, the throttle valve manipulates the inlet air mass flow-rate while the intake manifold is responsible for the distribution of atmospheric air into the internal combustion engine cylinders. The natural gas powered internal combustion engine is the system's prime mover responsible for heat and power generation. The exhaust manifold, gathers the exhaust gases

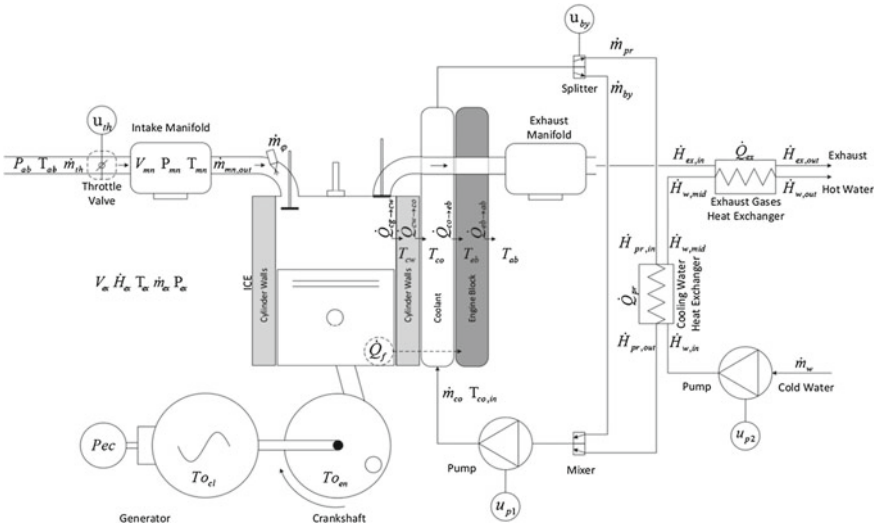

Fig. 16.1 Graphical representation of the CHP (acquired from Diangelakis et al. 2014)

from the internal combustion engine cylinders. The cooling system accounts for the internal combustion engine cooling while an external circuit interacts with the cooling system and the exhaust gases via heat exchangers.

The mass flow of the atmospheric air is controlled by the throttle valve and it is assumed that it enters the system at a relatively high pressure. Following that, the air passes through the intake manifold before blending with the injected fuel. The combustion of the air-fuel mixture in the engine cylinders results in the production of kinetic energy and heat. The kinetic energy is transferred through the engine crankshaft to the electrical generator. The heat is accumulated in the pistons, the cylinder walls and the engine block. The gases produced from the afore described chemical reaction are released through the exhaust manifold while the accumulated heat is induced by the engine cooling system. The coolant of the engine cooling system and the hot exhaust gases interact with an external circuit which consists of two heat exchangers and a volumetric pump for liquid mass flow-rate manipulation. With the aforementioned thermal interactions the following are achieved:

- The temperature of the recirculating coolant of the engine cooling system is reduced and eventually the engine temperature remains into certain bounds, thus preventing its malfunction.
- The temperature of the exhaust gases is dropped to a temperature range suitable for emission to the environment.
- The water of the external circuit is heated up.
- The electrical power produced by the electrical generator and the hot water of the external circuit are suitable for end use.

An individual system component approach is used in the mathematical modeling of the CHP. The model in this work assumes no losses of air, fuel or exhaust gases occur throughout the process. Furthermore, it is assumed that,

- The combustion of the atmospheric air with the directly injected fuel results into the production of carbon dioxide and water, in terms of chemical species,
- The gases in this work are considered ideal and perfect mixing among them is achieved and
- According to the mean value model approach, the total mass and energy balance of the reciprocating ICE is assumed to be in steady state.

16.2.1 Throttle Valve

A typical butterfly throttle valve controls directly the mass flow-rate of the atmospheric air entering the system; regulating, hence, indirectly the mass flow-rate of the injected fuel. The model describing the throttle valve is based on the following assumptions,

- No losses occur.
- The flow is compressible according to Guzzella and Onder (2010) (Eq. 16.1).
- Equation 16.2 accounts for both laminar and turbulent flow Moran and Shapiro (1992).

$$\dot{m}_{th} = c_d A_{th} \frac{P_{ab}}{\sqrt{R_\beta T_{ab}}} \psi \left(\frac{P_{ab}}{P_{mn} s} \right) \tag{16.1}$$

$$\frac{P_{ab}}{P_{mn}} < \frac{1}{2} \rightarrow \psi \left(\frac{P_{ab}}{P_{mn}} \right) = \frac{1}{\sqrt{2}} \tag{16.2a}$$

$$\frac{P_{ab}}{P_{mn}} \geq \frac{1}{2} \rightarrow \psi \left(\frac{P_{ab}}{P_{mn}} \right) = \sqrt{2 \frac{P_{mn}}{P_{ab}} \left(1 - \frac{P_{mn}}{P_{ab}} \right)} \tag{16.2b}$$

where $\dot{m}_{th}$ is the mass flow rate in the throttle valve, c_d is the throttle valve discharge coefficient, A_{th} is the throttle valve open area, P_{ab} is the ambient pressure, P_{mn} is the manifold pressure, R_β is the ideal gas constant and T_{ab} the ambient temperature.

The open area of the valve is determined by the angle formed by the butterfly and the valve walls. Assuming that the butterfly in closed valve position is tilted against the valve walls, Eqs. 16.3 are derived. Equation 16.3c is used to ensure continuity in the valve position setting during position changes.

$$\phi = \phi_0 + \left(\frac{\pi}{2} - \phi_0 \right) u_{th,cont} \tag{16.3a}$$

$$A_{th} = \frac{\pi D_{th}^2}{4} \left(1 - \frac{cos\phi}{cos\phi_0} \right) \tag{16.3b}$$

$$\frac{d}{dt} u_{th,cont} = l_{th} \left(u_{th} - u_{th,cont} \right) \tag{16.3c}$$

where ϕ is the angle between the butterfly and the valve walls, ϕ_0 is ϕ at a closed valve position, u_{th} is the discrete throttle valve control signal, D_{th} is the throttle valve diameter, and l_{th} represents the throttle valve latency in the butterfly position change.

A variable control volume approach is used to define the control volume of the valve (Eq. 16.4).

$$m_{th} = \frac{\pi D^3 sin(\phi)}{4cos(\phi_0)} \rho_{air,th}(T_{th}, P_{ab}, x_{air}) \tag{16.4}$$

where $\rho_{air,th}$ is the air density in the throttle valve T_{th} is the air temperature in the throttle valve and x_{air} is the air mass fraction.

Equations 16.5a and 16.5b describe the mass balance of the throttle valve during steady state and the energy balance respectively.

$$\frac{d}{dt}m_{th} = 0 \tag{16.5a}$$

$$m_{th}c_{p,air}\frac{d}{dt}T_{th} = \dot{m}_{th}(h_{th,in} - h_{th,out}) \tag{16.5b}$$

where $\dot{m}_{th,in}$ and $\dot{m}_{th,out}$ are the mass flow rate in the throttle valve inlet and utle respectively, $c_{p,air}$ is the mass specific heat capacity of the air, $h_{th,in}$ and $h_{th,out}$ are the air mass specific enthalpy in the inlet and the outlet of the throttle valve respectively.

16.2.2 Intake and Exhaust Manifold

The intake manifold is responsible for the even distribution of atmospheric air to the engine cylinders. In previous internal combustion engines, the intake manifold was also used for the premixing of the air-fuel mixture. In this work, the fuel is assumed to be sprayed directly into the engine cylinders, thus minimizing any fuel losses and maximizing the engine efficiency Heywood (1989). In this model, the intake manifold is assumed to be a small tank that manipulates the air intake pressure. The shape of the tank can be either rectangular parallelepiped or cylindrical (Eqs. 16.6)

$$V_{mn,cyl} = \frac{\pi D_{mn,cyl}^2}{4}L \tag{16.6a}$$

$$V_{mn,rec} = LHWi \tag{16.6b}$$

where $V_{mn,cyl}$ and $V_{mn,rec}$ are the volume of the cylindrical and rectangular parallelepiped manifold respectively, $D_{mn,cyl}$ is the diameter of the cylindrical manifold, L is the length, H is the height and Wi is the width of the rectangular parallelepiped manifold.

Given that the suggested intake manifold model is developed assuming (i) perfect mixing and (ii) ideal gas behavior Eqs. 16.7 describe the dynamic mass balance, Eqs. 16.8 the characterization of the air mass flow in the outlet of the tank and Eqs. 16.9 the dynamic energy balance.

$$\frac{d}{dt}m_{mn} = \dot{m}_{mn,in} - \dot{m}_{mn,out} \tag{16.7a}$$

$$m_{mn} = V_{mn}\rho_{air,mn} \tag{16.7b}$$

$$\dot{V}_{mn,out} = cpf(P_{mn} - P_{mn,out} \tag{16.8a}$$

$$\dot{m}_{mn,out} = \dot{V}_{mn,out}\rho_{air,mn} \tag{16.8b}$$

$$V_{mn}\frac{d}{dt}E = \dot{m}_{mn,in}h_{mn,in} - \dot{m}_{mn,out}h_{mn,out} \tag{16.9a}$$

$$E = \rho_{air,mn}\left(h_{mn} - \frac{P_{mn}}{\rho_{air,mn}}\right) \tag{16.9b}$$

where m_{mn} is the mass of air in the manifold, $\dot{m}_{mn,in}$ and $\dot{m}_{mn,out}$ are the inlet and outlet manifold mass flow rates of air respectively, V_{mn} is the manifold volume, $\rho_{air,mn}$ is the air density in the manifold, $\dot{V}_{mn,out}$ is the volumetric flow rate of air in the outlet of the manifold, cpf is the pressure–flow coefficient, $P_{mn,out}$ is the pressure in the outlet of the manifold, $\dot{m}_{mn,in}$ and $\dot{m}_{mn,out}$ is the mass flow rate in the inlet and the outlet of the manifold respectively, E is the manifolds internal energy, $h_{mn,in}$, $h_{mn,out}$ and h_{mn} is the air mass specific enthalpy in the inlet, the outlet and inside the manifold, respectively.

The shape and the volume of the manifold are set according to the engine specifications. The pressure of the manifold outlet is assumed to be equal to atmospheric pressure since the internal combustion engine is assumed to be naturally aspired. The exhaust manifold is modeled in a similar way. Note that, the exhaust manifold outlet pressure is set equal to 1.2 atm.

16.2.3 Internal Combustion Engine

The mean value model for the ICE presented in this section describes the behavior of an internal combustion engine using a volumetric pump approach. The model assumes "evenly distributed process and effects", of the ICE operation similarly to Guzzella and Onder (2010), Heywood (1989).

16.2.3.1 Engine Characterisitcs

Equation 16.10 describes the mass of the cylinder walls. The total wall volume is calculated using the cylinders geometrical characteristics (Eq. 16.11). Lubrication between the pistons and the cylinder walls is not taken into consideration in this work Payri et al. (2011), Aghdam and Kabir (2010).

$$m_{cw} = Vcw\rho_{cw} \tag{16.10}$$

$$V_{cw} = NoCS_{cyl}Wi_{cw}(2B_{cyl} + Wi_{cw})\frac{\pi}{4} \tag{16.11}$$

where M_{cw} is the mass of a single cylinder wall, V_{cw} is the volume of the cylinder wall, ρ_{cw} is the density of the cylinder wall material, NoC is the number of cylinders in the engine, S_{cyl} is the cylinder stroke, B_{cyl} is the cylinder bore and Wi_{cyl} is the width of the cylinder wall.

The use of simple cylindrical volume formulas and the stroke-to-bore ratio was used in order to correlate the stroke length and piston bore to the engine displacement volume, similarly to Rakopoulos et al. (2011) (Eq. 16.12).

$$V_d = NoC\frac{\pi B_{cyl}^2}{4}S_{cyl} \tag{16.12a}$$

$$SR = \frac{S_{cyl}}{B_{cyl}} \tag{16.12b}$$

where V_d is the engine displacement volume and SR is the cylinder stroke to bore ratio.

The engine block is assumed to be a rectangular parallelepiped the size of which is determined by the size of the piston bore, the piston stroke and the cylinder walls width. The cooling channels are assumed triangular and the pistons are assumed to be in series. The engine block surface area in contact with the coolant (Eq. 16.13a), the engine block area in contact with the ambient environment (Eq. 16.13b) and the cylinder wall surface area in contact with the coolant (Eq. 16.13c) are derived.

$$A_{eb} = NoC2\pi(B_{cyl} + 2Wi_{cw})S_{cyl}WS \tag{16.13a}$$

$$A_{ab} = 2(2Wi_{cw} + B_{cyl})(S_{cyl} + (2Wi_{cw} + B_{cyl})NoC + S_{cyl}) \tag{16.13b}$$

$$A_{co} = NoC\pi(B_{cyl} + 2Wi_{cw})S_{cyl}WS \tag{16.13c}$$

Equation 16.14 describes the mass of the coolant in the engine cooling channels.

$$m_{co} = NoC\frac{\sqrt{3}WS^2}{4}\pi^2(B_{cyl} + 2Wi_{cw})^2S_{cyl}\rho_{co} \tag{16.14}$$

where A_{eb} is the engine block area in contact with the cylinder walls, WS the wetting surface, A_{ab} is the engine area in contact with the ambient environment, A_{co} is the cooling channel area m_{co} is the mass of the coolant and ρ_{co} is the coolant density.

16.2.3.2 Mass Flow and Combustion

The atmospheric air is assumed to consist mainly of nitrogen and oxygen with ratio (0.79/0.21), while the fuel is assumed to be natural gas with the following composition (see Table 16.3).

The model does not take into consideration the reaction between nitrogen and oxygen. Consequently NOx emissions are neglected. The reaction between the

Table 16.3 Natural gas composition

	Mass fraction (%)
Methane	90.68
Ethane	4.46
Propane	0.52
n-Butane	0.21
n-Pentane	0.04
iso-Pentane	0.04
n-Hexane	0.05
Carbon Dioxide	1.83
Nitrogen	2.16

hydrocarbons and the atmospheric air is assumed to produce carbon dioxide and water (Eq. 16.15). The ICE operates in the presence of excess oxygen, the amount of which is defined according to the engine characteristics Williams (1985), Rausen et al. (2005). Equations 16.16 and 16.17 present the total mass balance of the combustion where represents the excess of atmospheric air. It is assumed that the ICE maintains fixed air to fuel ratio under any operational condition.

$$C_x H_y + \frac{4x+y}{4} O_2 + \frac{y}{2} H_2 O \tag{16.15}$$

$$\dot{m}_{ex} = \dot{m}_{mn,out} + \dot{m}_{\phi} \tag{16.16}$$

$$\dot{m}_{mn,out} = (1+\lambda)\sigma_0 \dot{m}_{\phi} \tag{16.17}$$

where m_{ex} is the exhaust gases mass flow rate, m_{ϕ} is the fuel mass flow rate, λ is the excessive air to fuel ratio and σ_0 is the stoichiometric air to fuel ratio.

Note that in an internal combustion engine that operates in a set point, the mass balance of each cylinder forms a cyclic steady state. Therefore, it is assumed that the total mass balance of the ICE is in steady state Guzzella and Onder (2010), Way (1976), Arsie et al. (1998).

16.2.3.3 Determination of the Angular Velocity

According to the MVM approach an ICE is modeled as a volumetric pump that aspires a volumetric flow proportional to its angular velocity Guzzella and Onder (2010). The latter is mathematically expressed in Eq. 16.18.

$$\dot{m}_{ex} = \frac{P_{mn}}{R_{\beta} T_{mn}} \frac{V_d}{4\pi} \eta_{vl} \omega_{en} \tag{16.18}$$

where η_{vl} is the engine volumetric efficiency and ω_{en} is the angular velocity.

The volumetric efficiency expresses the engine's ability to aspire a mixture of gases and describes how the engine differs from an ideal volumetric pump. In this ICE model, a multi-linear approach has been adopted for the volumetric efficiency (Eq. 16.19) Heywood (1989), Guzzella and Onder (2010), Onovwiona and Ismet (2007), Arsie et al. (1998), Cook and Powell (1988).

$$\eta_{vl} = \eta_{vl,1}(\omega_{en})\eta_{vl,2}(P_{mn}) \tag{16.19}$$

The term $\eta_{vl,1}(\omega_{en})$ represents the angular velocity dependent part and is assumed to be a quadratic function with a negative γ_2 term. This ensures maximum efficiency at the equivalent engine speed of 4500–5500 rounds per minute (Eq. 16.20).

$$\eta_{vl,1}(\omega_{en}) = \gamma_0 + \gamma_1\omega_{en} + \gamma_2\omega_{en}^2 \tag{16.20}$$

On the other hand, $\eta_{vl,2}(P_{mn})$ represents the dependence of the volumetric efficiency from pressure. For the development of Eq. 16.21 characteristics of the engine such as the displacement and compression volume (the chamber's volume when the cylinder is at the highest position, known as TDC) as well as the engine's upstream and downstream pressures, which are assumed constant, have been taken into account Guzzella and Onder (2010), Cook and Powell (1988).

$$\eta_{vl,2}(P_{mn}) = \frac{V_{cyl,c} + V_{cyl,d}}{V_{cyl,d}} - \frac{V_{cyl,c}}{V_{cyl,d}}\left(\frac{P_{ex}}{P_{mn}}\right)^{\frac{1}{1.4}} \tag{16.21}$$

where $V_{cyl,c}$ and $V_{cyl,d}$ are the compression and displacement volumes of a single cylinder and P_{ex} is the exhaust outlet pressure.

The engine compression volume is determined via its compression ratio (Eq. 16.22) which is assumed fixed, regardless the ICE operation level Angulo-Brown et al. (1994), Bhattacharyya (2000), Al-Hinti et al. (2008).

$$CR = \frac{V_{cyl,c}}{V_{cyl,d}} \tag{16.22}$$

16.2.3.4 Torque Generation

Equation 16.23 illustrates the torque production, using the characteristic displacement volume of the ICE and the break mean effective pressure.

$$To_{en} = \frac{P_{meb}V_d}{4\pi} \tag{16.23}$$

where To_{en} is the gernerated torque and P_{meb} is the break mean effective pressure.

The latter is the pressure that has to act during an expansion stroke over a piston in order for the amount of work produced by the piston to be the same as the work produced from the real engine in two engine revolutions for a four stroke engine Videla and Lie (2006). In naturally aspired engines, P_{meb} reflects the ability to induct air, the effectiveness of air utilization in combustion and the fuel conversion efficiency Heywood (1989). The break mean effective pressure is determined by taking into consideration the thermodynamic efficiency of the engine, the fuel mean effective pressure, the total external losses and the upstream and downstream pressure difference of the ICE inlet and oulet (Eq. 16.24).

$$P_{meb} = \eta_{td} P_{me\phi} - P_{mef} - P_{mepg} \tag{16.24}$$

where η_{td} is the thermodynamic efficiency of the engine, $P_{me\phi}$ is the fuel mean effective pressure, P_{mef} is the total external pressure losses and P_{mepg} is the engine inlet and outlet pressure difference.

The thermodynamic efficiency of the engine follows a quadratic correlation with the engine's angular velocity. The fuel mean effective pressure is the break mean effective pressure that an engine would have if its thermodynamic efficiency was equal to 100 % Guzzella and Onder (2010). Its expression is based on the lower heating value of the fuel (Eq. 16.25).

$$\eta_{td} P_{me\phi} = (\nu_0 + \nu_1 \omega_{en} + \nu_2 \omega_{en}^2) \frac{H_l \dot{m}_\phi 4\pi}{\omega_{en} V_d} \tag{16.25}$$

where ν_0, ν_1 and ν_2 are the engine efficiency coefficients and H_l is the fuel lower heating value.

The total external losses represent the friction in the engine and the pumping work required during the gas exchange process. Equation 16.26 demonstrates this estimation using a quadratic approach Videla and Lie (2006).

$$P_{mef} = (\beta_0 + \beta_1 \omega_{en} + \beta_2 \omega_{en}^2) \frac{4\pi}{V_d} \tag{16.26}$$

where β_0, β_1 and β_2 are the engine efficiency coefficients.

According to the MVM assumption, the exhaust outlet pressure is considered fixed throughout the engine cycle. Equation 16.27 represents the differences between the upstream and the downstream pressure of the ICE inlet and outlet.

$$P_{pemg} = \Delta P_{en} \tag{16.27}$$

16.2.3.5 Energy Balance

Equation 16.28 demonstrates the total energy balance of the combustion. The enthalpy difference between the combustion reactants and products is equal to the heat transferred to the cylinder walls, the work of compression and the usable work of expansion. The energy balance takes into account the amount of heat lost due to friction. The total energy balance is assumed to be in steady state, following the cyclic steady state operation of the ICE.

$$\begin{aligned}\dot{Q}_f + \dot{Q}_{cg\to cw} + \dot{W}_c + \dot{W}_{en} = {} & \dot{m}_{mn,out}\left(h_{mn,out} + \sum_{i=\text{air components}} x_{air,i}h^o_{f,air,i}\right) \\ & + \dot{m}_{\phi}\left(h_{\phi} + \sum_{j=\text{fuel components}} x_{\phi,j}h^o_{f,fuel,j}\right) \\ & - \dot{m}_{ex}\left(h_{ex} + \sum_{k=\text{exhaust components}} x_{ex,k}h^o_{f,ex,k}\right)\end{aligned} \tag{16.28}$$

where $h^o_{f,air,i}$, $h^o_{f,fuel,i}$ and $h^o_{f,ex,i}$ are the standard mass specific enthalpies of formation of the ith air,jth fuel and kth exhaust gas component respectively, h_{ϕ} and h_{ex} are the mass specific enthalpies of the fuel and exhaust gases respectively, $x_{air,i}$, $x_{\phi,j}$ and $x_{ex,k}$ are the ith air component, the jth fuel component and the kth exhaust gas component mass fraction respectively, $\dot{Q}_f$ are the heat loses due to friction, $\dot{Q}_{cg\to cw}$ is the heat transfer rate from the cylinder gasket to the cylinder wall, $\dot{W}_c$ is the work of combustion per time unit and $\dot{W}_{en}$ is the usable work per time unit.

The determination of the rate of heat release from the cylinder gasket to the cylinder walls follows a cubic function with respect to the break mean effective pressure (Eq. 16.29).

$$\dot{Q}_{cg\to cw} = \left(\delta_0 + \delta_1 P_{meb} + \delta_2 P^2_{meb} + \delta_3 P^3_{meb}\right)\frac{\omega_{en}}{4\pi} \tag{16.29}$$

The kinetic power which is transferred to the engine crankshaft and the work of expansion are presented in Eqs. 16.30a and 16.30b respectively.

$$\dot{W}_{en} = \omega_{en} To_{en} \tag{16.30a}$$

$$\dot{W}_c = \omega_{en}\frac{P_{mepg}V_d}{4\pi} \tag{16.30b}$$

Equations 16.31 and 16.32 used to define the lower heating value of the fuel. Equation 16.31 defines the standard enthalpy change of combustion.

$$\begin{aligned}\Delta\dot{H}_c = \dot{m}_{mn,out}\left(h_{mn,out} + \sum_{i=\text{air components}} x_{air,i}h^o_{f,air,i}\right)\\ + \dot{m}_{\phi}\left(h_{\phi} + \sum_{j=\text{fuel components}} x_{\phi,j}h^o_{f,fuel,j}\right)\\ - \dot{m}_{ex}\left(h_{ex} + \sum_{k=\text{exhaust comp}} x_{ex,k}h^o_{f,ex,k}\right)\end{aligned} \tag{16.31}$$

$$\Delta\dot{H}_c = H_l\dot{m}_{\phi}\frac{4\pi}{\omega_{en}V_d} \tag{16.32}$$

The heat loss due to friction losses is calculated in Eq. 16.33. The heat loss due to friction is not directly lost into the ambient environment but it is transferred to the engine.

$$\dot{Q}_f = \frac{P_{mef}V_d}{4\pi}\omega_{en} \tag{16.33}$$

The dynamic modeling of the heat transfer among the engine components, is undoubtedly of great importance for both the engine's and the CHP's operation simulation Arici et al. (1999). Equation 16.34 is a simple energy balance on the boundaries of the cylinder wall.

$$\frac{d}{dt}T_{cw} = \frac{\dot{Q}_{cq\to cw} - \dot{Q}_{cw\to co}}{m_{cw}c_{p,cw}} \tag{16.34}$$

where T_{cw} is the temperature of the cylinder walls.

Equation 16.35a calculates the heat rate of the convection between the cylinder walls and the coolant, while Eq. 16.35b describes the energy balance of the coolnat in the engine cooling channels.

$$\dot{Q}_{cw\to co} = TC_{co}A_{co}(T_{cw} - T_{co}) \tag{16.35a}$$

$$\frac{d}{dt}T_{co} = \frac{\dot{m}_{co}(h_{co,in} - h_{co,out}) + \dot{Q}_{cw\to co} - \dot{Q}_{co\to eb}}{m_{co}c_{p,co}} \tag{16.35b}$$

where $\dot{Q}_{cw\to co}$ is the heat transfer rate from the cylinder walls to the coolant, T_{co} is the coolant temperature, $\dot{m}_{co}$ is the coolant mass flow rate, $h_{co,in}$ and $h_{co,out}$ are the mass specific enthalpies of the coolant in the inlet and outlet of the cooling channels respectively, $\dot{Q}_{co\to eb}$ is the heat transfer rate from the coolant to the engine block and $c_{p,co}$ is the mass specific heat capacity of the coolant. TC_{co} is the coefficient of heat transfer between the cylinder walls and the coolant (assumed to be water).

Equation 16.36 is formulated in a similar way to Eq. 16.35 for the heat transfer among the coolant, the engine block and the ambient environment.

$$\dot{Q}_{co\to eb} = TC_{eb}A_{eb}(T_{co} - T_{eb}) \tag{16.36a}$$

$$\frac{d}{dt}T_{eb} = \frac{\dot{Q}_f + \dot{Q}_{co\to eb} - \dot{Q}_{eb\to ab}}{m_{eb}c_{p,eb}} \tag{16.36b}$$

$$\dot{Q}_{eb\to ab} = TC_{ab}A_{ab}(T_{eb} - T_{ab}) \tag{16.36c}$$

where TC_{eb} is the coefficient of heat transfer between the coolant and the engine block, T_{eb} is the temperature of the engine block, $\dot{Q}_{eb\to ab}$ is the heat transfer rate from the engine block to the ambient environment, m_{eb} is the mass of the engine block, $c_{p,eb}$ is the mass specific heat capacity of the engine block and TC_{ab} is the coefficient of heat transfer between the engine block and the ambient environment.

16.2.3.6 Electric Power Generation

The torque of the engine is transformed to generator torque via the crankshaft. The generator torque is calculated in Eq. 16.37.

$$\frac{d}{dt}To_{cl} = \frac{1}{Fl}(To_{en} - To_{cl}) \tag{16.37}$$

where To_{cl} is the torque and $1/Fl$ is used to denote the inertia of the flywheel.

An electric generator efficiency coefficient is used in Eq. 16.38 to account for any losses in electric power. The efficiency value is assumed fixed regardless the operation state of the engine or the generator.

$$Pec = \eta_{en}To_{cl}\omega_{en} \tag{16.38}$$

where Pec is the generated electrical power and η_{en} the generator electrical efficiency.

16.2.4 Engine Cooling System

The cooling system model describes the circulation of the coolant through the internal combustion engine and through the heat exchangers of the external circuit system. It consists of a simple pump, a stream splitter and a mixer. The purpose of the system is to keep the engine temperature between certain bounds, for efficiency purposes using water as a coolant Heywood (1989).

The pump manipulates the coolant mass flow rate ($\dot{m}_{co}$) and regulates the ICE temperature. The flow splitter divides the coolant flow into two streams. The stream characterized by ($\dot{m}_{pr}$) mass flow-rate interacts with the external circuit, while the

stream characterized by ($\dot{m}_{by}$) by-passes the interaction. The total steady state mass balance of the splitter is presented in Eq. 16.39a. Equations 16.39b and 16.39c demonstrates that the mass flow-rate of the by-pass stream is a fraction of the coolant mass flow-rate. The fraction is the control variable that manipulates the stream division for the more efficient temperature manipulation of the ICE components.

$$\dot{m}_{co} = \dot{m}_{pr} + \dot{m}_{by} \tag{16.39a}$$

$$\dot{m}_{by} = u_{by,cont}\dot{m}_{co} \tag{16.39b}$$

$$u_{by,cont} \in [0, 1] \tag{16.39c}$$

Equation 16.40 is used to ensure continuity of the control variable. The l_{by} is the splitter latency.

$$\frac{d}{dt}u_{by,cont} = l_{by}(u_{by} - u_{by,cont}) \tag{16.40}$$

where u_{by} is the control action of the splitter and $u_{by,cont}$ its continuous counterpart.

The steady state energy balance of the splitter is presented in Eq. 16.41. Adiabatic behavior of the splitter is assumed.

$$h_{co,out} = h_{by} = h_{pr,in} \tag{16.41}$$

where h_{by} is the mass specific enthalpy of the coolant stream that by-passes the interaction with the external circuit while $h_{pr,in}$ is the mass specific enthalpy of the coolant stream that interacts with the external circuit at the inlet of the splitter.

The flow mixer combines the two streams before they re-enter the engine's cooling channels. The temperature of the by-pass stream is assumed to remain constant. The temperature of the second stream is lowered due to the interaction with the external circuit. While the total mass balance of the mixer is identical to that of Eq. 16.40, the steady state energy balance differs (Eq. 16.42). The process is assumed adiabatic.

$$h_{co,in} = u_{by,cont}h_{by} + (1 - u_{by,cont})h_{pr,out} \tag{16.42}$$

16.2.5 External Circuit

The external circuit consists of two heat exchangers and a simple pump. It is used for the heat transfer through convection from the electricity production and converts the internal combustion engine, into an efficient cogeneration system. There are two sources of heat in this system:

- The recirculating engine coolant.
- The hot exhaust gases.

The interactions with the two heat sources will be hereon referred to as the stage one and stage two interaction, respectively. Regardless of the need for hot water

production, stage one interaction is vitally important to the engine's cooling. The heat transfer in this stage lowers the temperature of the recirculating engine coolant, keeping the engine temperature at relatively low levels. Its contribution to the CHP system is the initial rise of the external water temperature. In stage one a shell tube heat exchanger is used for heat transfer between liquids.

In stage two, the already heated external water interacts with the hot exhaust gases. In comparison to the stage one heat transfer, the purpose of stage two is purely to raise the temperature of the water to a higher temperature. A plate–fin heat exchanger is utilized. A simple heat transfer equation is used for the two heat exchangers (Eq. 16.43).

$$Q = UA\Delta T_{mean} \tag{16.43}$$

U is defined as the overall heat transfer coefficient and ΔT_{mean} denotes the logarithmic mean temperature difference and is defined in Eq. 16.44.

$$\begin{aligned} \Delta T_{out} > \Delta T_{small} \rightarrow \Delta T_{mean} &= \frac{\Delta T_{in} - \Delta T_{out}}{log\left(\left|\frac{\Delta T_{in}}{\Delta T_{out}}\right|\right)} \\ \Delta T_{out} \leq \Delta T_{small} \rightarrow \Delta T_{mean} &= \left(\frac{\Delta T_{in} - \Delta T_{out}}{log\left(\left|\frac{\Delta T_{in}}{\Delta T_{out}}\right|\right)}\right)\frac{\Delta T_{out}}{\Delta T_{small}} + \Delta T_{out}\left(1 - \frac{\Delta T_{out}}{\Delta T_{small}}\right) \end{aligned} \tag{16.44}$$

ΔT_{in} is defined as the temperature difference between the inlet streams of the heat exchangers, whereas ΔT_{out} as the temperature difference between its outlet streams. The overall heat transfer coefficient and the heat exchanging area of both the shell tube and plate–fin heat exchangers was based on Yokell (1990) and Afgan and Schünder (1974).

Throughout the CHP model Eq. 16.45 for the calculation of the density and enthalpy of the species are utilized.

$$\begin{aligned} \rho &= f(T, P, n) \\ h &= f(T, P, n) \end{aligned} \tag{16.45}$$

The integrated DAE system has a differentiation index of 1 and features 6 degrees of freedom which are the throttle valve outlet pressure, the pressure in the inlet and outlet of the ICE, the pressure of the exhaust gases the stream division and the stream flow of the engine cooling system. The model was implemented in gPROMS(©) Software Process Systems Enterprise (1997–2016). The complete model, as implemented in gPROMS(©) comprises of a total of 379 differential and algebraic equations.

The open loop simulation of the CHP system is presented in Fig. 16.2. The simulation represents the electrical power output of the system and the water temperature

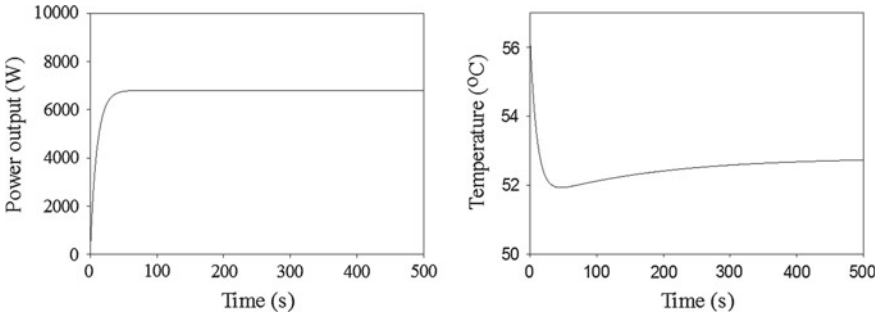

Fig. 16.2 Power output and external circuit water temperature (acquired from Diangelakis et al. 2014)

in the outlet of the external circuit. More specifically, the system was initiated from rest at t_0. The start-up, in this case, was chosen to be a step change in the throttle vale opening control variable u_{th} equal to 30 % of the maximum opening. The temperature of the water at the inlet of the external circuit was fixed at 25 °C. The water mass flow rate of the external circuit at t_0 was equal to 0 kg/s. A step change of 0.3 kg/s is applied to the external circuit pump. The splitter control variable u_{by} was set to 50 % and remained fixed. The manifolds were assumed rectangular parallelepiped and the displacement volume of the internal combustion engine was set to 1300 cc.

The proposed model follows the principles of Guzzella and Onder (2010) and Heywood (1989). Both authors have focused on ICE modeling and their work has been validated and in depth tested. The validation of the model through direct comparison with experimental data was not part of this work.

The dynamic approach to the behavior of the system tackles the discontinuities during changes in both the operational level of the power generation subsystem and the external circuit. In particular, a steady state modeling approach to the throttle valve and manifolds would result into a discontinuous power generation profile. This does not represent the system adequately as:

1. It implies discontinuous operation during operational level shifts.
2. It does not take into account the cold start-up of the system but rather assumes an immediate steady state power generation.

Furthermore, the modeling of the external circuit attempts to describe the dynamic behavior of the heat convection from the power generation subsystem to the water of the external circuit. A steady state approach would result into mass flow rate discontinuities and possibly physical constraint violation i.e. the rise of the temperature of liquid water above 100 °C. For further simulation sudies the reader is referred to Diangelakis et al. (2014).

16.3 Operation of the CHP System

Any cogeneration system can be normally viewed as comprised of two distinct subsystems: (a) the power generation subsystem and (b) the heat recovery subsystem which continuously interact with each other in order to produce usable heat and power through a single fuel source (Fig. 16.3).

Domestic cogeneration systems can be considered as multi-product processes. More specifically, the process cannot produce at the same time (a) electrical power at a desired level and (b) hot water of certain temperature and flow rate. The reason behind this is that the operation of the system is restricted by the electrical and heat efficiency of the prime mover, in this case the internal combustion engine. Figure 16.4 represents the ratio between the electrical power and the usable heat throughout the system's operation. In other words, consider the following example assuming that:

- The system operates at 4000 rpm,
- At this level a 0.5 kW electical power demand is satisfied,
- Water needs to be heated to 70 °C.

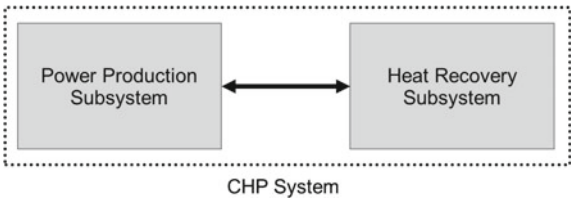

Fig. 16.3 Concept of cogeneration (acquired from Diangelakis et al. 2016)

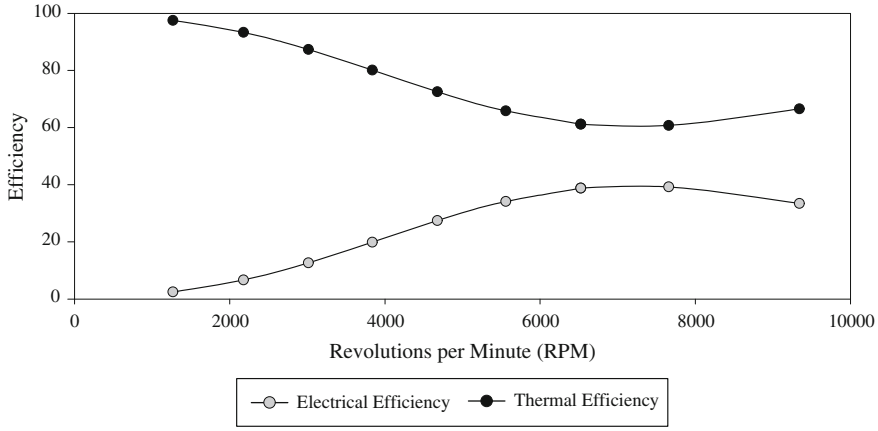

Fig. 16.4 CHP system heat to power ratio (acquired from Diangelakis et al. 2014)

Due to the restrictive power to heat ratio (in this case 20–80 %), 2 kW of the fuel's power is transformed into usable heat. The desired temperature difference can only be satisfied for a maximum of about 9 g/s of water flow rate, assuming standard liquid water heat capacity of 4.18 J/g K and inlet water temperature of 15 °C.

Equivalently, a desired flow rate of hot water of a certain temperature dictates the operating level of the CHP system, therefore the power production level cannot be determined independently. In most domestic applications, the temperature of hot water produced via the use of electrical or thermal boilers is fixed to temperatures close to 70 °C. Based on the capacity of the boilers, the flow rate of the water that can be heat up to such temperatures is determined. We have followed a similar approach where we have identified the two products of the CHP system as well as the two modes of operation of the system (also presented graphically in Fig. 16.5):

Operation mode 1: The power production driven operation denotes the operation during which a certain electrical output is guaranteed but the flow rate of the water is such that the produced hot water is of a certain temperature. In this mode of operation the main product is the electrical power.

Operation mode 2: The heat recovery driven operation corresponds to the mode of operation during which a certain flow rate of hot water of a certain temperature is guaranteed. During this mode of operation the power production level is not guaranteed. The main product in this case is the hot water.

It is clear from the above that a control policy that takes into account only one of the aforementioned operating modes is unable to capture the full production potential of the residential CHP plant. Furthermore, a CHP plant that ignores the ability to produce one of two products at a time, at a given set point, restricts the economic advantages that the process could have, i.e. at times of high electrical power demand, the operation of the domestic CHP plant could be such that electrical power production would partially cover the demand. The by-product hot water would be, in this case, stored in sufficiently large insulated tanks in order to cover the heating demands at a later time. This leads to reduced electricity costs from the electricity grid at times

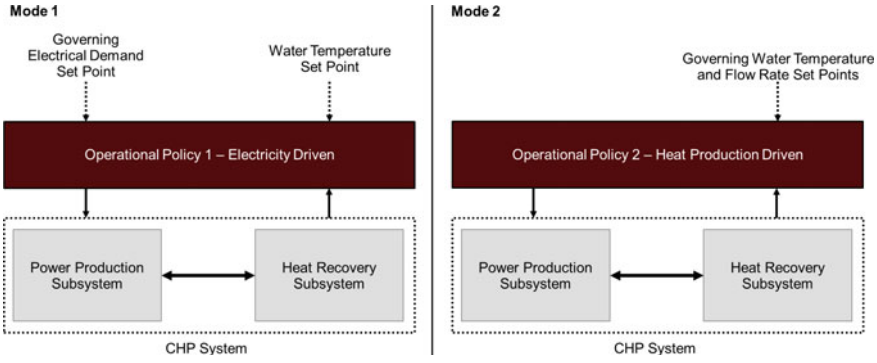

Fig. 16.5 Residential CHP modes of operation (acquired from Diangelakis et al. 2016)

of high demand, as well as reduced costs of water heating in the near future when hot water is required. On the contrary, during times of high heating demand the operation of the plant can be switched to mode 2, where hot water production is the driving force. The by-product electrical power can be immediately traded to the electrical grid, thus providing a small profit from the CHP usage. At this point, it should be noted that the problem of switching from one mode of operation to another for a single domestic CHP plant (or for a series of them) can be considered as a supply scheduling problem with economic evaluation criteria. More specifically, optimality in terms of operation of the plant during each operating mode is the objective of the control scheme presented in this work, while the economically optimal mode of the operation which inherently addresses the question of "which operating mode should be used and when" is among the objectives of a scheduling policy. The issue of the interactions between the control of the system and its scheduling will be visited at a later stage based on the principles presented in Baldea and Harjunkoski (2014).

It has become clear from the development of the process model that any CHP system can be considered as the coupling of *two subsystems*. Furthermore, the system can be viewed as a *two-product process*, according to the mode of *two modes of operation* (dual-operation) being used at every point in time.

16.4 Desing and Control Optimization via MIDO

Based on the inherent characteristics of the CHP system we derive a framework for the simultaneous design and control optimization of the CHP system at hand. The framework is based on the PAROC framework Pistikopoulos et al. (2015) and utilizes classical PI controllers, approximate state-space models based on the CHP subsystems and a Mixed-Integer Dynamic Optimization (MIDO) that determines (a) the system operation for different electrical and hot water needs at different times of the day (b) the optimal internal combustion engine size.

16.4.1 Model Approximation

The original high fidelity model described in the previous sections is highly complex. Although it is adequate to describe and simulate the system it is not suitable for control and advanced optimization studies. Therefore, an approximation step is required in order to reduce the system's complexity and produce model(s) suitable for MIDO. It is fundamental that the accuracy of the approximate model should be comparable to that of the original high fidelity model. It is clear that there is a trade–off between the two aspects. Following the principles described in Pistikopoulos et al. (2015) and Diangelakis et al. (2016), we derive approximate state–space models for the power generation subsystem and the heat recovery subsystem of the CHP system.

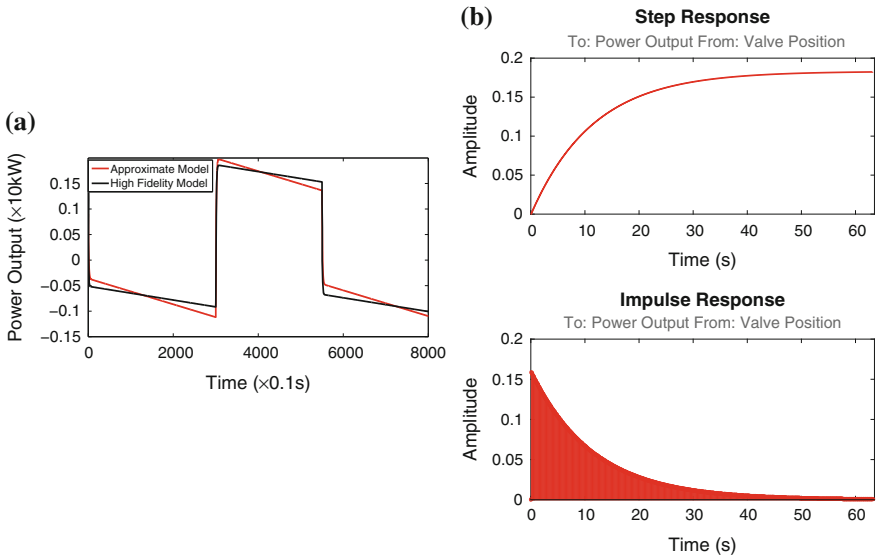

Fig. 16.6 Model approximation results for the powergeneration subsystem. **a** Output mismatch. **b** Step and impulse response

Figures 16.6 and 16.7 shows how the approximate models behave in comparison to the original high fidelity model for a set internal combustion engine design. Note that the approximate models used in this work are design dependent (Eqs. 16.46 and 16.47).

$$\begin{aligned} \frac{d}{dt}x &= -0.2029\,x(t) + 0.06393\,u_{th}(t) + 6.582 \times 10^{-7}\,V_d \\ Pec(t) &= 7.839\,x(t) \end{aligned} \tag{16.46}$$

where u_th is the discrete throttle valve control signal, Pec is the electrical power generated from the power generation subsystem and V_d is the displacement volume of the internal combustion engine, in this case treated as a measured uncertainty.

$$\begin{aligned} \frac{d}{dt}x &= \begin{bmatrix} -1.401 & -0.01659 \\ -1.316 & -2.282 \end{bmatrix} x(t) + \begin{bmatrix} 0.129 \\ -0.3904 \end{bmatrix} \quad Pec(t) + \begin{bmatrix} 1.121 \\ 5.467 \end{bmatrix} u_{p_2}(t) \\ T_{water}(t) &= \begin{bmatrix} 4.584 & 36.27 \end{bmatrix} x(t) \end{aligned} \tag{16.47}$$

where u_{p_2} is the water flow rate of the external circuit (heat recovery subsystem) and T_{water} is the temperature of water at the outlet of the heat recovery subsystem. Based on the mode of operation of the CHP unit, Pec is treated as a control variable when the unit is operated in a hot water production driven mode and u_{p_2} is treated as uncertain. The value for the latter is determined by the hot water demand. The exact opposite holds true for a power production driven mode of operation.

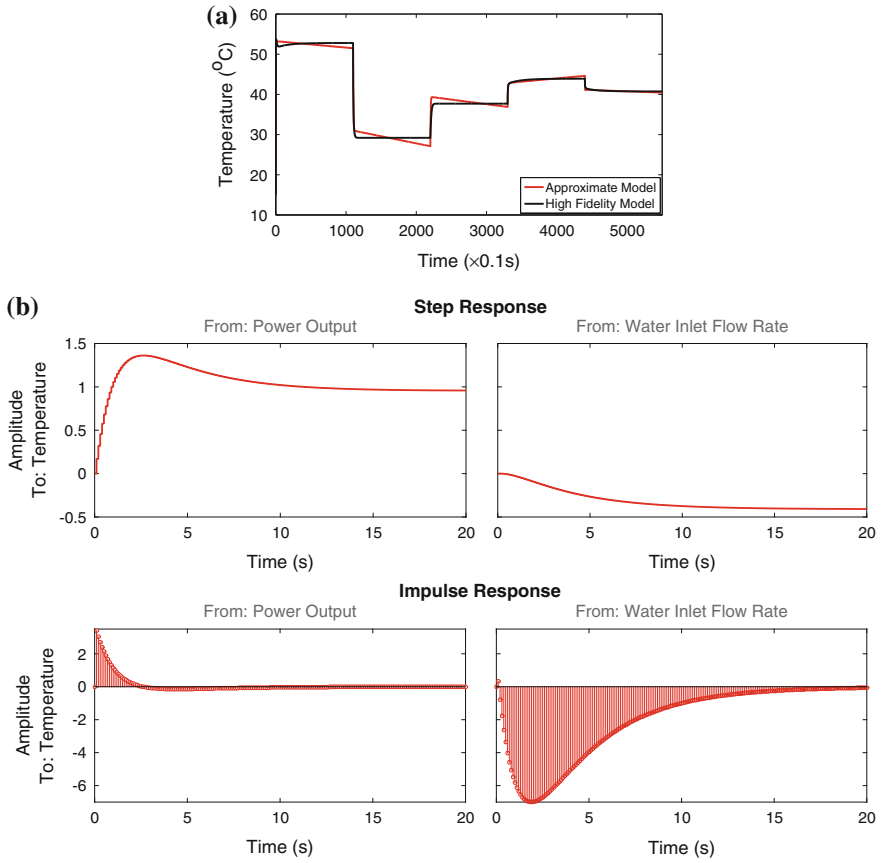

Fig. 16.7 Model approximation results for the heat recovery subsystem. **a** Output mismatch. **b** Step and impulse response

Note that the state variables for both the approximate models don't have a physical meaning as the models are derived via the System Identification Toolbox of MATLAB(©).

16.5 Simultaneous Design and Control Optimization

The approximate models derived in the previous section for the two subsystems are used to design and solve the simultaneous equipment design and control optimization. The objective of the optimization is the derivation of:

- The size of the internal combustion engine displacement volume V_d,
- The characteristics of the PI controllers that control the CHP system,
- The mode of operation of the CHP system.

The first two objectives are decisions that remain fixed throughout the lifetime of the plant but the mode of operation is subject to change based on several criteria. Most residential scale CHP units are operated in a heat production driven manner thus treating the electrical power as a by-product. Through this work we attempt to show that such an approach is erroneous as (a) the demand for hot water and electricity and (b) the price for buying electrical power and fuel from a central grid vary with respect to time even during the day. In order to ensure that outage in electrical power or hot water is prevented, we assume that the CHP unit is operated in a residential environment where the houses are connected to the electrical power grid as well as a central natural gas network. The electrical power that is not covered by the CHP operation is purchased from the grid and hot water produced in quantities less than the demand (or in lower temperatures than the demand) is countered by the use of a natural gas boiler. Similarly, excess electrical power is traded to the grid but excess hot water is discarded at a cost.[1] The following subsections describe the PI control scheme used in this work and the handling of the discrete component of the CHP system operation.

16.5.1 Classical Control

The CHP system is controlled by classical PI controllers. Three controllers are utilised.

1. The controller for the power generation subsystem manipulates the throttle valve opening to produce electrical power of a certain setpoint.
2. The controller for the heat recovery subsystem:
   - Manipulates the water flow rate of the external circuit of the subsystem to produce water of a certain temperature setpoint while the power generation subsystem follows the an electrical demand or
   - It manipulates the required electrical power in such a manner so that a temperature setpoint is met for a water stream of a certain flow. In that case the electrical power demand is dictated by the PI controller of the heat recovery subsystem.
3. A third PI controller is required for the heat recovery driven operation that controls the water flow in the external circuit.

The general form of the PI controller used in this work is presented in Eq. 16.48.

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) dt \tag{16.48}$$

[1] In future works, the utilization of a hot water storage tank is worth investigating in.

where $u(t)$ is the control variable, $e(t)$ is defined as the deviation of the output compared to a setpoint. K_i and K_p are the design parameters of control and are subjected to optimization.

16.5.2 Handling the Discrete Alternatives

In order to handle the discrete alternatives in the MIDO problem a smooth approximation is used similarly to Bansal (2000). According to this a function is used to relax the binary variables to continuous. It is well known that such an approach can settle for relaxed values that are different than 0 or 1. The formulation we use in this work though is free of such a constraint. The underlying idea of such an approximation is presented in Eq. 16.49.

$$y = tanh(My_{relax}) \tag{16.49}$$

where $y \in \{0, 1\}$ and $y \in [0, 1]$. M is a large number. The function in Eq. 16.49 settles to 0 when $y_{relax} = 0$ and to 1, otherwise. The key for the successful application of the method is the choice of M. As mentioned before, the entire optimization problem is handled by gPROMS(©). By setting the effective zero of the optimization algorithm of the software to a value of 10^{-x}, where x is a positive number and $M = 10^x$ we achieve the following: the only values of y_{relax} that would result into a value of y other than 0 or 1 are in the scale of 10^{-x}. Such values of x though are treated as zero by the optimization algorithm because of the effective zero setting. In this work $x = 8$, therefore $M = 10^8$ and the effective zero is set to 10^{-8}.

16.5.3 Cost Function

The cost function of the dynamic optimization problem takes into consideration several aspects.

- The annualized cost of acquisition of the CHP plant is a function of the size of the internal combustion engine. For this, a 25 year life time has been assumed.
- The cost of electricity from the grid is a function of time. In particular, electricity from the grid is cheaper to consume during night time than day time. The difference can be up to 30 % according to the regulations of the electricity provider.
- The cost of the natural gas does not fluctuate with time. Note that typically an internal combustion engine with a larger displacement volume consumes more fuel for a certain valve opening setting than an engine with a smaller displacement volume. Furthermore, the size of the engine affects the thermal and electrical efficiency of the plant. Therefore a larger engine could be more useful for a standalone system whereas a smaller one could be more suitable for every-day mild use.

- The trading of electricity with the grid is taken into consideration in the objective function. Fluctuation to the price is also taken into consideration.
- Scaled terms that ensure the minimization of the error associated with the control formulation are also present in the objective function formulation.

Following these points the objective function is of the form of Eq. 16.50.

$$\begin{aligned} f(V_d, K_p, K_i, y_{relax}, t) = &\text{ Cost of equipment acquisition} \\ &+ \text{ Cost of production} \\ &+ \text{ Revenue of trading} \\ &+ \text{ Scaled deviation penalty} \end{aligned} \tag{16.50}$$

The overall form of the design and control optimization problem is presented in Eq. 16.51.

$$\begin{aligned} \min_{d(t),K_p,K_i,y_{relax}(t)} \quad & J = \int_0^t f(d(\tau), K_p, K_i, y_{relax}(\tau), \tau)dt \\ \text{s.t.} \quad & \frac{d}{dt}x = A\,x(t) + B\,u_(t) + C\,d_(t) \\ & y(t) = D\,x(t) \\ & u(t) = y_{bin}(t)(u(t) = K_p e(t) + K_i \int_0^t e(\tau)dt) \\ & e(t) = y_{sp}(t) - y(t) \\ & y_{bin}(t) = tanh(My_{relax}(t)) \\ & 0 \leq K_p \leq K_{p,max} \\ & 0 \leq K_i \leq K_{i,max} \\ & 0 \leq y_{relax}(t) \leq 1 \\ & d_{min}(t) \leq d(t) \leq d_{max}(t) \end{aligned} \tag{16.51}$$

where $x(t)$ are the states, $u(t)$ are the control variables, $y(t)$ and $y_{sp}(t)$ are the outputs and output setpoints respectively, $d(t)$ are the disturbances, $e(t)$ are the setpoint deviations, $y_{bin}(t)$ are the binary variables, y_{relax} are the relaxed variables, K_i and K_p are the integral and proportional gains for the controllers and M a sufficiently large number. Note that the displacement volume of the internal combustion engine is a design variable that remains fixed throughout the optimization horizon. The same holds true for the gains of the PI controllers.

The objective of this optimization problem is to show the significance of the dual control approach and the effect of the design. More specifically, the fluctuation of (a) the electricity price and the (b) electrical power and (c) hot water demand throughout a time span $0 \rightarrow t$ should affect the operation of the system in such a manner that shows that both operating modes are of significance (i.e. an alternating power generation driven and heat recovery driven operation is expected to occur as a function of

the aforementioned fluctuations). Furthermore, the size of the internal combustion engine displacement volume is expected to affect not only the control setting of the degrees of freedom of the system but the ratio of electricity and heat that is produced from the CHP or is bought from the central supply.

16.6 Optimization Results and Discussion

The optimization problem is performed for 1000 time intervals. The profile of the hot water and power demand during that time period is presented in Fig. 16.8a, b. Note that the temperature of the hot water is 70 °C. The CHP system is allowed to cover partial demands in both operating modes. The price of electricity drops between 21st and 310th time interval by 70 %, a fluctuation that reflects the night time electricity price. The natural gas price is set such that an ideal gas boiler and an ideal electrical boiler would produce the same amount of hot water for the same cost. The results show that for "cheap electricity" time periods the hot water production driven operation is favorable, while the electricity driven operation is favorable for most of the rest of the time period. The fact that hot water production driven operation is favorable in some time intervals of "expensive electricity" is a result of the demand profile for the products, i.e. there is a threshold where even when electricity is expensive to buy from the grid, hot water production is favorable due to the demand ratio. Furthermore, it is clear from the results that the demands in rare cases are covered entirely by the CHP system. A combination of CHP operation and grid usage is more favorable. The PI control is able to adjust the control settings so that the setpoints are met, for the optimized engine displacement volume. The results of the optimization problem are presented in Table 16.4.

Figure 16.9 shows the switch between the operations. It is clear from the results that, in terms of cost, the heat recovery driven operation is more favorable though both operating modes are significant in the overall operation of the plant.

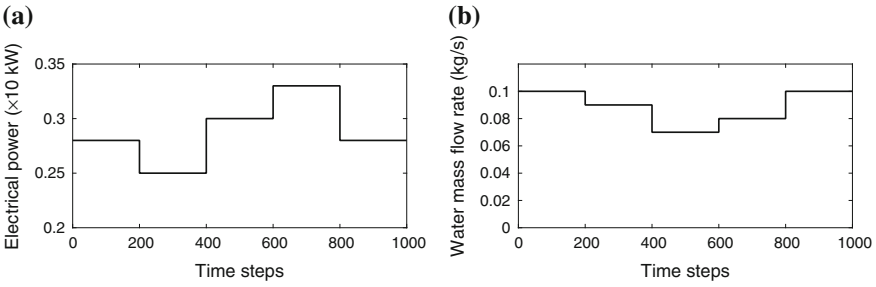

Fig. 16.8 Electrical power and hot water mass flow-rate demand considered in the optimization formulation. **a** Power demand. **b** Hot water demand

Table 16.4 Optimization result for the engine displacement volume and the controller parameters

Optimization variable	Value	
	Heat recovery driven operation	Electricity driven operation
K_i for power generation subsystem	9948.278	
K_p for power generation subsystem	8.032	
K_i for heat recovery subsystem	987.775	0.022
K_p for heat recovery subsystem	25.817	0.567
V_d engine displacement volume	1478.5 cc	

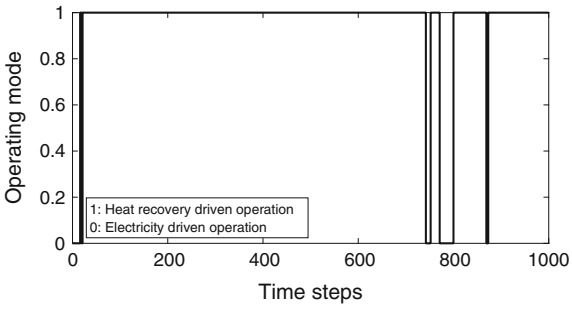

Fig. 16.9 Operation switch between electricity driven and heat recovery driven operation (1: heat recovery driven operation)

16.6.1 Future Steps

Two future steps directions can be identified from the current work:

- An advanced control scheme can be developed and implemented in the design and control framework. More specifically, the application of a model based control methodology would improve the approach as, through that approach, the ability to account for constraints is given. Furthermore, the process design aspect can be explicitly incorporated into the control framework. A multi-parametric model predictive control approach that treats the process design variables as parameters can be incorporated into the current framework seamlessly as the optimal control action functions (a) are available prior to the design optimization problem and (b) include the design variables in the formulation, i.e. $u = f(x, d, y, y_{sp})$, where x are the system states, d is the design variable, y is the system output y_{sp} is the output set point and f is an affine function. The reader is referred to Pistikopoulos and Diangelakis (2015) for more information.
- A scheduling component should be incorporated that accounts for long term operation. It is clear from the alternation between electricity and hot water driven operation that the electricity price is not enough to determine the operating mode of the system. Therefore, a scheduler that takes into account economic criteria and provides the operating set points for the CHP system is necessary Kopanos and Pistikopoulos (2014).

Acknowledgments Financial support from EPSRC (EP/I014640), Texas A&M University and Texas A&M Energy institute is gratefully acknowledged.

References

Afgan, N., & Schünder, E. (1974). *Heat Exchangers: Design and theory sourcebook*. Scripta Book Company.

Aghdam, E., & Kabir, M. (2010). Validation of a blowby model using experimental results in motoring condition with the change of compression ratio and engine speed. *Experimental Thermal and Fluid Science*, *34*(2), 197–209.

Al-Hinti, I., Akash, B., Abu-Nada, E., & Al-Sarkhi, A. (2008). Performance analysis of air-standard diesel cycle using an alternative irreversible heat transfer approach. *Energy Conversion and Management*, *49*(11), 3301–3304.

Angulo-Brown, F., Fernández-Betanzos, J., & Diaz-Pico, C. (1994). Compression ratio of an optimized air standard otto-cycle model. *European Journal of Physics*, *15*(1), 38–42.

Arici, O., Johnson, J., & Kulkarni, A. (1999). The vehicle engine cooling system simulation part 1—Model development. *SAE Technical Paper 1999-01-0240*.

Arsie, I., Pianese, C., & Rizzo, G. (1998). Models for the prediction of performance and emissions in a spark ignition engine—A sequentially structured approach. *SAE Technical Paper*, *980779*.

Bahri, P., Bandoni, J., & Romagnoli, J. (1997). Integrated flexibility and controllability analysis in design of chemical processes. *AIChE Journal*, *43*(4), 997–1015.

Baldea, M., & Harjunkoski, I. (2014). Integrated production scheduling and process control: A systematic review. *Computers & Chemical Engineering*, *71*, 377–390.

Bansal, V. (2000). *Anaylsis, design and control optimization of process systems under uncertainty*. Ph.D. thesis, Imperial College, London.

Bansal, V., Perkins, J. D., & Pistikopoulos, E. N. (2000). Flexibility analysis and design of linear systems by parametric programming. *AIChE Journal*, *46*(2), 335–354.

Bansal, V., Perkins, J. D., & Pistikopoulos, E. N. (2002). Flexibility analysis and design using a parametric programming framework. *AIChE Journal*, *48*(12), 2851–2868.

Bhattacharyya, S. (2000). Optimizing an irreversible diesel cycle—Fine tuning of compression ratio and cut-off ratio. *Energy Conversion and Management*, *41*(8), 847–854.

Brengel, D. D., & Seider, W. D. (1992). Coordinated design and control optimization of nonlinear processes. *Computers & Chemical Engineering*, *16*(9), 861–886.

Chi, R., Hou, Z., Huang, B., & Jin, S. (2015). A unified data-driven design framework of optimality-based generalized iterative learning control. *Computers and Chemical Engineering*, *77*, 10–23.

Cook, J., & Powell, B. (1988). Modeling of an internal combustion engine for control analysis. *IEEE Control Systems Magazine*, *8*(4), 20–26.

Diangelakis, N. A., Avraamidou, S., & Pistikopoulos, E. N. (2016). Decentralized multiparametric model predictive control for domestic combined heat and power systems. *Industrial & Engineering Chemistry Research*, *55*(12), 3313–3326.

Diangelakis, N. A., Manthanwar, A. M., & Pistikopoulos, E. N. (2014). A framework for design and control optimisation: Application on a chp system. In *Computer aided chemical engineering. Proceedings of the 8th international conference on foundations of computer-aided process design* (Vol. 34, pp. 765–770). Elsevier.

Diangelakis, N. A., Panos, C., & Pistikopoulos, E. N. (2014). Design optimization of an internal combustion engine powered CHP system for residential scale application. *Computational Management Science*, *11*(3), 237–266.

Fazlollahi, S. B., Mandel, P., Becker, G., & Maréchal, F. (2012). Methods for multi-objective investment and operating optimization of complex energy systems. *Energy*, *45*(1), 12–22.

Figueroa, J., Bahri, P., Bandoni, J., & Romagnoli, J. (1996). Economic impact of disturbances and uncertain parameters in chemical processes—A dynamic back-off analysis. *Computers and Chemical Engineering*, *20*(4), 453–461.

Flores-Tlacuahuac, A., & Biegler, L. T. (2007). Simultaneous mixed-integer dynamic optimization for integrated design and control. *Computers & Chemical Engineering*, *31*(5–6), 588–600.

Fraga, E., Hagemann, J., Estrada-Villagrana, A., & Bogle, I. (2000). Incorporation of dynamic behaviour in an automated process synthesis system. *Computers and Chemical Engineering*, *24*(2–7), 189–194.

Gani, R. (2004). Chemical product design: Challenges and opportunities. *Computers and Chemical Engineering*, *28*(12), 2441–2457.

Gebreslassie, B., Yao, Y., & You, F. (2012). Design under uncertainty of hydrocarbon biorefinery supply chains: Multiobjective stochastic programming models, decomposition algorithm, and a comparison between CVaR and downside risk. *AIChE Journal*, *58*(7), 2155–2179.

Gutierrez, G., Ricardez-Sandoval, L., Budman, H., & Prada, C. (2014). An MPC-based control structure selection approach for simultaneous process and control design. In *Computers and chemical engineering*.

Guzzella, L., & Onder, C. H. (2010). *Introduction to modeling and control of internal combustion engine systems* (2nd ed.). Springer.

Hamid, M., Sin, G., & Gani, R. (2010). Integration of process design and controller design for chemical processes using model-based methodology. *Computers and Chemical Engineering*, *34*(5), 683–699.

Heywood, J. B. (1989). *Internal combustion engine fundamentals*. McGraw-Hill, Inc.

Konstantinidis, D., Varbanov, P., & Klemeš, J. (2010). Multi-parametric control and optimisation of a small scale CHP. *Chemical Engineering Transactions*, *21*, 151–156.

Kookos, I., & Perkins, J. (2001). An algorithm for simultaneous process design and control. *Industrial and Engineering Chemistry Research*, *40*(19), 4079–4088.

Kopanos, G. M., & Pistikopoulos, E. N. (2014). Reactive scheduling by a multiparametric programming rolling horizon framework: A case of a network of combined heat and power units. *Industrial & Engineering Chemistry Research*, *53*(11), 4366–4386.

Lee, H., Koppel, L., & Lim, H. (1972). Integrated approach to design and control of a class of countercurrent processes. *Industrial and Engineering Chemistry: Process Design and Development*, *11*(3), 376–382.

Lewin, D., Seider, W., & Seader, J. (2002). Integrated process design instruction. *Computers and Chemical Engineering*, *26*(2), 295–306.

Liu, P., Georgiadis, M. C., & Pistikopoulos, E. N. (2013). An energy systems engineering approach for the design and operation of microgrids in residential applications. *Chemical Engineering Research and Design*, *91*(10), 2054–2069. The 60th Anniversary of the European Federation of Chemical Engineering (EFCE).

Luyben, M. L., & Floudas, C. A. (1994). Analyzing the interaction of design and control-1. A multiobjective framework and application to binary distillation synthesis. *Computers & Chemical Engineering*, *18*(10), 933–969.

Malcolm, A., Polan, J., Zhang, L., Ogunnaike, B., & Linninger, A. (2007). Integrating systems design and control using dynamic flexibility analysis. *AIChE Journal*, *53*(8), 2048–2061.

Mehleri, E. B., Sarimveis, H., Markatos, N., & Papageorgiou, L. (2011). Optimal design and operation of distributed energy systems. *Computer Aided Chemical Engineering*, *29*, 1713–1717.

Mehleri, E. D., Papageorgiou, L. G., Markatos, N. C., & Sarimveis, H. (2012). A model predictive control framework for residential microgrids. *Computer Aided Chemical Engineering*, *30*, 327–331.

Menon, R., Paolone, M., & Maréchal, F. (2013). Study of optimal design of polygeneration systems in optimal control strategies. *Energy*, *55*, 134–141.

Mohideen, M. J., Perkins, J. D., & Pistikopoulos, E. N. (1996). Optimal synthesis and design of dynamic systems under uncertainty. *Computers & Chemical Engineering*, *20*(Supplement 2(0)), S895–S900.

Moran, M., & Shapiro, H. N. (1992). *Fundamentals of engineering thermodynamics*. New York: Wiley.

Narraway, L., Perkins, J., & Barton, G. (1991). Interaction between process design and process control: Economic analysis of process dynamics. *Journal of Process Control*, *1*(5), 243–250.

Olsen, D., Svrcek, W., & Young, B. (2005). Plantwide control study of a vinyl acetate monomer process design. *Chemical Engineering Communications*, *192*(10–12), 1243–1257.

Onovwiona, H., & Ugursal, V. (2006). Residential cogeneration systems: Review of the current technology. *Renewable and Sustainable Energy Reviews*, *10*(5), 389–431.

Onovwiona, H. I., Ugursal, V. I., & Fung, A. S. (2007). Modeling of internal combustion engine based cogeneration systems for residential applications. *Applied Thermal Engineering*, *27*(5–6), 848–861.

Organisation for economic co-operation and development/International Energy Agency. (2008). Combined Heat and Power. *Technical report, OECD/IEA Head of communication and information office*, 9 rue de la Fédération, 75739 Paris Cedex 15, France.

Payri, F., Olmeda, P., Martn, J., & Garca, A. (2011). A complete 0D thermodynamic predictive model for direct injection diesel engines. *Applied Energy*, *88*(12), 4632–4641.

Pistikopoulos, E., & Diangelakis, N. (2015). Towards the integration of process design, control and scheduling: Are we getting closer? *Computers and chemical engineering*.

Pistikopoulos, E., Diangelakis, N., Oberdieck, R., Papathanasiou, M., Nascu, I., & Sun, M. (2015). PAROC—An integrated framework and software platform for the optimisation and advanced model-based control of process systems. *Chemical Engineering Science*, *136*, 115–138.

Process Systems Enterprise. gPROMS, 1997–2016.

Rakopoulos, C., Kosmadakis, G., Dimaratos, A., & Pariotis, E. (2011). Investigating the effect of crevice flow on internal combustion engines using a new simple crevice model implemented in a CFD code. *Applied Energy*, *88*(1), 111–126.

Rausen, D., Stefanopoulou, A., Kang, J.-M., Eng, J., & Kuo, T.-W. (2005). A mean-value model for control of homogeneous charge compression ignition (HCCI) engines. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, *127*(3), 355–362.

Ricardez-Sandoval, L. (2012). Optimal design and control of dynamic systems under uncertainty: A probabilistic approach. *Computers and Chemical Engineering*, *43*, 91–107.

Ricardez-Sandoval, L., Budman, H., & Douglas, P. (2009). Integration of design and control for chemical processes: A review of the literature and some recent results. *Annual Reviews in Control*, *33*(2), 158–171.

Sakizlis, V., Perkins, J. D., & Pistikopoulos, E. N. (2003). Parametric controllers in simultaneous process and control design optimization. *Industrial & Engineering Chemistry Research*, *42*(20), 4545–4563.

Sakizlis, V., Perkins, J. D., & Pistikopoulos, E. N. (2004). Recent advances in optimization-based simultaneous process and control design. *Computers & Chemical Engineering*, *28*(10), 2069–2086.

Sánchez-Sánchez, K., & Ricardez-Sandoval, L. (2013). Simultaneous process synthesis and control design under uncertainty: A worst-case performance approach. *AIChE Journal*, *59*(7), 2497–2514.

Savola, T., & Fogelholm, C.-J. (2007). MINLP optimisation model for increased power production in small-scale CHP plants. *Applied Thermal Engineering*, *27*(1), 89–99.

Savola, T., & Keppo, I. (2005). Off-design simulation and mathematical modeling of small-scale CHP plants at part loads. *Applied Thermal Engineering*, *25*(8–9), 1219–1232.

Seferlis, P., & Georgiadis, M. (2004). The integration of process design and control-summary and future directions. *Computer Aided Chemical Engineering*, *17*(C), 1–9.

Van Schijndel, J., & Pistikopoulos, E. (2000). Towards the integration of process design, process control and process operability—Current stares and furore trends. *Foundations of Computer-Aided Process Design*, *96*, 99–112.

Vega, P., Lamanna, R., Revollar, S., & Francisco, M. (2014a). Integrated design and control of chemical processes—Part II: An illustrative example. *Computers and Chemical Engineering*.

Vega, P., Lamanna de Rocco, R., Revollar, S., & Francisco, M. (2014b). Integrated design and control of chemical processes—Part I: Revision and classification. *Computers and Chemical Engineering*.

Videla, J. I., & Lie, B. (2006). Simulation of a small scale SI ICE based cogeneration system in modelica/dymola. In *SIMS conference*, Helsinki, Finland.

Washington, I., & Swartz, C. (2014). Design under uncertainty using parallel multiperiod dynamic optimization. *AIChE Journal*, *60*(9), 3151–3168.

Way, R. (1976). Methods for determination of composition and thermodynamic properties of combustion products for internal combustion engine calculations. *Proceedings of the Institution of Mechanical Engineers (London)*, *190*(60), 686–697.

Williams, F. (1985). *Combustion theory* (2nd ed.). The Benjamin Cummings Publishing Company, Inc.

Wu, D., & Wang, R. (2006). Combined cooling, heating and power: A review. *Progress in Energy and Combustion Science*, *32*(5–6), 459–495.

Würth, L., Hannemann, R., & Marquardt, W. (2011). A two-layer architecture for economically optimal process control and operation. *Journal of Process Control*, *21*(3), 311–321.

Yokell, S. (1990). *A working guide to shell and tube heat exchangers*. Mcgraw-Hill (Tx).

Yuan, Z., Chen, B., Sin, G., & Gani, R. (2012). State-of-the-art and progress in the optimization-based simultaneous design and control for chemical processes. *AIChE Journal*, *58*(6), 1640–1659.

Chapter 17
Energy Management of Smart Homes with Microgrid

Di Zhang, Songsong Liu and Lazaros G. Papageorgiou

Abstract Over a third of the world's primary energy is consumed by buildings, smart planning of energy supply to buildings is important to conserve energy and protect the environment. Most energy-consuming domestic tasks can be performed within a time period rather than at specific times. Energy cost or emissions could be reduced if these flexible tasks can be scheduled co-ordinately among multiple homes. This chapter addresses the problem of energy management of smart homes with microgrid, where the operation of distributed energy resources (DERs) and electricity-consumption household appliances are scheduled. A review of relevant literature works for smart homes with microgrid is presented. Then an optimisation-based framework is proposed to describe the related energy management problems of smart homes with microgrid. A mixed integer linear programming (MILP) model for three different objectives is developed: total cost minimisation, fair cost distribution, and cost versus CO_2 emissions. The application of this model is illustrated through an illustrative example of a smart building. The modelling approach developed in this work and the results obtained suggest that optimisation-based energy management of smart homes with microgrid results in cost saving and CO_2 emissions reduction. Moreover, the optimal operation schedules of the DERs, including thermal/electrical storage, are discussed.

17.1 Introduction

The current energy system is dominated by centralised generation, based on relatively few large plants using conventional fossil fuels and operating in central locations. Over 65 % of energy is lost along the electricity generation, transmission

D. Zhang · S. Liu · L.G. Papageorgiou (✉)
Department of Chemical Engineering, Centre for Process Systems Engineering, University College London, London WC1E 7JE, UK
e-mail: l.papageorgiou@ucl.ac.uk

S. Liu
School of Management, Swansea University, Bay Campus, Fabian Way, Skewen, Swansea SA1 8EN, UK

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_17

and distribution (Evangelisti et al. 2015), while in the UK 7 % of generation is lost in the transmission and distribution network (Basu et al. 2011). Microgrid equipped with intelligent elements from smart grid has been adopted to avoid these losses. Together with real-time pricing (RTP) scheme, it encourages and enables consumers to participate actively scheduling their home appliances to save energy, reduce cost and help environmental sustainability (Hu et al. 2010).

17.1.1 Microgrid and Smart Homes

As an alternative to the current energy providing system, microgrid is emerging to provide energy locally by utilising distributed energy resources (DER). The DERs include energy generation system, energy storage, load management options and energy generation units which are commonly located near the end users (Mallikarjun and Lewis 2014). The U.S. Department of Energy defines a microgrid as "a group of interconnected loads and distributed energy resources within clearly defined electrical boundaries that acts as a single controllable entity with respect to the grid and can connect and disconnect from the grid to enable it to operate in both grid-connected or island mode" (U.S. Department of Energy 2012). Microgrids have attracted increasing public attention in recent years since they can integrate renewable and distributed energy resources and energy storage systems, which provide useful control services during peak-demand hours or an electric outage (Zakariazadeh et al. 2014, Fares and Webber 2015). The main goals of a microgrid are to provide reliable and affordable energy to customers, reduce greenhouse gas emissions and relieve the burdens on transmission lines and save energy (Ciabattoni et al. 2015). Microgrids with intelligent elements of smart grid enable the widespread of DERs and demand response programs in distribution systems (Mitra and Suryanarayanan 2010).

Energy management of residential buildings is important to conserve energy and protect environment, since the domestic sector represents a significant proportion of energy use. The UK residential energy consumes 30 % of final energy and 35 % of all electricity in 2006 (Quiggin et al. 2012). Two-way information flow between energy markets and customers is available in a smart microgrid, where domestic appliances data can be collected by smart meters and sensors (Vega et al. 2015) and customer can play a more active and intelligent role in energy consumption to reduce energy cost and reshape load profile (Safamehr and Rahimi-Kian 2015). Domestic energy consumption depends on the dwelling physical properties, appliances' efficiencies and customers' behaviour. Energy consumption can be reduced by 10–30 % from changing living behaviours through a demand-side management (DSM) approach, which aims at matching generation values with demand by controlling the operation of appliances from the customer side (Tascikaraoglu et al. 2014). The most present manifestation of domestic DSM is the smart control of domestic appliances, such as washing machine, and this DSM approach is often related to the term 'smart home' (Müller et al. 2015). Smart homes are considered as a promising solution based on the common benefits to

customers, including lower energy costs, provision of comfort, security and home-based health care and assistance to elderly or disabled users (Pedrasa et al. 2010). Energy management in smart homes can be achieved by scheduling energy resource or domestic appliances.

17.1.2 Optimal Resource Scheduling in Microgrids

DER operation scheduling in microgrids have attracted much attention for energy savings. The operation costs are minimised by scheduling energy resources in the daily equipment operation planning based on a given energy demand profile in some recent literature. Logenthiran et al. (2011) present a multi-agent system for energy resource scheduling of power system with DERs for several microgrids with lumped loads. It schedules each microgrid individually first and then the microgrids are rescheduled to satisfy the total demand. A dynamic model is proposed for the energy management of a household through a Model Predictive Control (MPC) by Dagdougui et al. (2012), it decides the operational management of DERs for a green building according to time varying demands and resource availability. A mixed integer linear programming (MILP) model is developed for operation scheduling in a microgrid with diverse energy sources (Naraharisetti et al. 2011). The objective is to maximise the profit by maintaining diversity in the electricity production. Mohamed and Koivo (2012) propose a Genetic Algorithm (GA) approach to determine the optimal operating strategy of a microgrid for residential application, where the total operating cost is minimised while satisfying the load demand and environmental requirement.

17.1.3 Task Scheduling of Smart Homes with Economic Objective

Energy management involving energy tasks scheduling has also been studied besides the energy resources scheduling mentioned above. A robust optimisation method is applied to schedule household load of smart homes with PV power output uncertainty by Wang et al. (2015), where the daily electricity payment is minimised. An Electrical Demand-Side Management system is implemented to schedule daily deferrable and non-deferrable tasks in a real house with a PV generation system as a DER (Castillo-Cagigal et al. 2011). Kriett and Salani (2012) propose a generic MILP model to minimise the operating cost of both electrical and thermal supply and demand of a residential microgrid by scheduling the domestic appliances within the feasible time window. RTP encourage customers adjust their energy consumption, so a demand response management application with RTP is proposed by Chen et al. (2012) to determine the optimal operation of the residential appliances of a single house in the next 5 min time interval while considering future

electricity price uncertainties. It compares the stochastic optimisation and robust optimisation approaches for the scheduling of the tasks. By considering time of use (TOU) electricity tariff, an optimal and automatic residential load commitment framework is presented by Rastegar et al. (2012), it minimises household payment by determining on/off status of flexible appliances, operation of battery storage and plug-in hybrid electric vehicles. Another appliance scheduling work for a single house with TOU tariff is proposed by Adika and Wang (2014). The electrical appliances are clustered based on their time of use probabilities and then the aggregate loads of the appliances with similar schedules are tracked for different time periods under certain power limits. Baraka et al. (2013) design and implement a remotely controlled and energy efficient smart home using Android tablet as the user interface. The domestic tasks are assigned with priority numbers and then scheduled based on the overall cost limit and power usage limit in any time slot. Derin and Ferrante (2010) also develop a model addressing domestic energy consumption tasks scheduling, including the operation time of electric vehicle batteries, a dishwasher and a washing machine. However, for only those three tasks in a time span of 7 h, the exhaustive search takes 35 min which is relative slow. The computation time has been reduced to be within seconds by Zhang et al. (2013), where multiple smart homes with a common microgrid and 12 domestic tasks of each home are available to be scheduled. Both DER operation and electricity-consumption household appliances are scheduled to minimise the total energy cost based on RTP and domestic appliances operation time window. Fair cost distribution among smart homes is proposed by Zhang et al. (2014), the model is based on the lexicographic minimax method using an MILP approach. Electrical bill of each customer is minimised by scheduling domestic appliances using a game theoretical framework in the work of Barbato et al. (2015), where the peak demand of the customers are reduced.

17.1.4 Task Scheduling of Smart Homes with Multiple Objectives

In the recent literature, different papers can also be found about the scheduling work in smart homes based on various objectives other than cost minimisation or energy savings, such as load profiling, customer lifestyle and emissions (Beaudin and Zareipour 2015). Real-time scheduling of domestic appliances is addressed by Caprino et al. (2014) to limit the peak load of power usage. The appliance loads are classified into time-triggered and event-triggered loads and the physical model of these loads are considered in the model, such as a refrigerator and a washing machine. Amato et al. (2015) investigate on maximizing the self-consumption within a neighbourhood by planning the time when appliances will be switched on or off, while using the PV energy as much as possible. A smart home energy management model is proposed for scheduling domestic electrical and thermal

appliances (Shirazi et al. 2015), where discomfort index is proposed as the deviation from the most desired temperature and load shifting from preferred time window although the objective is still to minimise the total energy cost. A comfortable lifestyle is defined within both electrical and thermal zones for the residential energy management system by Anvari-Moghaddam et al. (2015). The two competitive objectives, energy cost and user's comfort level, are optimised in a mixed integer non-linear programing (MINLP) model by scaling the two goals into a single objective function by multiplying each objective with a relative weight. In EU, 25 % of emissions from energy use are relevant to households (Barisa et al. 2015), CO_2 emissions reduction is another important objective for domestic energy management. Within the environmental context, demand side management of a domestic dishwasher is investigated by Finn et al. (2013) according to renewable energy generation and pricing signals. Three optimisation objectives are examined separately: cost minimisation, demand on wind generation maximisation and associated carbon emissions minimisation. White good cycles and battery electric vehicles are scheduled by introducing demand response on system operation which impacts the system costs, reliability and emissions (Dupont et al. 2014). A dynamic optimisation model is presented by Bracco et al. (2015) to minimise operating costs and CO_2 emissions costs of the microgrid, while the CO_2 emissions are only minimised by being considered in a cost term due to emissions within the overall objective (total cost). A household behaviour simulation model is developed to investigate the joint influence of price and CO_2 signals in a demand response programme using a weighted sum approach (Song et al. 2014). Environmental and economic reasons are both considered in the work of Cheong et al. (2013), where the optimal household appliances scheduling is proposed for one home. An MILP model is proposed by Zhang et al. (2016) to schedule DER operations and the energy consumption of smart homes within a common microgrid under three price schemes. The domestic appliances are scheduled by coupling environmental and economic sustainability in a multi-objective optimisation with the ε-constraint method.

The major goal of this chapter is to propose an optimisation-based decision-support framework for the optimal energy management of smart homes with microgrid. Three optimisation models are developed to schedule DER operations and the energy consumption of smart home within a common microgrid, for total cost minimisation, fair cost distribution, and sustainability, respectively. The daily power consumption tasks are scheduled in this work based on daily electricity tariff, CO_2 intensity and given task operation time window. Lexicographic minimax method is applied for the fair cost distribution of the smart homes with microgrid. In addition, ε-constraint method is used for the sustainability model, by coupling environmental and economic sustainability.

The remainder of this chapter is organised as follows. The problem is described briefly first with relevant assumptions. Next, the mathematical programming models are provided. The proposed model is implemented on an illustrative example with 10 homes. Results for the example are presented and summary is drawn about this work.

17.2 Problem Description

In this work, multiple homes in a smart building are considered as shown in Fig. 17.1. A microgrid is available to provide local energy to the smart homes. The smart homes share common DERs of the microgrid, such as CHP generator, boiler, thermal or electrical storage. Electricity can be obtained from grid all the time when there is no sufficient local electricity generated. Each smart home has its own energy (heat and electricity) demands, which depend on the household types, available electrical appliances and living habits. Heat demand for each home is provided according to the types of household. While the electricity demand of each home depends on its own daily flexible and non-flexible domestic appliance tasks. Typical flexible tasks include dishwasher, washing machine and spin dryer. Thus, the electricity demand profile depends on the operation time of domestic electrical appliances. It assumes that local controllers are located for each DER and communication system is available to distribute the energy consumption scheme in the smart building. Electricity tariff and emission intensity are forecasted and given one day in advance. The energy cost for each smart home is calculated based on their respective energy consumption rate. Since DERs cannot fulfil demands for all smart

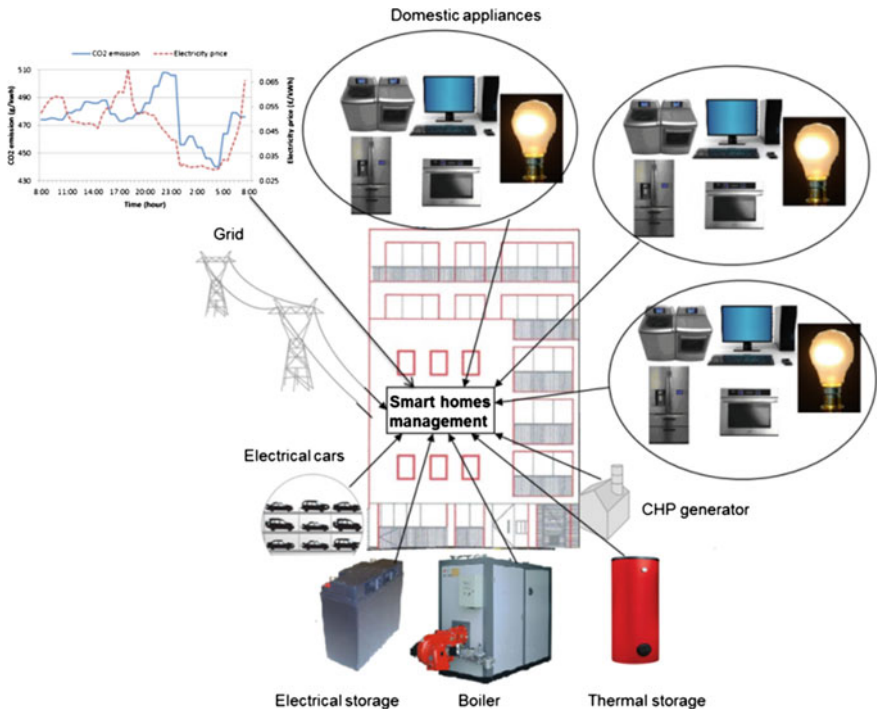

Fig. 17.1 Example of smart building

homes all the time, the smart homes need to compete with each other for the energy generated from DERs for minimum energy cost. On the other hand, electricity tariff and emission intensity vary over time, the electrical tasks tend to be operated in low tariff/emission intensity periods subject to the given task operation time window.

The overall problem can be stated as follows:

Given are (a) a time horizon split into a number of equal intervals, (b) heat demand of each smart home, (c) equipment capacities, (d) efficiencies of technologies, (e) maintenance cost of all equipment, (f) heat-to-power ratio of CHP generator, (g) charge and discharge limit rates for thermal/electrical storage, (h) gas price, real-time electricity prices from grid, (i) earliest starting and latest finishing times, (j) task capacity profiles, (m) task duration, (n) CO_2 emission intensity,

Determine (a) energy production plan, (b) equipment output sharing plan, (c) task starting time, (d) thermal/electrical storage plan, (e) electricity bought from grid, *so as to* (a) minimise daily total cost, (b) distribute optimal cost fairly among multiple smart homes, (c) obtain optimal solution considering both economic and environmental aspects.

17.3 Mathematical Formulation

In this section, a mixed integer linear programming model is presented, for the energy management of smart homes with microgrids, under three objectives: cost minimisation, fair cost distribution and cost versus CO_2 emissions. Firstly, the notation of the symbols used is given below, the superscript is used to indicate equipment and the subscript is used for indices:

Indices

i task
j smart home in the smart building
t time interval
θ task operation period

Parameters

b_t electricity buying price from grid at time t (£/kWh$_e$)
$C_{i\theta}$ power consumption capacity of task i at operation period θ (kW$_e$)
C^{CHP} CHP generator capacity (kW$_e$)
C^B boiler capacity (kW$_{th}$)
C^E electrical storage capacity (kWh$_e$)
C^T thermal storage capacity (kWh$_{th}$)
D^E electrical storage discharge limit (kW$_e$)
D^T thermal storage discharge limit (kW$_{th}$)
G^E electrical storage charge limit (kW$_e$)
G^T thermal storage charge limit (kW$_{th}$)
H_{jt} heat demand of smart home j at time t (kW$_{th}$)
r price of natural gas (£/kWh)

P_{ji} processing time of task i of home j
q electricity selling price to grid (£/kWh$_e$)
T^F_{ji} latest finishing time of task i of home j
T^S_{ji} earliest starting time of task i of home j
α CHP heat-to-power ratio
δ time interval duration (h)
μ^E cost per unit input (maintenance) for electrical storage unit (£/kWh$_e$)
μ^T cost per unit input (maintenance) for thermal storage unit (£/kWh$_{th}$)
η^B boiler efficiency
η^{CHP} CHP generator electrical efficiency
η^E electrical storage charge/discharge efficiency
η^T thermal storage charge/discharge efficiency
ξ^B CO_2 intensity of boiler thermal output (kg CO_2/kWh$_{th}$)
ξ^{CHP} CO_2 intensity of CHP electrical output (kg CO_2/kWh$_e$)
ξ^G_t CO_2 intensity of grid electricity at time t (kg CO_2/kWh$_e$)

Continuous Variables

f_{jt} thermal discharge rate to smart home j at time t (kW$_{th}$)
g_{jt} thermal charge rate from smart home j at time t (kW$_{th}$)
I_{jt} electricity imported from the grid to smart home j at time t (kW$_e$)
S^{IE} initial state of electrical storage (kWh$_e$)
S^{IT} initial state of thermal storage (kWh$_{th}$)
S^{ET}_t total electricity in electrical storage at time t (kWh$_e$)
S^E_{jt} electricity stored for smart home j at time t (kWh$_e$)
S^{TT}_t total heat stored at time t (kWh$_{th}$)
S^T_{jt} heat in thermal storage of smart home j at time t (kWh$_{th}$)
u_{jt} electricity output from CHP generator for smart home j at time t (kW$_e$)
x_{jt} heat output from boiler for smart home j at time t (kW$_{th}$)
y_{jt} electrical discharge rate to smart home j at time t (kW$_e$)
z_{jt} electrical charge rate from smart home j at time t (kW$_e$)
φ_j daily electricity cost of home j (£)
Φ_1 objective function, total cost
Φ_2 objective function, total CO_2 emission

Binary Variables

E_{jit} 1 if task i of home j starts at time t, 0 otherwise
X^G_{jt} 1 if electricity is bought from grid by home j at time t, 0 otherwise
X^E_{jt} 1 if electrical storage is charged by home j at time t, 0 otherwise
X^T_{jt} 1 if thermal storage is charged by home j at time t, 0 otherwise

17.3.1 Constraints

Firstly, the constraints are presented here.

17.3.1.1 Capacity Constraints

For the equipment included, CHP generator, boiler, electrical storage and thermal storage, the output utilised by each home is summed which should be limited by their designed capacities,
CHP generator:

$$\sum_j u_{jt} \leq C^{CHP} \quad \forall t \tag{17.1}$$

Boiler:

$$\sum_j x_{jt} \leq C^B \quad \forall t \tag{17.2}$$

Electrical storage:

$$\sum_j S^E_{jt} \leq C^E \quad \forall t \tag{17.3}$$

Thermal storage:

$$\sum_j S^T_{jt} \leq C^T \quad \forall t \tag{17.4}$$

17.3.1.2 Energy Storage Constraints

A central electrical storage is available for all the homes in the building. Each home can send or receive electricity from it, while the charging or discharging amount from battery for each home is recorded. No electricity can be consumed from the battery unless electricity has been charged before by that home. Each home is assumed to have its own sub-battery, but the capacity for them is flexible and the total capacity of the battery for the whole building is defined and provided. For each sub-battery j, electricity stored in the electrical storage at time t is equal to the amount stored at $t-1$ plus the electricity charged minus the electricity discharged. Electricity lost during the charging and discharging processes is counted here, for example if amount of electricity δz_{jt} is sent to the electrical storage, net amount

$\delta\eta^E z_{jt}$ will be charged, and the rest being lost, where η^E is turn-around efficiency of electrical storage. Meanwhile, during the discharging process, in order to discharge δy_{jt} of electricity to home consumption or grid, $\delta y_{jt}/\eta^E$ of electricity is needed.

$$S_{jt}^E = S_{j,t-1}^E + \delta\eta^E z_{jt} - \delta y_{jt}/\eta^E \quad \forall j,t \tag{17.5}$$

The discharged amount should be within the storage amount from the previous time interval.

$$S_{j,t-1}^E \geq \delta y_{jt}/\eta^E \quad \forall j,t \tag{17.6}$$

For each home, charging and discharging of electricity at the same time is not allowed:

$$z_{jt} \leq MX_{jt}^E \quad \forall j,t \tag{17.7}$$

$$y_{jt} \leq M(1-X_{jt}^E) \quad \forall j,t \tag{17.8}$$

At each time interval, the total electrical storage is the sum of amount over all sub-batteries in the building.

$$S_t^{ET} = \sum_j S_{jt}^E \quad \forall t \tag{17.9}$$

Initial storage state is set for the electrical storage at the beginning of each sample day. In order to avoid net accumulation, the electrical storage must return to its initial value at the end of the day. The initial storage state value is decided through optimising the model for one day utilisation. Otherwise, the initial state can be remained from the previous day. At the end of the day, the electrical storage must return to be over minimum limit to protect the equipment.

$$S_0^{ET} = S_T^{ET} = S^{IE} \tag{17.10}$$

The rates of discharge or charge of electricity should be bounded by the electrical storage discharge and charge limits defined by the battery manufacturer, so as to prevent excessive discharge/charge rates that would damage the battery or reduce its capacity:

$$\sum_j y_{jt} \leq D^E \quad \forall t \tag{17.11}$$

$$\sum_j z_{jt} \leq C^E \quad \forall t \tag{17.12}$$

Similarly, the smart building has a central thermal storage which can be considered as the sum of the sub-thermal storages from each home. Heat stored in the thermal storage at time t is equal to the amount stored at $t-1$ plus the heat amount charged minus the heat discharged. The heat loss during the heat storage process is represented in the similar way as presented for the electrical storage. At each time interval, the total thermal storage sums up all sub-thermal storage in the building.

$$S_{jt}^{T} = S_{j,t-1}^{T} + \delta\eta^{T} g_{jt} - \delta f_{jt}/\eta^{T} \quad \forall j,t \tag{17.13}$$

The discharged amount cannot exceed the storage amount obtained from the previous time interval.

$$S_{j,t-1}^{T} \geq \delta f_{jt}/\eta^{T} \quad \forall j,t \tag{17.14}$$

At each time interval, the total thermal storage is the sum of the storage amount over all sub-thermal storage in the building.

$$S_{t}^{TT} = \sum_{j} S_{jt}^{T} \quad \forall t \tag{17.15}$$

Stored heat must return to the initial state by the end of the day to avoid heat accumulation over one day. The initial storage state value is also attained through the model optimisation.

$$S_{0}^{TT} = S_{T}^{TT} = S^{IT} \tag{17.16}$$

The rates of discharge and charge of heat cannot go beyond the thermal storage discharge and charge limits which are constrained by the type of storage medium, mass and latent heat of the material:

$$\sum_{j} f_{jt} \leq D^{T} \quad \forall t \tag{17.17}$$

$$\sum_{j} g_{jt} \leq G^{T} \quad \forall t \tag{17.18}$$

Charge and discharge of heat cannot happen simultaneously for each home:

$$g_{jt} \leq M X_{jt}^{T} \quad \forall j,t \tag{17.19}$$

$$f_{jt} \leq M(1 - X_{jt}^{T}) \quad \forall j,t \tag{17.20}$$

17.3.1.3 Energy Balances

Electricity is supplied by the CHP generator, electricity received from the electrical storage and grid, while charged to the electrical storage. The power consumption capacity of some tasks fluctuates over the operation time periods, e.g. washing machine has different capacity profiles over washing and spinning processes. So the electricity consumption of each home at any time interval is calculated by summing up all the capacities of different tasks during task operation period θ.

$$\sum_{i} \sum_{\theta}^{P_{ji}-1} C_{i\theta} E_{j,i,t-\theta} = u_{jt} + y_{jt} + I_{jt} - z_{jt} \quad \forall j, t \tag{17.21}$$

For each home, buying and selling of electricity from/to the grid at the same time is forbidden:

$$I_{jt} \leq M X_{jt}^{G} \quad \forall j, t \tag{17.22}$$

$$R_{jt} \leq M(1 - X_{jt}^{G}) \quad \forall j, t \tag{17.23}$$

For each time interval, heat demand is fulfilled by the CHP generator, boiler and heat discharged from the thermal storage, minus heat charged to the thermal storage.

$$H_{jt} = \alpha u_{jt} + x_{jt} + f_{jt} - g_{jt} \quad \forall j, t \tag{17.24}$$

17.3.1.4 Starting Time and Finishing Time

The operation time of each task must be bounded by its given time window. Each task can neither start earlier than the starting time of each task, nor finish after the latest finishing time. For each task from each home, it has to start only once.

$$\sum_{t=T_{ji}^{S}}^{T_{ji}^{F}-P_{ji}} E_{jit} = 1 \quad \forall j, i \tag{17.25}$$

17.3.1.5 Daily Cost

For each home, the total daily electricity cost consists of: the operation and maintenance cost of the CHP generator, boiler, electrical storage and thermal storage; the cost of electricity purchased from the grid; the revenue from electricity sold to the grid.

$$\begin{aligned}\phi_j = {} & \delta \sum_t ru_{jt}/\alpha \quad \text{CHP operation cost} \\ & + \delta \sum_t rx_{jt}/\eta^B \quad \text{boiler operation cost} \\ & + \delta \sum_t \mu^E y_{jt} \quad \text{electrical storage maintenance cost } \forall j \\ & + \delta \sum_t \mu^T f_{jt} \quad \text{thermal storage maintenance cost} \\ & + \delta \sum_t b_t I_{jt} \quad \text{electricity buying cost from grid}\end{aligned} \tag{17.26}$$

17.3.2 Objective Functions

Next, the three objectives of the proposed optimisation model are presented respectively, for total cost minimisation, fair cost distribution, and cost versus CO_2 emissions.

17.3.2.1 Total Cost Minimisation

As the first objective, the total cost of the all smart homes is minimised. Thus, the objective function is the sum of each home's cost.

$$\Phi_1 = \sum_j \phi_j \tag{17.27}$$

The proposed model for the total cost minimisation includes Eqs. (17.1)–(17.26) as constraints and Eq. (17.27) as the objective function.

17.3.2.2 Fair Cost Distribution

For the fair cost distribution objective, each home in the smart building targets in minimising the total cost subject to fair cost distribution among homes. The lexicographic minimax approach is applied for this problem.

A lexicographic minimax problem is defined as follows:

$$\underset{x \in X}{lex} \min\left\{\Theta(\phi_j(x))\right\} \tag{17.28}$$

where $\phi_j(x)$ is vector of the objective value under fairness scenario of each home, which is defined as the relative cost based on the pre-determined cost boundaries from each home, maximum and minimum energy cost ($\phi_j^{\max}$ and $\phi_j^{\min}$),

$\phi_j = \frac{\phi_j - \phi_j^{\min}}{\phi_j^{\max} - \phi_j^{\min}}$, $x \in X \subset R^n$ is a n-dimensional vector of decision variables, X is the decision space defined by Eqs. (17.1)–(17.26) and $\Theta: R^J \to R^J$ maps orders of the component of vectors in a non-decreasing order. With a given vector $e = (e_1 \ldots e_J)$, $\Theta(e) = (\theta_1(e), \ldots, \theta_J(e))$, where $\theta_j(e) \in \{e_1, \ldots, e_J\}$ is the jth component in vector $\Theta(e)$ and $\theta_1(e) \geq \ldots \geq \theta_J(e)$. Then in the lexicographic minimax problem, the objective values are minimised in the non-increasing order of the objective values, which means the highest objective value is minimised first, then the second and so on. The optimal solution of the lexicographic minimax problem is the Pareto-optimum (Safamehr and Rahimi-kian 2015), and can be achieved by solving the following lexicographic minimisation problem (Erkut et al. 2008):

$$\begin{aligned} &\text{Lexmin}\left\{\lambda_1 + \sum_{j=1}^{J} d_{1j}, \lambda_2 + \frac{1}{2}\sum_{j=1}^{J} d_{2j}, \ldots, \lambda_J + \frac{1}{J}\sum_{j=1}^{J} d_{Jj}\right\} \\ &\text{s.t. } \lambda_{n'} + d_{n'j} \geq \frac{\phi_j - \phi_j^{\min}}{\phi_j^{\max} - \phi_j^{\min}} \quad \forall j, n' = 1, \ldots, n \\ &\quad d_{n'j} \geq 0 \quad \forall j, n' = 1, \ldots, n \\ &\quad Eqs.(17.1) - (17.26) \end{aligned} \tag{17.29}$$

17.3.2.3 Cost Versus CO_2 Emissions

Both total cost and total CO_2 emissions are considered as the objective functions in the third objective. The total cost is defined as Eq. (17.27), while the total CO_2 emission, including the CO_2 emissions from the use of CHP generator and boiler, and from the conventional electricity grid, is formulated as follows:

$$\Phi_2 = \sum_j \sum_t \left[\delta\left(\xi^{CHP} u_{jt} + \xi_t^G I_{jt} + \xi^B x_{jt}\right)\right] \tag{17.30}$$

Thus, the proposed multi-objective optimisation formulation can be presented as follows:

$$\text{Min}\{\Phi_1, \Phi_2\} \tag{17.31}$$

s.t. Eqs. (17.1)–(17.27), and (17.30).

To solve the above multi-objective optimisation model, the classic ε-constraint method is applied to transform the multi-objective optimisation to single-objective optimisation. Here, Φ_1 is kept as the objective function, while Φ_2 is limited by an upper bound, ε, in the constraints. By changing the value of ε, a set of Pareto optimal solutions can be achieved.

17.4 Illustrative Example

The proposed model for energy consumption scheduling under three objectives, including cost minimisation, fair cost distribution and cost versus CO_2 emissions, are applied for a smart building with 10 smart homes. A sample summer day is considered in this work, where the domestic tasks are the same while the heat demands are different. The common DERs are shared by all homes, whose capacities are obtained from the energy profiles while the technical parameters and costs are taken from Hawkes and Leach (2009):

- One CHP generator with a capacity of 4 kW_e and electrical efficiency of 40 %. Heat to power ratio is assumed to be 1.2, and natural gas cost of 2.7 p/kWh;
- One boiler with capacity of 30 kW_{th} and efficiency of 85 %, and natural gas cost of 2.7 p/kWh;
- One electrical storage unit with a capacity of 4 kWeh, charge/discharge efficiency of 95 %, both discharge limit and charge limit of 4 kW_e, and the maintenance cost of 0.5 p/kWh_e;
- One thermal storage unit with a capacity of 6 kW_{th} h; charge/discharge efficiency of 98 %, both discharge limit and charge limit of 6 kW_{th}, and the maintenance cost of 0.1 p/kWh_{th};
- A grid connection (allowing import of electricity when operating parallel to the conventional grid).

The heat demand is generated from the Community's Heating Demand Profile Generator which is developed by the University of Strathclyde (2007). The 10 smart homes are assumed to be selected from a flat building built during 1998 −2002 and they are classified into 3 types of occupancy as presented in Table 17.1. Smart home 4, 5, 6 are top/ground flats while other homes are mid flats. For a summer day, the heat demands of the 10 smart homes are presented in Fig. 17.2.

There are 12 electrical appliances available to be scheduled in smart homes, which are shown in Table 17.2. All tasks, except the dishwasher and washing machine, have constant power consumption rates during operation, while the electrical profiles for dish washer and washing machine are illustrated in Fig. 17.3. and Table 17.3.

Table 17.1 Household occupancy types (University of Strathclyde, 2007)

	Type 1	Type 2	Type 3
Household type	A household that at least one member has a part time job during the morning session	A household that all members are working on a full time scheme	A household that there is one or more pensioners, disabled persons or unemployed
Unoccupied period	9:00−13:00	9:00−18:00	N/A
Smart homes	1, 4, 7, 10	2, 5, 8	3, 6, 9

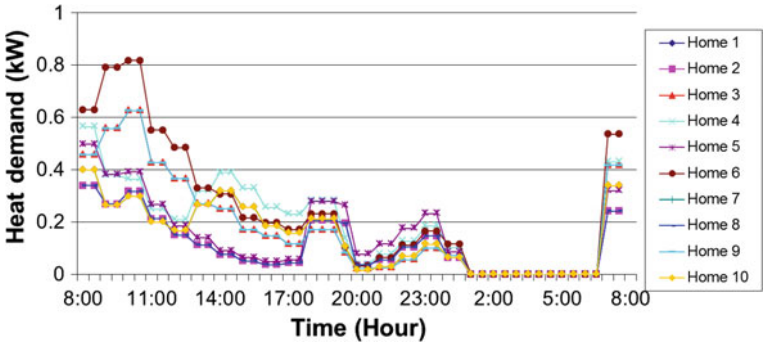

Fig. 17.2 Heat demands of 10 smart homes for a summer day

Table 17.2 Electricity consumption for tasks (Electropaedia 2016)

	Task	Power (kW)	Duration (hour)
1	Dishwasher	–	2
2	Washing machine	–	1.5
3	Spin dryer	2.5	1
4	Cooker hob	3	0.5
5	Cooker oven	5	0.5
6	Microwave	1.7	0.5
7	Interior lighting	0.84	6
8	Laptop	0.1	2
9	Desktop	0.3	3
10	Vacuum cleaner	1.2	0.5
11	Fridge	0.3	24
12	Electrical car	3.5	3

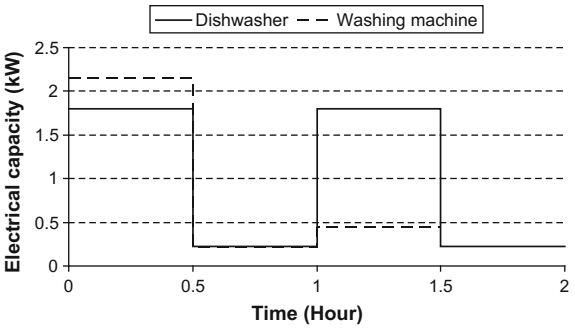

Fig. 17.3 Electricity utilisation profiles of dishwasher and washing machine

The earliest starting time of each task from each smart home is given in Table 17.4 and the time window length is presented in Table 17.5. The electricity demands of 10 smart homes when all the tasks start at the earliest starting time are

Table 17.3 Electrical task of each smart home (Zhang et al. 2014)

Home	Tasks
1	1–12
2	1–6
3	7–12
4	1–8
5	4–12
6	1–12
7	1–4, 9–12
8	1–4, 6–10
9	1–12
10	5–8, 10–12

Table 17.4 Electrical task earliest starting time in hour

	Home	1	2	3	4	5	6	7	8	9	10
1	Dishwasher	12	11	–	13	–	18	14	16	11	–
2	Washing machine	16	14	–	11	–	22	22	20	16	–
3	Spin dryer	19	17	–	14	–	25	25	23	19	–
4	Cooker hob	15	10	–	13	10	14	18	11	10	–
5	Cooker oven	11	15	–	20	13	13	–	–	19	20
6	Microwave	21	13	–	20	12	17	–	18	20	10
7	Interior lighting	18	–	20	20	22	19	–	17	20	21
8	Laptop	19	–	17	17	19	21	–	18	19	19
9	Desktop	17	–	16	–	14	19	20	22	20	–
10	Vacuum cleaner	18	–	19	–	20	16	22	21	21	21
11	Fridge	1	–	1	–	1	1	1	–	1	1
12	Electrical car	21	–	20	–	19	18	17	–	21	19

Table 17.5 Electrical task time window length in hour

	Home	1	2	3	4	5	6	7	8	9	10
1	Dishwasher	8	7.5	–	6.5	–	5.5	5	4.5	4	–
2	Washing machine	3	2.5	–	3	–	3	2.5	2.5	2.5	–
3	Spin dryer	5	4.5	–	3.5	–	2.5	2.5	2	1.5	–
4	Cooker hob	1	1.5	–	2.5	3	3.5	5.5	4.5	5	–
5	Cooker oven	1	1.5	–	2.5	3	3.5	–	4.5	5	5.5
6	Microwave	1	1.5	–	2.5	3	3	–	2	1.5	1
7	Interior lighting	6	–	6	6	6	6	–	6	6	6
8	Laptop	6	–	5.5	3.5	5	6	–	4.5	5	5.5
9	Desktop	6	–	4	–	5.5	6	5	3.5	4.5	–
10	Vacuum cleaner	8	–	4.5	–	5.5	6	6.5	7	7.5	8
11	Fridge	24	–	24	–	24	24	24	–	24	24
12	Electrical car	10	–	7	–	4	8	8.5	–	9.5	10

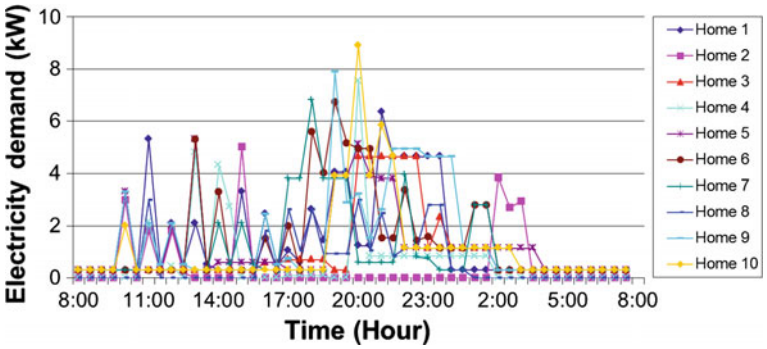

Fig. 17.4 Electricity demands of 10 smart homes under the earliest starting time

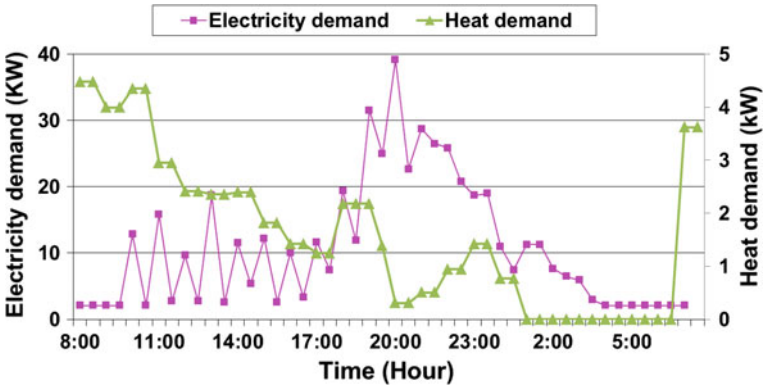

Fig. 17.5 Total energy demands of 10 smart homes in a summer day under earliest starting time

shown in Fig. 17.4 while the total energy demands of the 10 smart homes for a summer day are given in Fig. 17.5.

The electricity tariff and CO_2 intensity profiles in the UK on 17th August 2013 is applied for the illustrative example. As shown in Fig. 17.6, the two profiles have different peak hours resulting from electricity generation of different energy sources over the day or even the importation of the electricity from the international markets.

The CO_2 emission rates from the CHP and the boiler operation are given in Table 17.6, and they are assumed to be constant over all the considered time periods. The carbon footprint of the use phase for the system is assessed considering a functional unit of 1 kWh of electrical output and 1 kWh of thermal output, for the CHP and the boiler respectively. A boiler efficiency of 85 % is assumed in this study. For the CHP, it is assumed as a fuel cell unit with 10 kW capacity and 40 % electrical efficiency. The "natural gas supply" impact is referred to the extraction and distribution of the natural gas up to the system and it is country

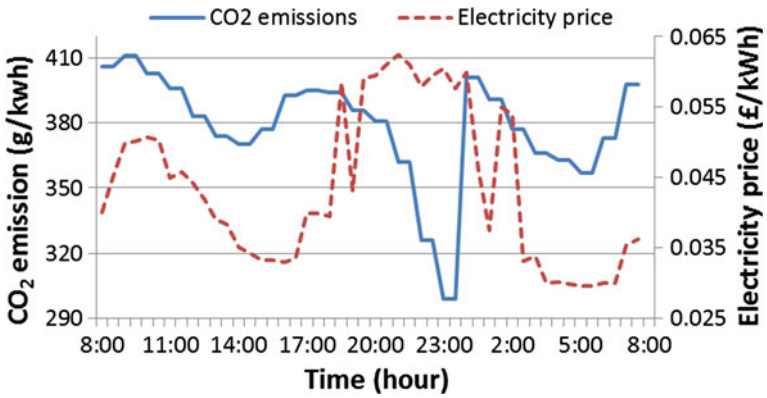

Fig. 17.6 Electricity tariff and CO_2 intensity of for 17th August 2013 (Balancing Mechanism Reporting System 2013, Earth Notes 2013)

Table 17.6 Carbon footprint for the CHP and the boiler

	Natural gas supply	Direct emissions	Total
CHP ($kgCO_2$ eq/kwh electrical output)	0.0396	0.5049	0.5445
Boiler ($kgCO_2$ eq/kwh thermal output)	0.0186	0.2923	0.3109

specific, while the "direct emissions" impact is referred to the specific use of the system. The carbon footprint is calculated by GaBi 6.0 sustainability software (PE International 2016).

17.5 Results and Discussion

The proposed model is implemented for the illustrative example under a sample summer day. The optimal results under the three objectives are presented in the following.

17.5.1 Total Cost Minimisation

By minimising the total cost in Eq. (17.27) subject to Eqs. (17.1)–(17.26), the obtained total daily energy consumption cost is £11.26. Under the earliest stating time scenario, the total cost is £13.97. There is about 19 % cost saving resulting from DERs utilisation and scheduling the operation of DERs and domestic appliances.

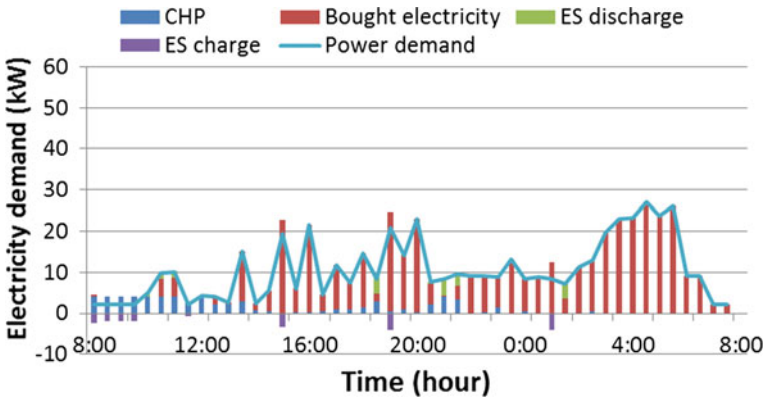

Fig. 17.7 Electricity balance of cost minimisation

Figure 17.7 shows the total electricity balances of the 10 homes for the sample summer day. CHP is providing constant maximum output 4 kW in the morning while there is no electricity output from the CHP during night time where no heat is needed. Electrical storage is applied several times, charged when tariff is low and discharged when it is high. Compared with the total electricity demand in Fig. 17.5, the peaks of the total power demand have been moved to the night time from evening. Meanwhile the peak demand has been reduced by 13−27 kW. As defined in the model, for each single home, electricity cannot be charged and discharged simultaneously. In this example, total electricity is charged in the early morning and discharged later in the evening. Figure 17.8 presents the heat balances of the 10 homes. CHP becomes the main heat provider for the smart homes, while the remaining minor heat demand is supplied by the boiler. Thermal storage is used quite frequently in this example. As seen from the figure, for some hours thermal

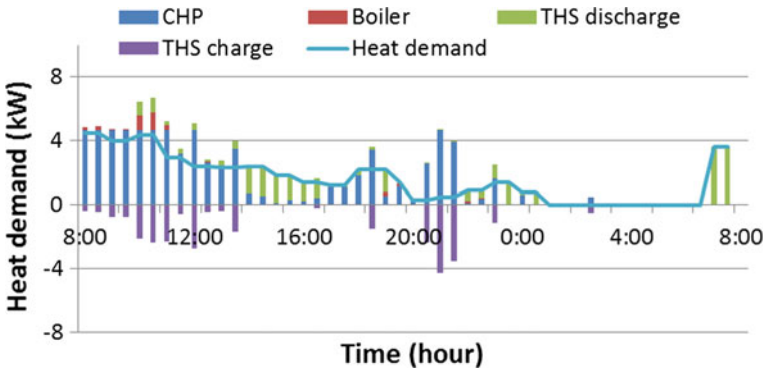

Fig. 17.8 Heat balance under fairness concern

storage charge and discharge happens at the same time. It is because the thermal storage works as bank system for heat deposit for the 10 homes. Ideally, only the amount of usage is counted in the system as depositing money in a bank and then cost is calculated based on the total usage over the time under concern. The thermal storage and electrical storage are shared in such a way because of the different energy demand patterns among homes.

17.5.2 Fair Cost Distribution with Lexicographic Minimax Method

If only total cost of the 10 homes is minimised, the cost of each home is presented as φ_j in Table 17.7. Since cost from respective home is not considered, the cost is distributed without fairness concern as shown in the table. The obtained cost from each home is compared under the proposed fairness concept proposed as $\hat{\varphi}_j(x)$. The upper bound of the cost of each smart home is obtained from the case when all heat is generated from boiler and electricity is solely bought from grid, while all tasks start at their earliest starting time. The cost lower bound of each smart home is achieved by minimising energy cost of each single smart home where microgrid is available. Values of both bounds are presented in the table. The $\hat{\varphi}_j$ values indicate the total cost is distributed unfairly among homes, as the normalised objective values range from 0.00 to 0.38.

In order to obtain the fair cost distribution among the 10 homes, the lexicographic minimax approach is applied. The fair solution is presented in the last two columns in Table 17.7. The total cost with fairness concern is £11.29, which is very

Table 17.7 Cost of each home from minimising total cost and fairness concern

			Total cost minimisation		Fair solution	
Home	$\phi_j^{\max}$(£)	$\phi_j^{\min}$(£)	ϕ_j(£)	$\hat{\phi}_j$	ϕ_j^*(£)	$\hat{\phi}_j^*$
1	1.89	1.42	1.48	0.13	1.51	0.18
2	0.53	0.43	0.43	0.00	0.45	0.18
3	1.45	0.99	1.06	0.14	1.07	0.18
4	0.93	0.79	0.84	0.38	0.82	0.23
5	1.57	1.17	1.25	0.2	1.25	0.19
6	1.91	1.47	1.53	0.13	1.55	0.18
7	1.33	1.13	1.19	0.28	1.18	0.23
8	0.86	0.65	0.72	0.33	0.70	0.23
9	1.96	1.45	1.54	0.17	1.54	0.18
10	1.54	1.14	1.22	0.21	1.22	0.21
Total	13.97	10.64	11.26	–	11.29	–

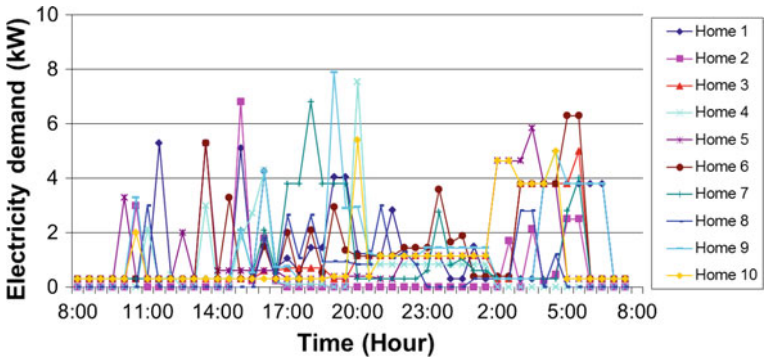

Fig. 17.9 Optimal electricity demands

close to the minimum total cost £11.26. The optimal objective values are close to 0.20. The costs are more fairly distributed according to the contribution from each home.

Figure 17.9 presents the optimal electricity demand of each home with fairness concern. Compared with Fig. 17.4, the electricity demands are shifted to the night time where the tariff is lower. The scheduling is based on the real-time electricity pricing and given task operation time window to minimise respective cost of each home with fairness concern.

17.5.3 Cost Versus CO_2 Emissions

When both economic and environmental aspects are considered, the model is applied with the objective in Eq. 17.31. The sample summer day profiles of electricity tariff and CO_2 emissions in Fig. 17.6 are used together with heat demands of the 10 smart homes shown in Fig. 17.2. Figure 17.10 presents the Pareto curves for cost and CO_2 emissions, where CO_2 emissions decrease while cost increases. The difference between the maximum and minimum values for cost and CO_2 emissions is 24 % and 6 % respectively. As shown, the curve presented is made of 21 points, the curve shows a steep decrease in CO_2 emissions over the first 11 points with values for the CO_2 emissions dropping from 101 to 98 kg, while the cost difference is less than £0.42. After point 12, the CO_2 emissions drop at a slower rate to 95 kg.

Figure 17.11 shows the electricity balances for the point with minimum CO_2 emissions in Fig. 17.10, where to minimise the total kg CO_2eq is the single objective, hence the electricity demand is scheduled based on minimising the carbon footprint of the electricity consumption of the smart homes. In this case, the CHP operates as much as possible in the morning because of its low CO_2 intensity and the remaining demand is satisfied by the electricity from the grid and the electrical storage discharge. The electricity from the grid is bought during the time

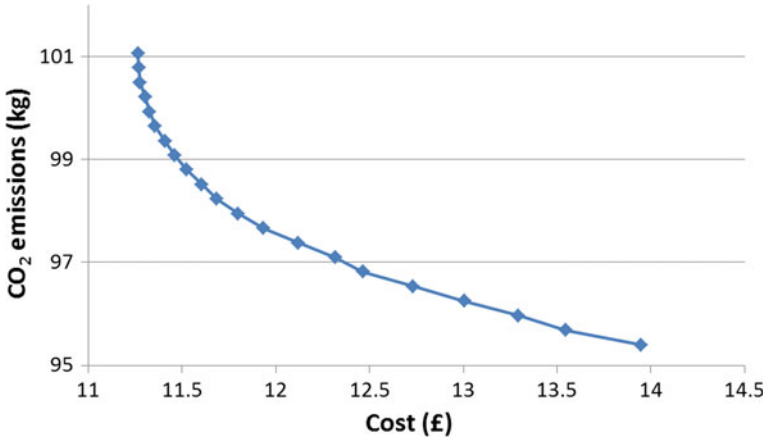

Fig. 17.10 Pareto curve for cost and carbon footprint

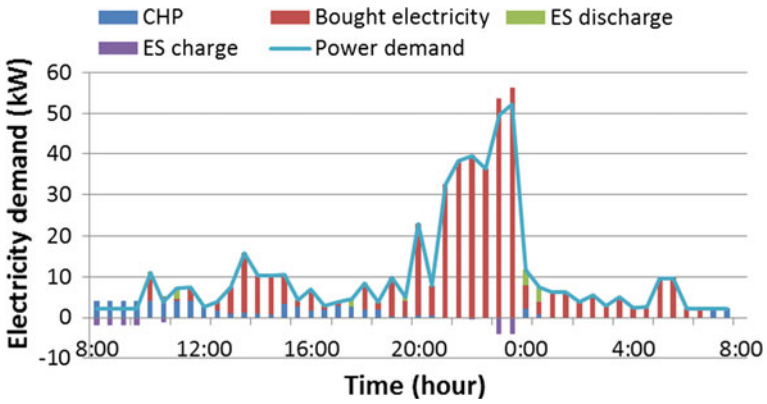

Fig. 17.11 Electricity balance for the minimum CO_2 emissions point

periods when the grid CO_2 intensity is low (see Fig. 17.6), mainly 20:00−00:00. Electricity is totally bought from the grid when there is no heat demand during the night. Electrical storage is charged to store the electricity with low emission, and discharged when the emission rate is high. The heat balances for the minimum CO_2 emissions point is given in Fig. 17.12, again the thermal storage is used quite often here. The thermal storage compensates the usage of CHP which has low CO_2 intensity. No heat is provided by the boiler because of the preference of CHP for the energy demand profiles of the sample summer day.

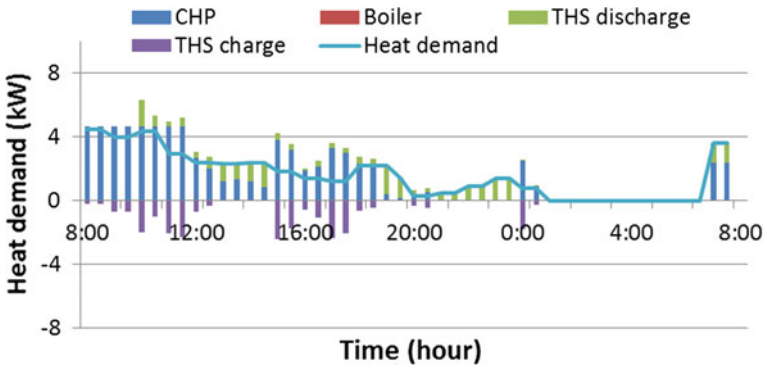

Fig. 17.12 Heat balance for the minimum CO_2 emissions point

17.6 Summary

Optimisation-based model for energy management of smart homes with microgrid provide valuable tools to schedule the operation of distributed energy resources (DERs) and electricity-consumption household appliances. Economic cost reduction and CO_2 emission reduction can be obtained through scheduling among multiple smart homes. This work has presented an MILP mathematical model under three objectives: total cost minimisation, fair cost distribution, and cost versus CO_2 emissions. Total cost minimisation is considered first, then the cost is distributed fairly among smart homes based on the lexicographic minimax method. The model with the last objective determines the optimal results with trade-off between economic cost and environmental emissions via ε-constraint method. The Pareto-optimal curve between cost and CO_2 emissions has been provided.

The energy management of smart homes depends heavily on the family living habits, on the other hand it could also affect people's behaviour and wider time window could result in more cost savings. Fair cost distribution is also influenced by the objective fairness definition. Different fairness criteria could be selected accordingly based on participants' preferences. Further more, priorities could be assigned to some participants for particular reasons, such as poverty or location. Different solutions for the scheduling of DER operation and electrical tasks can be obtained under various objectives, depending on the energy demand patterns from different seasons, the electricity tariff and CO_2 intensity profiles. More environmental impact factors can be added in the model besides CO_2 emissions, such as acidification potential (AP) and primary energy (PE).

Acknowledgments Authors gratefully acknowledge the financial supports from EPSRC Centre for Innovative Manufacturing in Emergent Macromolecular Therapies (to SL and LGP, EP/I033270/1), and Qatar National Research Fund (via GORD) under NPRP 6-588-2-243 for DZ.

References

Adika, C. O., & Wang, L. (2014). Smart charging and appliance scheduling approaches to demand side management. *International Journal of Electrical Power & Energy Systems, 57*, 232–240.

Amato, A., Di martino, B., Scialdone, M., & Venticinque, S. (2015). Design and evaluation of P2P overlays for energy negotiation in smart micro-grid. *Computer Standards & Interfaces*.

Anvari-Moghaddam, A., Monsef, H., & Rahimi-Kian, A. (2015). Cost-effective and comfort-aware residential energy management under different pricing schemes and weather conditions. *Energy and Buildings, 86*, 782–793.

Balancing Mechanism Reporting System (2016). The New Electricity Trading Arrangements. Retrieved January 07, 2016 from http://www.bmreports.com/bsp/SystemPricesHistoric.htm.

Baraka, K., Ghobril, M., Malek, S., Kanj, R. & Kayssi, A. (2013). Low Cost arduino/android-based energy-efficient home automation system with smart task scheduling. In: *2013 Fifth International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN),* 5–7 June 2013 (pp. 296−301).

Barbato, A., Capone, A., Chen, L., Martignon, F., & Paris, S. (2015). A distributed demand-side management framework for the smart grid. *Computer Communications, 57*, 13–24.

Barisa, A., Rosa, M., Laicane, I., & Sarmins, R. (2015). Application of low-carbon technologies for cutting household GHG emissions. *Energy Procedia, 72*, 230–237.

Basu, A. K., Chowdhury, S. P., Chowdhury, S., & Paul, S. (2011). Microgrids: Energy management by strategic deployment of DERs—A comprehensive survey. *Renewable and Sustainable Energy Reviews, 15*, 4348–4356.

Beaudin, M., & Zareipour, H. (2015). Home energy management systems: A review of modelling and complexity. *Renewable and Sustainable Energy Reviews, 45*, 318–335.

Bracco, S., Delfino, F., Pampararo, F., Robba, M., & Rossi, M. (2015). A dynamic optimization-based architecture for polygeneration microgrids with tri-generation, renewables, storage systems and electrical vehicles. *Energy Conversion and Management, 96*, 511–520.

Caprino, D., Della Vedova, M. L., & Facchinetti, T. (2014). Peak shaving through real-time scheduling of household appliances. *Energy and Buildings, 75*, 133–148.

Castillo-Cagigal, M., Gutiérrez, A., Monasterio-Huelin, F., Caamaño-Martín, E., Masa, D., & Jiménez-Leube, J. (2011). A semi-distributed electric demand-side management system with PV generation for self-consumption enhancement. *Energy Conversion and Management, 52*, 2659–2666.

Chen, Z., Wu, L., & Fu, Y. (2012). Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization. *IEEE Transactions on Smart Grid*, 1–9.

Ciabattoni, L., Comodi, G., Ferracuti, F., Fonti, A., Giantomassi, A., & Longhi, S. (2015). Multi-apartment residential microgrid monitoring system based on kernel canonical variate analysis. *Neurocomputing, 170*, 306–317.

Dagdougui, H., Minciardi, R., Ouammi, A., Robba, M., & Sacile, R. (2012). Modeling and optimization of a hybrid system for the energy supply of a "Green" building. *Energy Conversion and Management, 64*, 351–363.

Derin, O., & Ferrante, A. (2010). Scheduling energy consumption with local renewable micro-generation and dynamic electricity prices. In *Proceedings of the First Workshop on Green and Smart Embedded System Technology: Infrastructures, Methods and Tools, Stockholm, Sweden.* Stockholm.

Dupont, B., Dietrich, K., de Jonghe, C., Ramos, A., & Belmans, R. (2014). Impact of residential demand response on power system operation: A Belgian case study. *Applied Energy, 122*, 1–10.

Erkut, E., Karagiannidis, A., Perkoulidis, G., & Tjandra, S. A. (2008). A multicriteria facility location model for municipal solid waste management in North Greece. *European Journal of Operational Research, 187*, 1402–1421.

Earth Notes. (2016).A note on variations in UK/GB grid electricity CO2 intensity with time. Retrieved January 07, 2016 from http://www.earth.org.uk/note-on-UK-grid-CO2-intensity-variations.html.

Cheong, K. S., Kordel, M., Wu, J., Sandberg, H., & Johansson, K. H. (2013). Energy and CO_2 efficient scheduling of smart home appliances. In *Control Conference (ECC), 2013 European*, 17–19 July 2013 (4051–4058).

Electropaedia (2016). Domestic electrical energy usage. Retrieved January 07, 2016 from http://www.mpoweruk.com/electricity_demand.htm.

Evangelisti, S., Lettieri, P., Clift, R., & Borello, D. (2015). Distributed generation by energy from waste technology: A life cycle perspective. *Process Safety and Environmental Protection, 93*, 161–172.

Fares, R. L., & Webber, M. E. (2015). Combining a dynamic battery model with high-resolution smart grid data to assess microgrid islanding lifetime. *Applied Energy, 137*, 482–489.

Finn, P., O'Connell, M., & Fitzpatrick, C. (2013). Demand side management of a domestic dishwasher: Wind energy gains, financial savings and peak-time load reduction. *Applied Energy, 101*, 678–685.

Hawkes, A. D., & Leach, M. A. (2009). Modelling high level system design and unit commitment for a microgrid. *Applied Energy, 86*, 1253–1265.

Hu, W., Chen, Z., & Bak-Jensen, B. (2010). *Optimal operation strategy of battery energy storage system to real-time electricity price in Denmark*. IEEE: IEEE PES General Meeting.

Kriett, P. O., & Salani, M. (2012). Optimal control of a residential microgrid. *Energy, 42*, 321–330.

Logenthiran, T., Srinivasan, D., & Khambadkone, A. M. (2011). Multi-agent system for energy resource scheduling of integrated microgrids in a distributed system. *Electric Power Systems Research, 81*, 138–148.

Mallikarjun, S., & Lewis, H. F. (2014). Energy technology allocation for distributed energy resources: A strategic technology-policy framework. *Energy, 72*, 783–799.

Mitra, J., & Suryanarayanan, S. (2010). *System analytics for smart microgrids*. IEEE: IEEE PES General Meeting.

Mohamed, F. A., & Koivo, H. N. (2012). Online management genetic algorithms of microgrid for residential application. *Energy Conversion and Management, 64*, 562–568.

Müller, D., Monti, A., Stinner, S., Schlösser, T., Schütz, T., Matthes, P., et al. (2015). Demand side management for city districts. *Building and Environment, 91*, 283–293.

Naraharisetti, P. K., Karimi, I. A., Anand, A., & Lee, D.-Y. (2011). A linear diversity constraint—Application to scheduling in microgrids. *Energy, 36*, 4235–4243.

PE International. (2016). GaBi sustainability software. Retrieved January 07, 2016 from http://www.gabi-software.com/uk-ireland/software/gabi-software.

Pedrasa, M. A. A., Spooner, T. D., & Macgill, I. F. (2010). Coordinated Scheduling of residential distributed energy resources to optimize smart home energy services. *IEEE Transactions on Smart Grid, 1*, 134–143.

Quiggin, D., Cornell, S., Tierney, M., & Buswell, R. (2012). A simulation and optimisation study: Towards a decentralised microgrid, using real world fluctuation data. *Energy, 41*, 549–559.

Rastegar, M., Fotuhi-Firuzabad, M., & Aminifar, F. (2012). Load commitment in a smart home. *Applied Energy, 96*, 45–54.

Safamehr, H., & Rahimi-Kian, A. (2015). A cost-efficient and reliable energy management of a micro-grid using intelligent demand-response program. *Energy, 91*, 283–293.

Shirazi, E., Zakariazadeh, A., & Jadid, S. (2015). Optimal joint scheduling of electrical and thermal appliances in a smart home environment. *Energy Conversion and Management, 106*, 181–193.

Song, M., Alvehag, K., Widén, J., & Parisio, A. (2014). Estimating the impacts of demand response by simulating household behaviours under price and CO2 signals. *Electric Power Systems Research, 111*, 103–114.

Tascikaraoglu, A., Boynuegri, A. R., & Uzunoglu, M. (2014). A demand side management strategy based on forecasting of residential renewable sources: A smart home system in Turkey. *Energy and Buildings, 80*, 309–320.

University of Strathclyde. (2007). Demand Profile Generators. Retrieved January 07, 2016 from http://www.esru.strath.ac.uk/EandE/Web_sites/06-07/Carbon_neutral/tools%20folder/Demand%20Profile%20Generators.htm.

U.S. Department of Energy. (2012). Summary Report: 2012 DOE microgrid workshop. Retrieved January -07, 2016 from http://energy.gov/sites/prod/files/2012%20Microgrid%20Workshop%20Report%2009102012.pdf.

Vega, A. M., Santamaria, F., & Rivas, E. (2015). Modeling for home electric energy management: A review. *Renewable and Sustainable Energy Reviews, 52*, 948–959.

Wang, C., Zhou, Y., Jiao, B., Wang, Y., Liu, W., & Wang, D. (2015). Robust optimization for load scheduling of a smart home with photovoltaic system. *Energy Conversion and Management, 102*, 247–257.

Zakariazadeh, A., Jadid, S., & Siano, P. (2014). Smart microgrid energy and reserve scheduling with demand response using stochastic optimization. *International Journal of Electrical Power & Energy Systems, 63*, 523–533.

Zhang, D., Shah, N., & Papageorgiou, L. G. (2013). Efficient energy consumption and operation management in a smart building with microgrid. *Energy Conversion and Management, 74*, 209–222.

Zhang, D., Liu, S., & Papageorgiou, L. G. (2014). Fair cost distribution among smart homes with microgrid. *Energy Conversion and Management, 80*, 498–508.

Zhang, D., Evangelisti, S., Lettieri, P., & Papageorgiou, L. G. (2016). Economic and environmental scheduling of smart homes with microgrid: DER operation and electrical tasks. *Energy Conversion and Management, 110*, 113–124.

Chapter 18
Optimal Demand Side Management for Cryogenic Air Separation Plants

Qi Zhang, Ignacio E. Grossmann and Jose M. Pinto

Abstract The management of electricity demand, also referred to as demand side management (DSM), has been recognized as an effective approach to improving power grid performance. For electricity consumers, DSM constitutes the opportunity to benefit from financial incentives by adjusting their electricity consumption. The cost of electricity is the single largest variable operating cost incurred in cryogenic air separation plants; hence, there is a strong interest in reducing the electricity cost in such plants through DSM. However, to perform effective DSM, we need to develop systematic and innovative decision-making tools that can help us answer questions such as the following: How much potential for load adjustment exists in the plant? How can production and energy management decisions be optimized in an integrated fashion? How can we make long-term strategic decisions while considering hourly changing electricity prices? How do we deal with uncertainty in process data and future information? In this chapter, we draw insights from multiple projects in which we have used mathematical optimization approaches to perform industrial DSM and demonstrated the potential for air separation plants using real-world case studies. In particular, we emphasize the importance of accurate integrated scheduling models, the impact of considering uncertainty and risk in the decision-making, the implementation of robust models, and the modeling of multiple time scales.

Q. Zhang · I.E. Grossmann (✉)
Department of Chemical Engineering, Carnegie Mellon University,
Pittsburgh, PA 15213, USA
e-mail: grossmann@cmu.edu

Q. Zhang
e-mail: qz@andrew.cmu.edu

J.M. Pinto
Business and Supply Chain Optimization R&D, Praxair, Inc.,
Danbury, CT 06810, USA
e-mail: jose_m_pinto@praxair.com

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_18

18.1 Introduction

The power grid is designed to reliably match electricity supply and demand. This task has become increasingly challenging due to high fluctuations in electricity demand, deregulation of electricity markets, and increasing penetration of intermittent renewable energy into the electricity supply mix. The steadily rising energy demand and the pressure to reduce greenhouse gas emissions have further amplified the need to improve the efficiency, reliability, and sustainability of the power grid. In recent years, the notion of a *smart grid* (Farhangi 2010) has been evolving, which represents the idea of effectively coordinating the major grid operations—electricity generation, transmission, distribution, and consumption—through improved communication, holistic optimization, and market design.

One major innovation in the smart grid concept is the utilization of the load adjustment capabilities on the electricity consumers' side, referred to as *demand side management* (DSM). A definition of DSM can be stated as follows (Zhang and Grossmann 2016): "DSM encompasses systematic activities at the interplay between grid operator and electricity consumer aiming at changing the amount and/or timing of the consumer's use of electricity in order to increase grid performance and consumer benefits. DSM activities on the grid operator side involve the assessment of the need for load adjustment and the creation of financial incentives for the consumer, while the consumer reacts to these financial incentives and performs the actual physical load adjustment operations." Ideally, DSM creates a win-win situation in which the grid performance is improved at lower cost, while the consumer providing DSM services benefits from the financial rewards.

One distinguishes between three consumer sectors: residential, commercial, and industrial. Although large untapped DSM potentials exist in all three sectors (Gellings et al. 2006), power-intensive industries are especially suited for DSM because of the high electricity consumption of individual plants and the already existing advanced metering and automation infrastructure. The high potential impact of industrial DSM is widely acknowledged (Paulus and Borggrefe 2011; Samad and Kiliccote 2012) and has been the focus of increased research efforts in recent years. Because of the time-sensitive nature of electricity prices and DSM activities, efficient planning, scheduling, and control of plant operations are crucial for enabling effective DSM (Merkert et al. 2014).

In a recent comprehensive review, Zhang and Grossmann (2016) identify the four main challenges in industrial DSM: (1) accurate modeling of operational flexibility, (2) integration of production and energy management, (3) optimization across multiple time scales, and (4) decision-making under uncertainty. Over the past few years, we have developed mathematical optimization frameworks that address these challenges. In particular, we consider the scheduling of power-intensive continuous processes under various market conditions and uncertainty. Several industrial case studies have been conducted, in which the proposed models have been primarily applied to real-world cryogenic air separation plants.

Cryogenic air separation is a highly complex process that requires effective heat integration and process control (Castle 2002). In order to reach the low temperatures in the cryogenic process, considerable compression work is required; hence, cryogenic air separation units (ASUs) consume very large amounts of electricity. For example, an industrial-scale ASU producing 500 tons/day of liquefied gas consumes approximately 300 MWh of electricity every day, approximately the same amount consumed by 10,000 households. The reduction of electricity cost is a major goal in the design and operation of ASUs, and in this context, air separation companies have discovered the benefit of DSM. However, while ASUs have the operational flexibility that enables relatively quick load adjustment, their DSM potential has only been partially exploited due to the lack of efficient systematic decision-making tools.

First proposed by Ierapetritou et al. (2002) and further developed by Karwan and Keblis (2007), the concept of operating modes has been widely applied in scheduling models for air separation plants. In a mode-based model, the process can only operate in one of the given operating modes, and each mode is defined by a specific feasible region in the product space and a power consumption function with respect to the production rates. Mitra et al. (2012, 2013) present improved formulations and introduce additional constraints to impose restrictions on the transitions between different modes. In the most recent development, Zhang et al. (2016c) have further generalized the model such that it can also be used to represent continuous process networks. Mitra et al. (2014) apply a stochastic programming approach to capacity expansion planning for air separation plants under demand uncertainty. In addition to uncertainty in product demand, Zhang et al. (2016a) also consider uncertainty in electricity price in the context of risk-based integrated production scheduling and electricity procurement. Zhang et al. (2015a) assess the operational benefit of adding a cryogenic energy storage system to an existing air separation plant. Such an integrated plant would allow the generation of electricity, which could be sold back to the market or used to provide operating reserve. To account for the uncertainty in reserve demand, a robust optimization approach has been applied. A similar approach is taken by Zhang et al. (2016b) when optimizing the provision of interruptible load.

In this chapter, we summarize the main results from several previous works and draw insights for the planning and scheduling of cryogenic air separation plants while considering various aspects of DSM. First, in Sect. 18.2, we present a deterministic discrete-time scheduling model that can accurately represent the operational flexibility of an air separation plant. The proposed model considers hourly changing electricity prices and incorporates various types of power contracts. The importance of modeling uncertainty and risk is highlighted in Sect. 18.3, where the problem of integrated production scheduling and electricity procurement is considered. In Sect. 18.4, we demonstrate the benefit of providing interruptible load where the demand for load reduction is not known in advance. Finally, in addition to uncertainty, the capacity planning problem presented in Sect. 18.5 also addresses the challenge of optimization across multiple time scales. We close with some concluding remarks in Sect. 18.6.

Note that a vital part of this work is the analysis of the results from real-world case studies for which the data are provided by Praxair; however, due to confidentiality reasons, we cannot disclose detailed information about plant specifications and actual product demands. Therefore, all results presented in this chapter are given as dimensionless quantities, and numerical values are normalized if necessary.

18.2 Discrete-Time Scheduling Model for Integrated Production and Energy Management

An air separation plant with multiple pieces of equipment or sets of interconnected pieces of equipment and multiple intermediate and final products can be conveniently represented as a process network (Zhang et al. 2016c). Figure 18.1 shows the process network representation of the air separation plant used in our case study. The air separation (AS) process produces gaseous oxygen (GO2), gaseous nitrogen (GN2), liquid oxygen (LO2), liquid nitrogen (LN2), and liquid argon (LAr). While GO2, LO2, LN2, and LAr can be directly sold to the market, GN2 has to be further compressed before it can be supplied to the customers. Two kinds of GN2 are sold: medium-pressure GN2 (MPGN2) and high-pressure GN2 (HPGN2). GN2 is compressed to MPGN2 through Process LMCompGN2 and can be further compressed to HPGN2 through Process MHCompGN2; it can also be directly converted to HPGN2 by running Process LHCompGN2. Furthermore, GN2 can be liquefied to LN2 through Process LiqGN2. Overproduced gaseous products can be vented through a venting process, and all liquid products can be converted into the corresponding gaseous products through a so-called driox process.

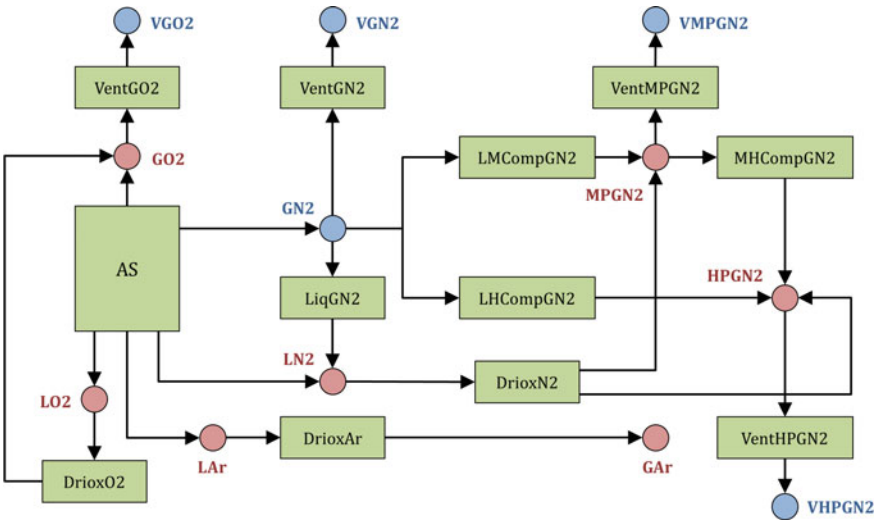

Fig. 18.1 Process network representation of the given air separation plant, consisting of process nodes (*rectangles*), material nodes (*circles*), and arcs depicting the material flows

Inventory capacities are given for the liquid products. Demand is given for all final products. Electricity can be purchased from various power contracts which may differ in price, availability, and penalty for under- or overconsumption. Here, we assume that the prices and other conditions for signed contracts are known in advance, and that accurate forecasts for the day-ahead or real-time electricity prices are available.

The goal is to find a production schedule over a given time horizon that minimizes the total electricity cost while satisfying all product demand. For every time period of the scheduling horizon, we determine

- The mode of operation for each process,
- The production rates in each process,
- The amounts of products stored,
- The amounts of products purchased,
- The amount of electricity purchased from each power contract.

18.2.1 Model Formulation

In the following, we present a mixed-integer linear programming (MILP) scheduling model with a discrete time representation, in which the time horizon is divided into periods of equal length (typically an hour). In a discrete-time model, time-varying electricity prices can be considered by simply assigning different price values to each time period. Although this approach may require a large number of time periods, the resulting models tend to be more efficient than existing continuous-time models (Castro et al. 2009). The notation for the time discretization is such that time period t starts at time point $t-1$ and ends at time point t. The scheduling horizon is defined by the set of time periods $\overline{T} = \{1, 2, \ldots, t^{\text{fin}}\}$, which is a subset of $T = \{-\theta^{\max} + 1, -\theta^{\max} + 2, \ldots, 0, 1, \ldots, t^{\text{fin}}\}$, which also includes time periods in the past. Note that all continuous variables in this model are constrained to be nonnegative.

18.2.1.1 Mass Balance Constraints

For a given process network operating continuously in each time period t, the mass balance constraints can be stated as follows:

$$Q_{jt} = Q_{j,t-1} + \sum_{i \in \hat{I}_j} P_{ijt} - \sum_{i \in \check{I}_j} P_{ijt} + W_{jt} - D_{jt} \quad \forall j,\, t \in \overline{T} \tag{18.1a}$$

$$Q_j^{\min} \le Q_{jt} \le Q_j^{\max} \quad \forall j,\, t \in \overline{T} \tag{18.1b}$$

$$W_{jt} \le W_j^{\max} \quad \forall j,\, t \in \overline{T} \tag{18.1c}$$

where Q_{jt} is the inventory level for material j at time t, and P_{ijt} is the amount of material j consumed or produced by process i in time period t. The set of processes

producing material j is denoted by $\hat{I}_j$, whereas $\bar{I}_j$ is the set of processes receiving material j. The additional purchase of material j in time period t is denoted by W_{jt}. Parameter D_{jt} denotes the demand for material j in time period t. Equation (18.1a) states that the inventory level of a material increases when it is produced or purchased, and the inventory level decreases when the material is consumed by other processes or used to meet product demand. Equation (18.1b) sets lower and upper bounds on the inventory levels. For nonstorable materials (gaseous products), $Q_j^{\min}$ and $Q_j^{\max}$ are zero. Equation (18.1c) limits the amount of material that can be purchased per time period.

18.2.1.2 Process Surrogate Model

It is assumed that each process can operate in different operating modes. Each mode represents a particular operating state, e.g. "off", "on", or "startup". The feasible operating region for each mode is defined by a union of convex subregions in the corresponding material space, and a linear electricity consumption function with respect to the production rates is given for each subregion. The key feature here is that every subregion has the form of a polytope. Such a model is referred to as a Convex Region Surrogate (CRS) model (Zhang et al. 2015b) and can be generated from plant data. Obviously, this representation is an approximation of the true feasible region; however, it allows us to formulate process models that can account for nonconvex feasible regions and are computationally efficient when embedded in an MILP formulation.

At any point in time, a process can only operate in one mode. For a given mode, the operating point has to lie in either one of the convex subregions. Any point in a subregion can be represented as a convex combination of the vertices of the polytope. These relationships are stated in the following:

$$P_{ijt} = \sum_{m \in M_i} \sum_{r \in R_{im}} \overline{P}_{imrjt} \quad \forall\, i, j \in J_i, t \in \overline{T} \tag{18.2a}$$

$$\overline{P}_{imrjt} = \sum_{l \in L_{imr}} \lambda_{imrlt}\, \phi_{imrlj} \quad \forall\, i, m \in M_i, r \in R_{im}, j \in J_i, t \in \overline{T} \tag{18.2b}$$

$$\sum_{l \in L_{imr}} \lambda_{imrlt} = \bar{y}_{imrt} \quad \forall\, i, m \in M_i, r \in R_{im}, t \in \overline{T} \tag{18.2c}$$

$$U_{it} = \sum_{m \in M_i} \sum_{r \in R_{im}} \left(\delta_{imr}\, \bar{y}_{imrt} + \sum_{j \in J_i} \gamma_{imrj}\, \overline{P}_{imrjt} \right) \quad \forall\, i, t \in \overline{T} \tag{18.2d}$$

$$\sum_{m \in M_i} y_{imt} = 1 \quad \forall\, i, t \in \overline{T} \tag{18.2e}$$

$$y_{imt} = \sum_{r \in R_{im}} \bar{y}_{imrt} \quad \forall\, i, m \in M_i, t \in \overline{T} \tag{18.2f}$$

where M_i is the set of modes in which process i can operate, R_{im} is the set of operating subregions in mode $m \in M_i$, L_{imr} is the set of vertices of subregion $r \in R_{im}$, and J_i is the set of input and output materials of process i. The binary variable y_{imt} equals 1 if mode $m \in M_i$ is selected in time period t, while the binary variable $\bar{y}_{imrt}$ equals 1 if subregion $r \in R_{im}$ is selected in time period t. The amount of material $j \in J_i$ consumed or produced by process i is denoted by P_{ijt}. Associated with P_{ijt} is the disaggregated variable $\overline{P}_{imrjt}$ for subregion $r \in R_{im}$, which is expressed as a convex combination of the corresponding vertices, ϕ_{imrlj}, as stated in Eqs. (18.2b)–(18.2c). Variable λ_{imrlt} can only take a nonzero value if subregion $r \in R_{im}$ is selected. As shown in Eq. (18.2d), the amount of electricity consumed, U_{it}, is a linear function of P_{ijt} with a constant δ_{imr} and coefficients γ_{imrj} specific to the selected subregion. Equation (18.2e) states that one and only one mode has to be selected for each process in each time period, and according to Eq. (18.2f), one region in R_{im} has to be selected if process i operates in mode $m \in M_i$.

18.2.1.3 Transition Constraints

A transition occurs when the system changes from one operating point to another. For changes between operating points belonging to the same operating mode, a bound on the rate of change, $\Delta^{\max}_{imj}$, can be set as follows:

$$-\Delta^{\max}_{imj} \le \sum_{r \in R_{im}} \left(\overline{P}_{imrjt} - \overline{P}_{imrj,t-1} \right) \le \Delta^{\max}_{imj} \quad \forall\, i,\, m \in M_i,\, j \in J_i,\, t \in \overline{T}. \tag{18.3}$$

In addition to the rate-of-change constraints, Eqs. (18.4)–(18.6) impose constraints on the transitions between different operating modes. Here, $z_{imm't}$ is a binary variable which equals 1 if and only if process i switches from mode m to mode m' at time t, which is enforced by the following constraint:

$$\sum_{m' \in \overline{TR}_{im}} z_{im'm,t-1} - \sum_{m' \in \widehat{TR}_{im}} z_{imm',t-1} = y_{imt} - y_{im,t-1} \quad \forall\, i,\, m \in M_i,\, t \in \overline{T} \tag{18.4}$$

where $\overline{TR}_{im} = \{m' : (m', m) \in TR_i\}$ and $\widehat{TR}_{im} = \{m' : (m, m') \in TR_i\}$ with TR_i being the set of all possible mode-to-mode transitions for process i.

The restriction that a process has to remain in a certain mode for a minimum amount of time after a transition is expressed by the following constraint:

$$y_{im't} \ge \sum_{k=1}^{\theta_{imm'}} z_{imm',t-k} \quad \forall\, i,\, (m, m') \in TR_i,\, t \in \overline{T} \tag{18.5}$$

with $\theta_{imm'}$ being the minimum stay time in mode m' after switching to it from mode m.

For predefined sequences, each of which defined as a fixed chain of transitions from mode m to mode m' to mode m'', we can specify a fixed stay time in mode m' by imposing the following constraint:

$$z_{imm',t-\bar{\theta}_{imm'm''}} = z_{im'm''t} \quad \forall\, i,\ (m, m', m'') \in SQ_i,\ t \in \overline{T} \tag{18.6}$$

where SQ_i is the set of predefined sequences and $\bar{\theta}_{imm'm''}$ is the fixed stay time in mode m' in the corresponding sequence.

18.2.1.4 Energy Balance Constraints

The required amount of electricity can be purchased from multiple sources, which is stated in Eq. (18.7a). The different power sources are available as power contracts, denoted by index c. Equation (18.7b) sets lower and upper bounds on the electricity purchase from contract c in time period t, E_{ct}. The lower bound, $E_{ct}^{\min}$, is typically zero, but could be nonzero if the contract conditions demand a minimum purchase.

$$\sum_i U_{it} = \sum_c E_{ct} \quad \forall\, t \in \overline{T} \tag{18.7a}$$

$$E_{ct}^{\min} \le E_{ct} \le E_{ct}^{\max} \quad \forall\, c,\ t \in \overline{T} \tag{18.7b}$$

18.2.1.5 Power Contract Model

Equations (18.7) are sufficient if power is only purchased from the spot market or from contracts that are merely defined by a unit price for each time period and possibly some minimum and maximum purchasing restrictions. However, large industrial electricity consumers typically commit themselves to power contracts that provide additional favorable conditions. There are a large variety of such power contract structures; in this model, we consider the two most common types: discount and penalty contracts. With a discount contract, the unit price decreases with increasing amount of purchased electricity. For penalty contracts, the consumer agrees to either purchase at least a certain amount of electricity and pay a penalty for underconsumption, or consume less than a certain amount to avoid penalty for overconsumption.

Discount prices and penalties are defined with respect to the amount of electricity purchased over a certain period of time. In practice, this means that the cumulative electricity purchase is recorded, and there are predefined meter reading times at which the amount of electricity purchased since the last meter reading is computed. According to this cumulative electricity purchase between consecutive meter readings, discount prices and penalties are issued. In this model, we track the cumulative electricity purchase by using the following equations:

$$F_{ct} = F_{c,t-1} + E_{ct} \quad \forall\, c \in \overline{C},\, t \in \overline{T} \setminus \widehat{T}_c \tag{18.8a}$$

$$F_{ct} = F_{c,t-1} + E_{ct} - G_{ct} \quad \forall\, c \in \overline{C},\, t \in \widehat{T}_c \tag{18.8b}$$

$$F_{ct} = 0 \quad \forall\, c \in \overline{C},\, t \in \widehat{T}_c \cup \{0\} \tag{18.8c}$$

where $\overline{C}$ is the set of discount and penalty contracts and $\widehat{T}_c$ is the set of meter reading times for contract c, which does not include time 0. The cumulative electricity purchase is denoted by F_{ct}, and it is reset to zero at every meter reading time. We further introduce the variable G_{ct}, which is the cumulative electricity purchase since the last meter reading before time t.

A block contract formulation is used to model both discount and penalty contracts. In a general block contract, each block is defined by an electricity price and the corresponding amount of electricity that one has to purchase in order to reach this block. The following disjunction represents a block contract:

$$\bigvee_{b \in B_c} \begin{bmatrix} X_{cbt} \\ H_{cb't} = H^{\max}_{cb'} \quad \forall\, b' \in B_c,\, b' < b \\ H_{cbt} \le H^{\max}_{cb} \\ H_{cb't} = 0 \quad \forall\, b' \in B_c,\, b' > b \end{bmatrix} \quad \forall\, c \in \overline{C},\, t \in \widehat{T}_c \tag{18.9a}$$

$$\bigvee_{b \in B_c} X_{cbt} \quad \forall\, c \in \overline{C},\, t \in \widehat{T}_c \tag{18.9b}$$

$$X_{cbt} \in \{\text{true}, \text{false}\} \quad \forall\, c \in \overline{C},\, b \in B_c,\, t \in \widehat{T}_c \tag{18.9c}$$

where B_c is the set of blocks for contract c, H_{cbt} denotes the amount of cumulative electricity purchased in block $b \in B_c$ at time t, and $H^{\max}_{cb}$ is the amount of electricity that one has to purchase in block $b \in B_c$ before reaching the next block. The Boolean variable X_{cbt} is true if block b is the highest block reached for contract c at time t, as stated in Eq. (18.9c). Disjunction (18.9a) states that if X_{cbt} is true, the maximum amount is purchased in all lower blocks $b' < b$, the electricity purchase in block b is bounded by $H^{\max}_{cb}$, and no electricity is purchased in higher blocks $b' > b$. According to logic constraint (18.9b), one and only one X_{cbt} has to be true.

By applying the hull reformulation (Balas 1985), Eqs. (18.9) can be transformed into the following mixed-integer linear constraints:

$$\overline{H}_{cb'bt} = H^{\max}_{cb'} x_{cbt} \quad \forall\, c \in \overline{C},\, b \in B_c,\, b' \in B_c,\, b' < b,\, t \in \widehat{T}_c \tag{18.10a}$$

$$\overline{H}_{cbbt} \le H^{\max}_{cb} x_{cbt} \quad \forall\, c \in \overline{C},\, b \in B_c,\, t \in \widehat{T}_c \tag{18.10b}$$

$$H_{cbt} = \sum_{b' \in B_c, b' \ge b} \overline{H}_{cbb't} \quad \forall\, c \in \overline{C},\, b \in B_c,\, t \in \widehat{T}_c \tag{18.10c}$$

$$\sum_{b \in B_c} x_{cbt} = 1 \quad \forall\, c \in \overline{C},\, t \in \widehat{T}_c \tag{18.10d}$$

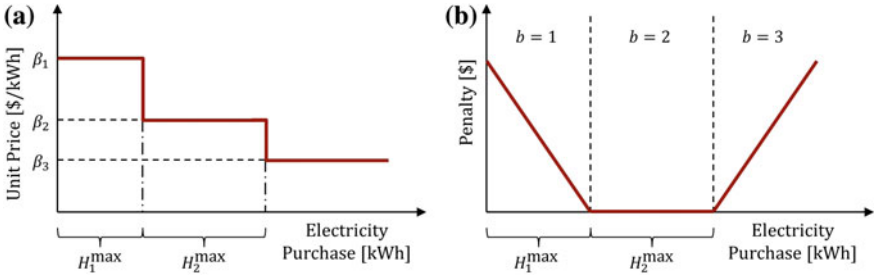

Fig. 18.2 **a** With a discount contract, unit price decreases with increasing electricity purchase **b** A simple penalty contract can be modeled as a three-block contract, with the first and third blocks corresponding to under- and overconsumption, respectively

where x_{cbt} is a binary variable, and $\overline{H}_{cbb't}$ is the disaggregated variable. Figure 18.2 illustrates how discount and penalty contracts can be modeled using this block contract formulation.

To obtain the amount of electricity purchased from a contract, we sum up the electricity purchased in all corresponding contract blocks:

$$G_{ct} = \sum_{b \in B_c} H_{cbt} \quad \forall c \in \overline{C},\, t \in \hat{T}_c. \tag{18.11}$$

18.2.1.6 Boundary Conditions

We solve the scheduling problem for a given time horizon, with boundary conditions stated as follows:

$$Q_{j,0} = Q_j^{\text{ini}} \quad \forall j \tag{18.12a}$$

$$y_{im,0} = y_{im}^{\text{ini}} \quad \forall i,\, m \in M_i \tag{18.12b}$$

$$z_{imm't} = z_{imm't}^{\text{ini}} \quad \forall i,\, (m,m') \in TR_i,\, -\theta_i^{\max} + 1 \le t \le -1 \tag{18.12c}$$

$$Q_{j,t^{\text{fin}}} \ge Q_j^{\text{fin}} \quad \forall j \tag{18.12d}$$

with $\theta_i^{\max} = \max\left(\max_{(m,m') \in TR_i}\{\theta_{imm'}\}, \max_{(m,m',m'') \in SQ_i}\{\bar{\theta}_{imm'm''}\}\right)$, which defines for how far back in the past the mode switching information has to be provided. The initial conditions (18.12a)–(18.12c) set the initial inventory levels to Q_j^{ini}, the initial operating modes to y_{im}^{ini}, and the mode switching history according to $z_{imm't}^{\text{ini}}$, while the terminal constraint Eq. (18.12d) sets the lower bounds on the final inventory levels to Q_j^{fin}.

18.2.1.7 Objective Function

The goal is to minimize the total electricity cost, *TC*, as expressed in the following objective function:

$$TC = \sum_{c} \sum_{t \in \overline{T}} \alpha_{ct} E_{ct} + \sum_{c \in \overline{C}} \sum_{t \in \widehat{T}_c} \left[\sum_{b \in B_c} \beta_{cbt} H_{cbt} + \zeta_{ct}^{\mathrm{u}} (H_{c,1}^{\max} - H_{c,1,t}) + \zeta_{ct}^{\mathrm{o}} H_{c,|B_c|,t} \right] \tag{18.13}$$

where α_{ct} and β_{cbt} are unit costs for purchased electricity, whereas ζ_{ct}^{u} and ζ_{ct}^{o} are unit penalty costs for under- and overconsumption, respectively. The first term in Eq. (18.13) represents the base cost and applies to all contracts, while the remaining terms only apply to block contracts.

Note that the price structure of a contract is defined by the cost coefficients α_{ct}, β_{cbt}, ζ_{ct}^{u}, ζ_{ct}^{o}, as well as $H_{cb}^{\max}$. Many combinations are possible, which provides the flexibility of modeling various different power contracts. Some common examples are listed in the following:

- Day-ahead or real-time market: $\alpha_{ct} > 0$, $\beta_{cbt} = 0$, $\zeta_{ct}^{\mathrm{u}} = 0$, $\zeta_{ct}^{\mathrm{o}} = 0$
- Pure discount contract: $\alpha_{ct} = 0$, $\beta_{cbt} > 0$, $\zeta_{ct}^{\mathrm{u}} = 0$, $\zeta_{ct}^{\mathrm{o}} = 0$
- Contract with partial price discount: $\alpha_{ct} > 0$, $\beta_{cbt} > 0$, $\zeta_{ct}^{\mathrm{u}} = 0$, $\zeta_{ct}^{\mathrm{o}} = 0$
- Contract with penalty for underconsumption: $\alpha_{ct} > 0$, $\beta_{cbt} = 0$, $\zeta_{ct}^{\mathrm{u}} > 0$, $\zeta_{ct}^{\mathrm{o}} = 0$
- Contract with penalty for overconsumption: $\alpha_{ct} > 0$, $\beta_{cbt} = 0$, $\zeta_{ct}^{\mathrm{u}} = 0$, $\zeta_{ct}^{\mathrm{o}} > 0$
- Contract with penalties for under- and overconsumption: $\alpha_{ct} > 0$, $\beta_{cbt} = 0$, $\zeta_{ct}^{\mathrm{u}} > 0$, $\zeta_{ct}^{\mathrm{o}} > 0$
- Combined discount and penalty contract: $\alpha_{ct} > 0$, $\beta_{cbt} > 0$, $\zeta_{ct}^{\mathrm{u}} > 0$, $\zeta_{ct}^{\mathrm{o}} > 0$

18.2.2 Case Study

The proposed model is applied to the plant shown in Fig. 18.1 with a scheduling horizon of one week. The model is implemented in GAMS 24.4.1 (GAMS Development Corporation 2015) and solved using the commercial solver CPLEX 12.6.1 on an Intel® Core™ i7-2600 machine at 3.40 GHz with 8 hardware threads and 8 GB RAM running Windows 7 Professional.

An hourly time discretization is applied, resulting in 168 time periods. Demand profiles for all products are given. Note that while there is continuous demand for gaseous products, the demand for each liquid product is assumed to occur only at the end of each day. The assumption here is that there is sufficient capacity in the inventory to handle flows into and out of the inventory tank throughout the day. The demand at the end of the day is then the total amount of product that needs to be drawn from the tank over the course of the day.

Figure 18.3 shows the electricity consumption profiles for each of the processes as suggested by the optimal solution of the MILP model. The vast majority of the elec-

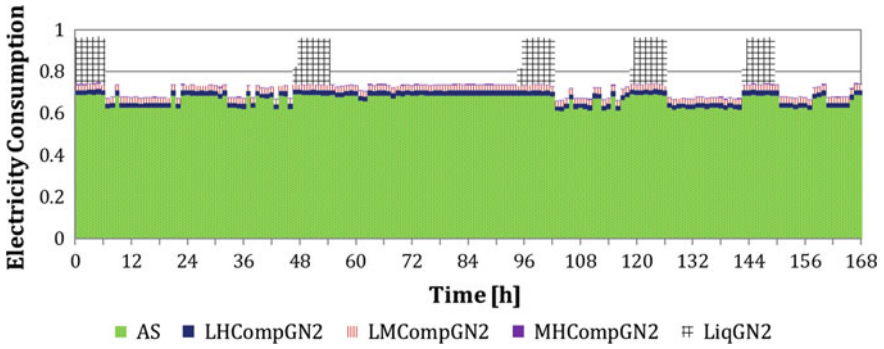

Fig. 18.3 Amount of electricity consumed by each process of the air separation plant

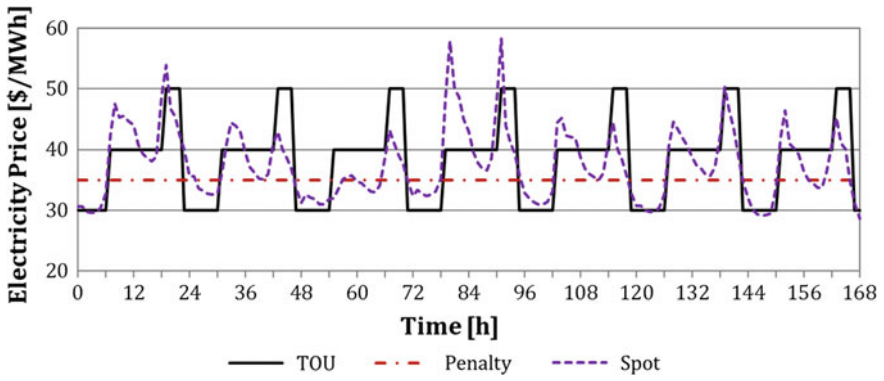

Fig. 18.4 Electricity prices for the TOU contract, the penalty contract, and the spot market

tricity consumption is attributed to the AS process. The GN2 liquefier also consumes a large amount of electricity but is only used five times, each time for a few hours. Compared to the ASU and the GN2 liquefier, the pipeline compressors contribute relatively little to the total electricity consumption. Significant load shifting can be observed in the schedule; this is mainly realized by operating the liquefier during low-price hours, which allows a fairly constant operation of the other processes.

Three different sources from which electricity can be purchased are considered: a TOU contract, a penalty contract with penalty for underconsumption, and the spot market. The corresponding electricity prices are shown in Fig. 18.4. The breakdown of the total electricity purchase into the purchases from the three different sources is shown in Fig. 18.5. One can observe that in each time period, the solution suggests purchasing from the source with the lowest price. Two sources are chosen in the same time period only when the maximum purchase amount is reached for one of the two sources. Also, sufficient amount of electricity is purchased from the penalty contract such that no penalty has to be paid.

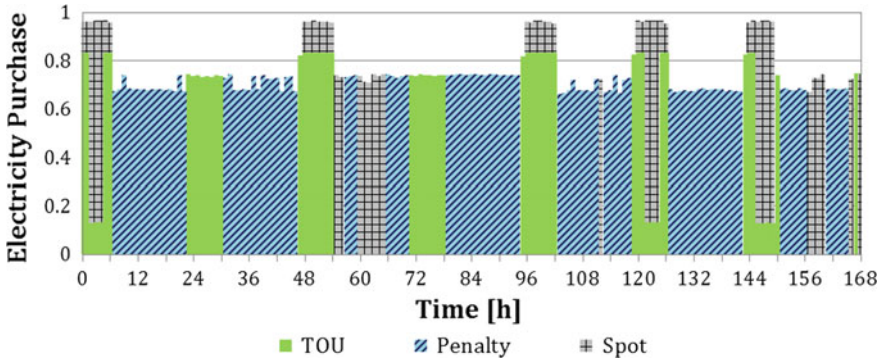

Fig. 18.5 Breakdown of electricity purchase into purchases from the three different sources

As an example, Fig. 18.6 shows the amount of GN2 produced by the ASU and the GN2 flows into the subsequent processes. Large portion of GN2 is compressed to feed the product pipelines. Another significant amount of GN2 is liquefied to increase the production of LN2. However, the majority of GN2 is vented. The reason for this overproduction of GN2 is that the plant is oxygen-limited, which means that the production is driven by the oxygen demand. At such a plant, the nitrogen production usually exceeds the demand, which makes venting of GN2 necessary.

Load shifting in an air separation plant is only possible because of its capacity for storing liquid products. Figure 18.7 shows the amounts of LN2 produced by the ASU and the GN2 liquefier in each time period. The inventory increases when LN2 is produced and it decreases when LN2 is drawn from the tank to satisfy demand, which occurs every 24 h and is not shown in the figure. One can see that in addition to the production through the ASU, a large amount of LN2 is obtained by liquefying GN2. In this solution, driox is not used in any of the time periods.

The MILP model for this industrial-scale problem has 28,808 continuous variables, 9,776 binary variables, and 113,810 constraints. It was solved to zero inte-

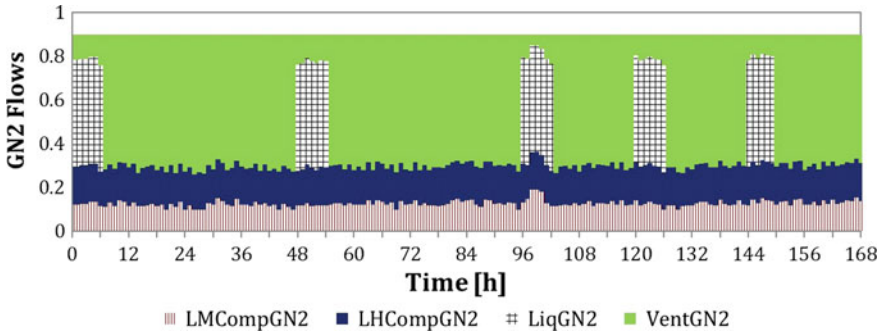

Fig. 18.6 GN2 production and its breakdown into feeds for different processes

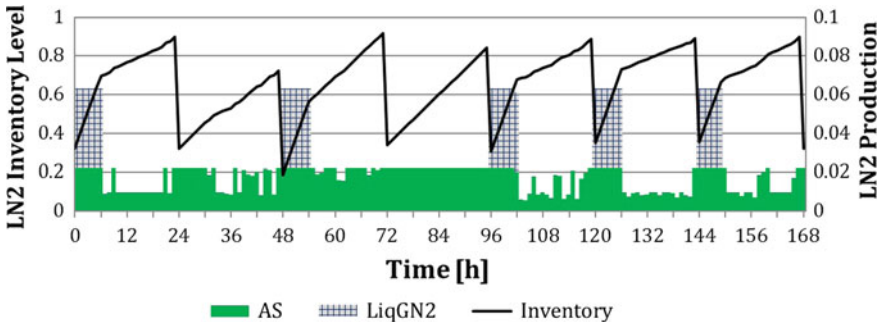

Fig. 18.7 LN2 production and inventory profile

grality gap in 7 seconds wall-clock time. A computational study was conducted by solving several instances of the problem with different input parameters. Almost all instances were solved to optimality within two minutes. The computational efficiency of the model allows its use in a real industrial setting. In fact, a scheduling tool using the proposed modeling framework is currently implemented in one of Praxair's air separation plants.

18.3 Risk-Based Integrated Production Scheduling and Electricity Procurement

While the scheduling model presented in the previous section integrates production and energy management, it does not account for the various sources of uncertainty that frequently arise in industrial DSM. The challenge of decision-making under uncertainty is addressed in the following sections. Each of the proposed optimization frameworks uses a specific version of the model described in Sect. 18.2 as a deterministic base model. In particular, in these variants of the model, the plant is modeled as a single process instead of a process network. In the following, we focus on the presentation of the conceptual ideas and the analysis of the case studies; hence, comprehensive descriptions of the mathematical models are omitted. For more details, we refer to the respective publications.

In this section, we present a stochastic programming approach to the problem of integrated production scheduling and electricity procurement involving the two most critical sources of uncertainty: spot electricity price and product demand (Zhang et al. 2016a). In general, electricity can be purchased from the spot market, or from power contracts that have fixed pre-agreed electricity prices and availability conditions. While purchases from the spot market can be made one day in advance (day-ahead) or on the spot (real-time), one has to commit to the electricity purchase from power contracts for a longer period of time, e.g. for one week. The goal is to optimize

the production and electricity procurement schedules in terms of expected profit and risk over a given time horizon.

The decisions are divided into two sets: one containing *here-and-now* decisions that have to be made at the beginning and cannot be changed over the course of the scheduling horizon; the other containing *wait-and-see* decisions that can be adjusted after realization of the uncertainty. In this problem, the here-and-now decisions are the mode of operation for the production process and the amount of electricity purchased from each power contract in each time period of the scheduling horizon. The wait-and-see decisions are the actual production rates, the amounts of products stored, the amounts of products purchased, and the amount of electricity purchased from the spot market.

18.3.1 Risk Modeling Using Conditional Value-at-Risk

In stochastic programming (Birge and Louveaux 2011), uncertainty is represented by discrete scenarios, and decisions are made at different stages, which are defined such that realization of uncertainty is observed between two stages, and at each stage, actions that depend on previous observations are taken. The given problem has a multistage character since in every time period, the spot electricity price changes and the production rates can be adjusted. However, the resulting multistage stochastic programming problem is extremely large and computationally intractable. Therefore, we approximate the multistage problem with a two-stage formulation where all uncertainty for the entire scheduling horizon is assumed to realize right after the here-and-now decisions are made. We define the set of product demand scenarios, S^{D}; similarly, S^{P} denotes the set of electricity price scenarios. Each pair of demand scenario and price scenario corresponds to a general scenario $s \in S$ with the probability φ_s. The total expected profit, TP, can then be expressed as follows:

$$TP = \sum_{s} \varphi_s \left[RV_s - \sum_{t \in \overline{T}} \left(\alpha_{ts}^{\mathrm{ES}} ES_{ts} + \sum_{j} \alpha_{jt}^{\mathrm{PC}} PC_{jts} \right) \right] - \sum_{c} BC_c - \sum_{t \in \overline{T}} \sum_{c} \alpha_{ct}^{\mathrm{EC}} EC_{ct} \tag{18.14}$$

where RV_s denotes the revenue in scenario s, ES_{ts} is the amount of electricity purchased from the spot market in time period t of scenario s, PC_{jts} is the amount of product j purchased in time period t of scenario s, BC_c is the base cost for purchasing from contract c, and EC_{ct} is the amount of electricity purchased from contract c in time period t. Cost coefficients are denoted by $\alpha_{ts}^{\mathrm{ES}}$, $\alpha_{jt}^{\mathrm{PC}}$, and $\alpha_{ct}^{\mathrm{EC}}$.

As a risk measure, the conditional value-at-risk (CVaR) (Rockafellar and Uryasev 2000) is applied, which is defined as the expected profit computed over the scenarios that have profit values smaller than the $(1 - \alpha)$-quantile of the profit distribution. By increasing the CVaR, we reduce the financial risk of encountering a scenario with very low profit. The CVaR, denoted by CV, is incorporated into the model by adding the following constraints to the formulation:

$$CV = \kappa - \frac{1}{1-\alpha} \sum_{s} \varphi_s \omega_s \tag{18.15a}$$

$$\kappa - \left[R_s - \sum_{t \in \overline{T}} \left(\alpha_{ts}^{\mathrm{ES}} ES_{ts} + \sum_{j} \alpha_{jt}^{\mathrm{PC}} PC_{jts} \right) - \sum_{c} BC_c - \sum_{t \in \overline{T}} \sum_{c} \alpha_{ct}^{\mathrm{EC}} EC_{ct} \right] \leq \omega_s \quad \forall s \tag{18.15b}$$

where κ and ω_s are continuous variables with $\kappa \in \mathbb{R}$ and $\omega_s \geq 0$. For each scenario in which the profit is less than κ, ω_s takes the value of the difference between κ and the profit; otherwise, ω_s is zero. When CV is maximized, it takes the value of the α-CVaR.

The objective is to maximize a weighted sum of the total expected profit and the CVaR, i.e.

$$\zeta\, TP + (1-\zeta) CV \tag{18.16}$$

with $\zeta \in [0, 1]$. Because there is usually a trade-off between expected outcome and risk in the sense that no solution can be found that maximizes both TP and CV, the weighting factor ζ can be used to specify which objective should be emphasized more. In this context, ζ can also be seen as a parameter that sets the desired level of risk aversion. The smaller ζ, the more risk-averse is the solution, since more weight is assigned to the CVaR. If $\zeta = 1$, we are risk-neutral and only maximize TP; if $\zeta = 0$, only CV is maximized.

18.3.2 Case Study

The stochastic optimization framework is applied to an air separation plant that has to meet demand for LO2 and LN2. The uncertainty in product demand is characterized by five scenarios that resemble a normal distribution. To model the uncertainty in electricity price, an ARIMAX model with temperature as exogenous input is created using data from four consecutive weeks as training data. Monte Carlo simulation is then applied to generate 1000 equiprobable price scenarios, each scenario corresponding to a price profile over the entire scheduling horizon. The price profiles for all scenarios are shown in Fig. 18.8 along with the expected price profile. Because lines are shown in the same color, the individual scenarios are indistinguishable in the diagram; however, the picture depicts the large spread in the price distribution. One can see that the level of uncertainty increases with time.

A model considering all 1000 price scenarios is computationally intractable; hence, we apply scenario reduction techniques (Dupacova et al. 2003) to reduce the number of scenarios to 50 such that the probability distribution represented by the reduced scenario set (denoted by $\tilde{S}^{\mathrm{P}}$) is close to the one represented by the original full scenario set. The price profiles for the reduced scenario set are shown in Fig. 18.9. Note that the scenario reduction process assigns different probabilities to the scenarios in the reduced set.

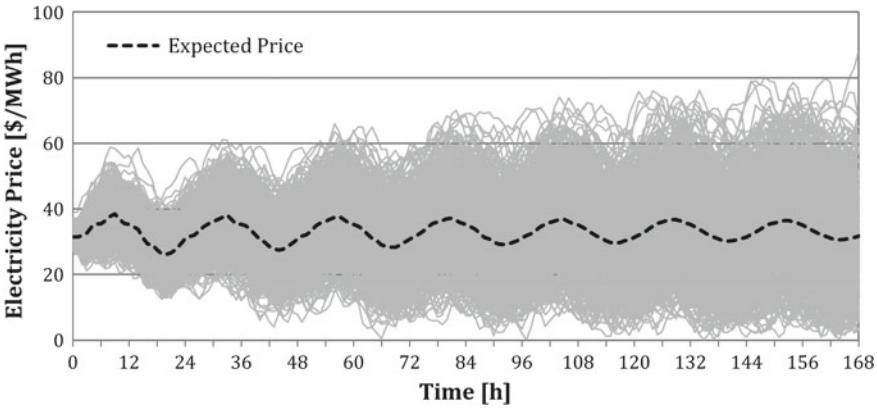

Fig. 18.8 Full set of spot electricity price profiles and the expected price profile

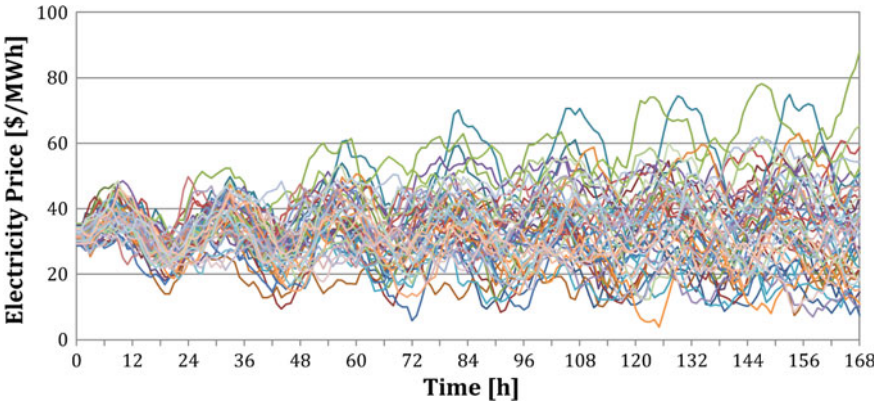

Fig. 18.9 Spot electricity price profiles associated with the reduced set of scenarios

To compare the deterministic and stochastic solutions, the value of stochastic solution (VSS) is used as a measure for the difference in the solutions. Here, the VSS is defined as

$$\mathrm{VSS} = \zeta \left(TP^{\mathrm{sto}} - TP^{\mathrm{det}} \right) + (1 - \zeta) \left(CV^{\mathrm{sto}} - CV^{\mathrm{det}} \right) \tag{18.17}$$

where the superscript "sto" denotes values obtained at the optimal solution of the stochastic problem, and "det" denotes values computed by solving the stochastic problem with fixed first-stage decisions obtained from the deterministic problem. Since the stochastic optimization is performed on a reduced scenario set, but the VSS is computed over the full scenario set, only an approximate VSS (AVSS) can be obtained. The AVSS would take the value of the true VSS if the obtained first-

stage decisions were the same as the ones that one would obtain from solving the stochastic problem with the full set of scenarios.

We create three cases with different levels of demand uncertainty (i.e. variances in the demand scenarios, denoted by Var^{D}): low, medium, and high. For each case, two instances are created, one neglecting electricity price uncertainty ($|\tilde{S}^{\text{P}}| = 1$), the other incorporating the 50 price scenarios from the reduced scenario set ($|\tilde{S}^{\text{P}}| = 50$). Risk-neutral and risk-averse ($\zeta = 0.5$, $\alpha = 0.9$) optimization are performed on all instances for which the results are shown in Tables 18.1 and 18.2, respectively.

The most notable observation is that in risk-neutral optimization, accounting for price uncertainty does not seem to provide any added value. In each of the three cases, the VSS remains the same when $|\tilde{S}^{\text{P}}|$ is increased from 1 to 50. In contrast, in risk-averse optimization, there is a clear increase in VSS when price uncertainty is considered in the stochastic optimization. This result suggests that accounting for electricity price uncertainty in risk-neutral optimization may be unnecessary, whereas it can lead to significant additional benefit in risk-averse optimization. The explanation for this phenomenon lies in a special characteristic of electricity price uncertainty: electricity prices in different scenarios may differ considerably in magnitude; however, all price profiles follow essentially the same trend. Therefore, in the risk-neutral case, the second-stage decisions tend to be the same across all scenarios with the same product demand; however, this statement does not apply to the risk-averse case because of the additional CVaR term in the objective function. For an in-depth explanation, see Zhang et al. (2016a).

Table 18.1 Expected cost and VSS from risk-neutral optimization

$\lvert S^{\text{D}}\rvert$	Var^{D}	$\lvert S^{\text{P}}\rvert$	$\lvert\tilde{S}^{\text{P}}\rvert$	TP^{det}	TP^{sto}	AVSS	$\overline{\text{AVSS}}$ [%]
5	Low	1000	1	15.51	15.51	0.00	0.0
			50		15.51	0.00	0.0
5	Medium	1000	1	14.60	14.92	0.32	2.1
			50		14.92	0.32	2.1
5	High	1000	1	12.02	13.18	1.16	8.7
			50		13.18	1.16	8.7

Table 18.2 Expected profits, CVaRs, and VSS from risk-averse optimization

$\lvert S^{\text{D}}\rvert$	Var^{D}	$\lvert S^{\text{P}}\rvert$	$\lvert\tilde{S}^{\text{P}}\rvert$	TP^{det}	CV^{det}	TP^{sto}	CV^{sto}	AVSS	$\overline{\text{AVSS}}$ [%]
5	Low	1000	1	15.51	12.84	15.51	12.84	0.00	0.0
			50			15.27	13.68	0.30	2.1
5	Medium	1000	1	14.60	11.15	14.69	11.39	0.17	1.3
			50			14.80	11.59	0.32	2.5
5	High	1000	1	12.02	5.78	12.82	6.34	0.68	7.6
			50			12.71	6.57	0.74	8.3

In Fig. 18.10, we compare for the high-Var^D case the solutions obtained from the deterministic, risk-neutral, and risk-averse optimization. One can see that these three approaches lead to very different decisions. In deterministic optimization, uncertainty is ignored such that decisions are primarily driven by the differences between the power contract prices and the expected spot price. A significant amount of electricity is procured from Contract 2 because the price discount at this purchasing amount makes it less expensive than purchasing from the spot market during on-peak hours. One issue with the deterministic solution is only implicitly shown in the diagram, namely that there is a very high expected cost of purchasing additional products because the selected operating modes do not have sufficient production capacities in the high-demand scenarios.

In risk-neutral stochastic optimization, first-stage decisions are to a large extent driven by the need for flexibility that has to be maintained in the second stage in order to react to different scenarios. Deterministic and risk-neutral optimization lead to similar schedules for the electricity procurement from power contracts. However,

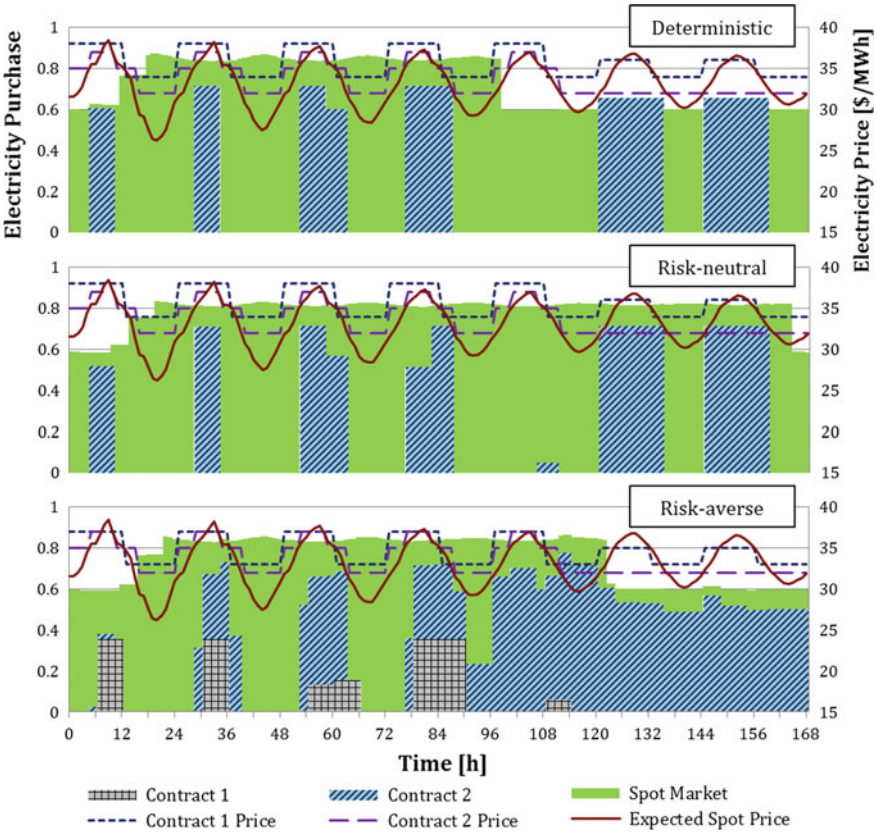

Fig. 18.10 Electricity purchase profiles for the high-Var^D case

the risk-neutral solution suggests selecting operating modes with higher production capacities in order to be able to accommodate high-demand scenarios. This strategy leads to considerably lower costs for purchasing additional products, especially during the last three days of the week.

The electricity procurement decisions resulting from risk-averse optimization are very different from the deterministic and risk-neutral solutions. Here, more than half of the required electricity is purchased from power contracts. Because no more electricity can be purchased from Contract 2 due to the specified purchase limit, electricity is also procured from Contract 1. Here, contracts are effectively used to hedge against the risk of low-profit scenarios. In particular, considerable amount of electricity is purchased from contracts toward the end of the week when the level of uncertainty in the spot electricity price is highest.

All instances were solved by applying a multicut Benders decomposition algorithm on the same machine as the one specified in Sect. 18.2.2. For details on the algorithm and computational results, we refer to Zhang et al. (2016a).

18.4 Robust Provision of Interruptible Load

The power grid is designed to match electricity supply and demand at all times. When real-time electricity supply falls below the demand, e.g. due to generator failures or sudden load changes, backup capacities are called upon in order to eliminate the supply-demand gap. One type of such backup capacity is called *operating reserve*, which has to be dispatched within minutes upon request. Providing reserve capacity is lucrative because the reserve provider is rewarded even when no actual dispatch is required. Operating reserve can be provided by generating facilities that are able to quickly increase electricity supply. Alternatively, the supply-demand gap can be eliminated by reducing demand. Therefore, electricity consumers also have the opportunity to provide operating reserve if they possess the flexibility to quickly reduce their electricity consumption. Such operating reserve provided by electricity consumers is also referred to as *interruptible load*.

To provide interruptible load, the electricity consumer has to specify a target power consumption rate and a maximum possible reduction with respect to the target power consumption. The amount of maximum load reduction is then the amount of interruptible load that the electricity consumer provides. Modern air separation plants can quickly ramp down their electricity consumption; hence, they can provide significant amounts of interruptible load. However, since load reduction is only requested by the grid operator in case of contingency, one does not know in advance when and how much load reduction will be required; yet dispatch upon request has to be guaranteed since otherwise one has to pay extremely high penalties. Therefore, the uncertainty in load reduction demand has to be taken into account when optimizing the amount of interruptible load to provide (Zhang et al. 2016b).

18.4.1 Applying Affinely Adjustable Robust Optimization

Since feasible load reduction has to be guaranteed, robust optimization (Ben-Tal et al. 2009) is the natural choice for solving the given problem. In robust optimization, the worst case is optimized while guaranteeing feasibility for all possible realizations of the uncertainty, which is described by an uncertainty set. Here, we apply an affinely adjustable robust optimization (AARO) approach, which, unlike traditional static robust optimization, incorporates recourse.

We define the uncertainty set as follows:

$$W(IL) = \left\{ w : \left(LR_t = IL_t\, w_t,\ 0 \le w_k \le 1\ \forall k = 1, \ldots, t,\ \sum_{k=1}^{t} w_k \le \Gamma_t \right) \forall t \in \overline{T} \right\} \tag{18.18}$$

where LR_t denotes the actual load reduction in time period t, IL_t is the amount of interruptible load provided in time period t, $w_t = LR_t/IL_t$ is the normalized required load reduction, and Γ_t is a budget parameter limiting the cumulative load reduction required up to time t. By changing Γ_t, the level of robustness can be adjusted. In practice, the budget parameters can be chosen based on historical data; alternatively, depending on the market, there may be a strict limit on the number of times in which load reduction can be requested during a specific time horizon, which can be used to set Γ_t.

In this problem, the here-and-now decisions are the modes of operation, the target production rates for each product, and the committed purchase amounts for each product in each time period of the scheduling horizon. The wait-and-see decisions are the changes in production rates and product purchases if load reduction is requested or has been requested in previous time periods. In general, the recourse variables related to the wait-and-see decisions are functions of the uncertain parameters. In order to obtain a tractable formulation, these functions are restricted to be affine functions of the uncertain parameters. For instance, $\overline{P}_{mrjt}$, the production rate for product j in subregion r of mode m in time period t, can be expressed as

$$\overline{P}_{mrjt} = \widehat{P}_{mrjt} + \sum_{k=t-\zeta_t}^{t} p_{mrjtk}\, w_k \tag{18.19}$$

where $\widehat{P}_{mrjt}$ is the target production rate, and the second term constitutes a linear decision rule determining the change in production rate depending on the realization of the uncertainty w. The linear decision rule can be optimized by changing the coefficients p_{mrjtk}. This formulation allows multistage decision-making since at each time period t, the recourse decision depends on the uncertain parameters that have been realized in the preceding ζ_t time periods as well as the current time period, i.e. w_k for $t - \zeta_t \le k \le t$. The parameter ζ_t can be any integer between zero and $t - 1$; it can be seen as a parameter for adjusting the level of conservatism. The greater ζ_t, the more realized uncertainty is considered in the recourse; hence, an improved solution

may be obtained. However, the problem size increases with ζ_t. Therefore, in large-scale problems, $\zeta_t < t - 1$ can be chosen in order to achieve a trade-off between level of conservatism and problem size.

With the uncertainty set and the linear decision rules outlined above, a tractable robust counterpart can be formulated by applying reformulation techniques that mainly rely on linear programming (LP) duality. For the full derivation of the AARO model, see Zhang et al. (2016b).

18.4.2 Case Study

We again optimize a one-week schedule for an air separation plant that produces LO2 and LN2. Figure 18.11 shows the LO2 and LN2 production, purchase, and inventory profiles for the case in which no interruptible load is provided. The optimal solution suggests to shift production as much as possible to times when the electricity price (shown in Fig. 18.12) is low. The resulting total cost is 100 (normalized).

For the case in which the provision of interruptible load is considered, Γ_t is chosen to increase by 1 every 24 h, i.e. load reduction can be requested up to 7 times over the whole week. This is a fairly conservative assumption, since typically dispatch of operating reserve is only requested a few times in months (EnerNOC 2014). We introduce an auxiliary parameter $\bar{\zeta}$, which denotes the maximum number of previously realized uncertain parameters that are considered in the decision rules, and set ζ_t such that $\zeta_t = \min\{\bar{\zeta}, t - 1\}$. The level of conservatism decreases, but the model size increases with $\bar{\zeta}$.

Various instances with different $\bar{\zeta}$ are created. For each instance, Table 18.3 lists the total cost, the model size, and the wall-clock computation time used to solve the

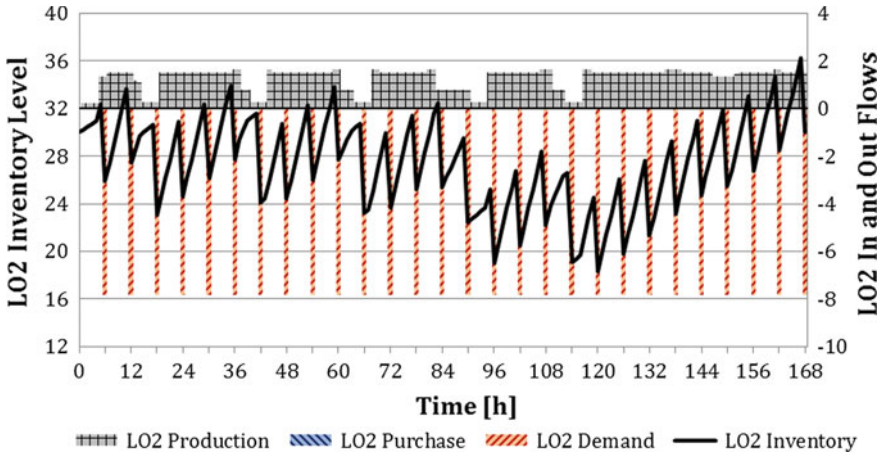

Fig. 18.11 LO2 flows and inventory profile for the case without interruptible load

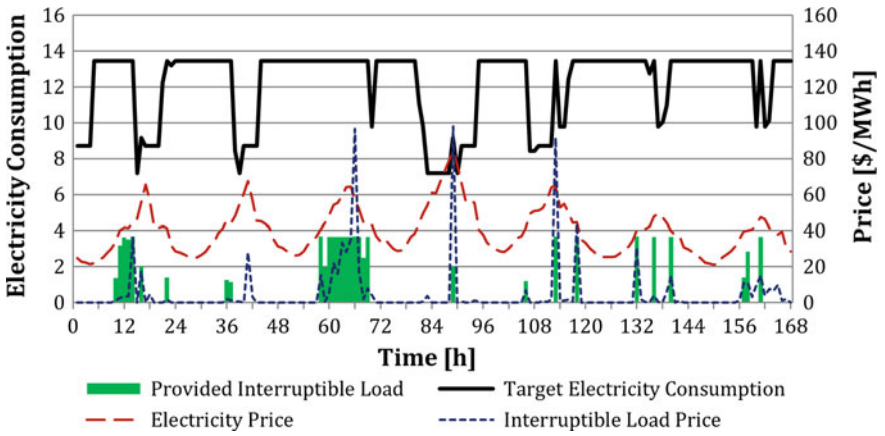

Fig. 18.12 Target electricity consumption profile and provided interruptible load for the case of $\bar{\zeta} = 23$, and price profiles

Table 18.3 Total costs, model sizes, and computation times for cases with different $\bar{\zeta}$

$\bar{\zeta}$	*TC*	# of bin. variables	# of cont. variables	# of constraints	Wall-clock time [s]
0	98.83	3,282	82,670	84,604	185
5	98.48	3,282	139,595	139,879	1,056
11	98.33	3,282	205,628	203,998	2,598
17	98.29	3,282	269,177	265,705	5,846
23	98.20	3,282	330,242	325,000	6,476
35	98.16	3,282	444,920	436,354	18,381
47	98.14	3,282	549,662	538,060	23,280

MILP to 0.1 % optimality gap. As expected, the total cost decreases with increasing $\bar{\zeta}$ due to the higher flexibility in the recourse. However, the numbers of continuous variables and constraints grow with $\bar{\zeta}$, and lead to dramatic increases in computation time. Here, we see the clear trade-off between the level of conservatism in the model and its computational performance. For our further analysis, we choose $\bar{\zeta}$ to be 23, as the required computation time is reasonable for practical purposes and only minor improvement is achieved for $\bar{\zeta} > 23$.

Besides the electricity and interruptible load prices, Fig. 18.12 shows the target load profile and the amount of interruptible load provided for the case of $\bar{\zeta} = 23$. It is worth pointing out that typically, as it is also the case here, high electricity prices coincide with high interruptible load prices. For those time periods, the optimal solution reveals whether reducing production to save electricity cost or increasing production to provide more interruptible load is more beneficial. Figure 18.13 shows the target production, purchase, and inventory profiles as well as the recourse actions. As one can see, the recourse actions allow the plant to ramp down its production

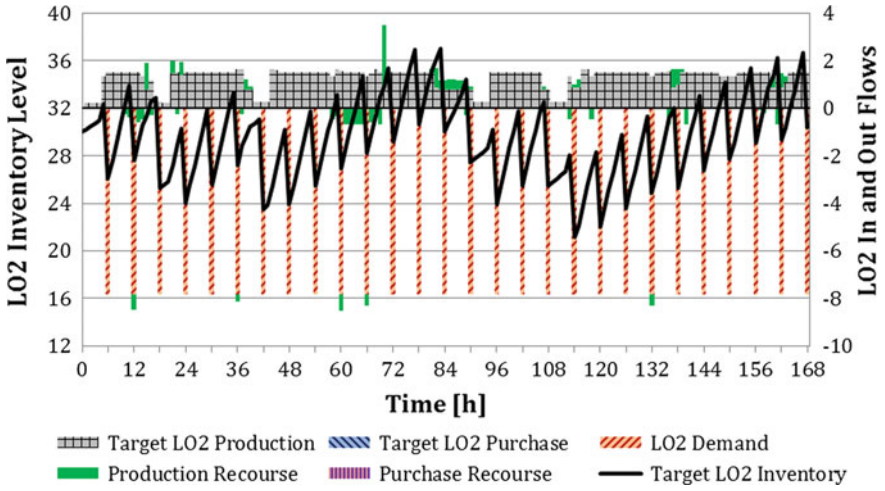

Fig. 18.13 Target and recourse LO2 flows and target inventory profile for the case of $\bar{\zeta} = 23$

if load reduction is required and to make up for the lost production by increasing production after load reduction or by purchasing additional products. Compared to the base case without interruptible load, the total cost reduces by 1.8 %, which may seem small but is actually significant at such industrial scale. Also, note that the cost savings are reported with respect to the worst case, i.e. the actual cost savings are likely to be higher.

18.5 Stochastic Multiscale Capacity Planning

While short-term scheduling problems have been considered in the previous sections, we now address the long-term strategic problem of capacity planning under demand uncertainty (Mitra et al. 2014). The objective is to decide which of the following investments to make over a time horizon of several years: (1) add new pieces of equipment, (2) perform upgrades or replacements of existing equipment, or (3) install additional storage facilities. The investment decisions can be conveniently incorporated in the mode-based scheduling model presented in Sect. 18.2; however, the model becomes prohibitively large if an hourly time discretization is applied to the entire time horizon.

Two-stage stochastic programming is applied to model the uncertainty in product demand. In this problem, the here-and-now decisions are the strategic investment decisions, and the wait-and-see decisions are all operational decisions such as the selection of the modes of operation and production rates.

18.5.1 Modeling Multiple Time Scales

The major modeling challenge is the integration of the different time scales that are involved in the problem. On the one hand, electricity prices fluctuate on an hourly basis in most electricity markets. On the other hand, strategic capacity planning decisions have to be justified for a time horizon of multiple years. However, electricity prices are known to exhibit a seasonal behavior, i.e. for each season, a typical weekly price profile can be identified that applies approximately to all weeks of the same season. Based on this insight, we consider four major periods of operation for each year, corresponding to the seasonal behavior of the electricity prices: spring, summer, fall, and winter. Furthermore, for each season, we consider a representative week that is repeated cyclically and in which electricity prices are specified on an hourly basis. This multiscale time representation is illustrated in Fig. 18.14.

Typically, we want to make decisions on investments in the foreseeable future, i.e. within the next few years; however, we still consider a longer time horizon in order to capture the trade-off between capital expenditures (CAPEX) and operating expenditures (OPEX). In this case, for computational reasons, we may aggregate the later years in which no investment decisions are made. For example, in the case study, the length of the planning horizon is 10 years, but investment decisions can only be made in the first 4 years. To reduce the size of the model, years 5–10 are aggregated into one year with operating costs that are weighted appropriately in the objective function. As shown in Fig. 18.15, this time representation results in a model considering 20 representative seasons.

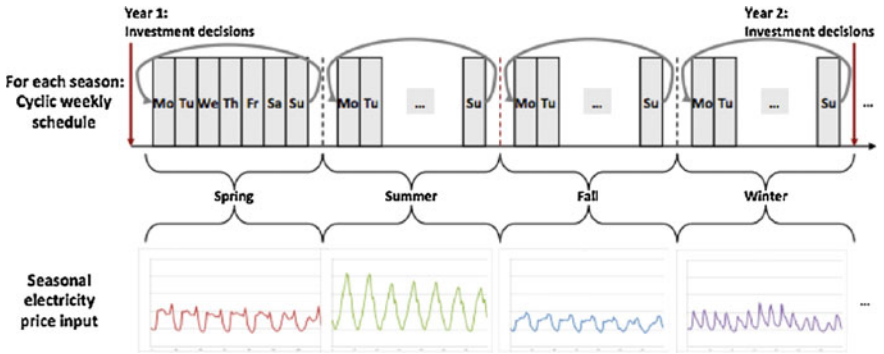

Fig. 18.14 Multiscale time representation of the capacity planning problem with hourly varying electricity prices

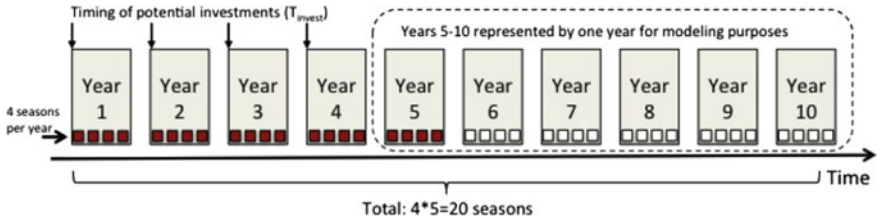

Fig. 18.15 The time horizon of ten years is represented by 5 years with 4 seasons each. Years 5–10 are aggregated into one year in order to maintain a manageable problem size

18.5.2 Case Study

The following investment options are given for the air separation plant in our case study: upgrade the existing liquefier, install an additional liquefier, and install additional storage tanks. We consider three demand scenarios and construct four cases with different baseline demands (i.e. current levels of plant utilization), demand growth rates, probability distributions, and prices for additional product purchase. Case 1 shows low current utilization and low growth with little uncertainty. Case 2 also has a low degree of uncertainty, but a high current utilization and high growth rate. Cases 3 and 4 show medium current utilization as well as medium growth. However, the forecasts have a high degree of uncertainty and the price for external product purchases is higher in Case 4. In each case, the objective is to minimize the total cost consisting of CAPEX and OPEX over the planning horizon of 10 years.

For each case, we would like to understand the value of the current process flexibility (without new investments) and the value of additional flexibility that can be achieved by retrofitting the air separation plant. Therefore, we investigate three setups for both deterministic demand and stochastic demand: (1) constant operation with the currently installed equipment, (2) flexible operation with the currently installed equipment, and (3) joint optimization of investment and operational decisions. The difference in total cost between (1) and (2) is the value of the current flexibility. The difference in total cost between (2) and (3) is the value of additional flexibility. These values may differ for the deterministic demand model and the stochastic demand model.

For deterministic demand, Fig. 18.16 shows the total costs for all four cases. As one can see, the value of current flexibility depends on the plant utilization. If utilization is relatively low, as in Case 1, the plant already has a significant amount of flexibility to react to variability in electricity price. Therefore, the value of current flexibility is equivalent to a reduction in total costs of 13.3 % in Case 1. In contrast, if utilization is very high, e.g. in Case 2, the value of current flexibility to shift production to low-price periods is very limited. Hence, cost savings are relatively small (0.3 %). Consequently, the value of current flexibility is intermediate if the utilization is also within a medium range as one can see in Cases 3 and 4. Only in Case 2, investments are made to increase operational flexibility, driven by the higher demand.

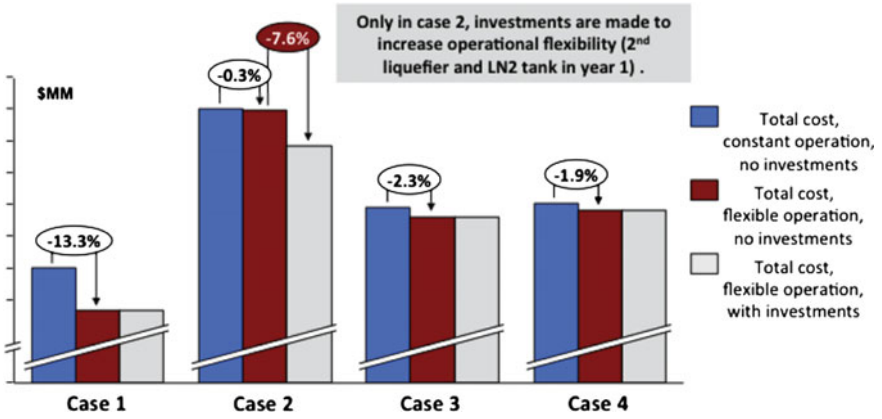

Fig. 18.16 Values of current and additional flexibility for the deterministic demand model

In the first time period, the second liquefier and an additional storage tank for LN2 are installed. The existing liquefier is not replaced, nor were additional storage tanks for LO2 or LAr purchased. The realized cost savings, i.e. the value of additional flexibility, are 7.6 %.

Figure 18.17 shows the total costs for all cases with stochastic demand. As in the deterministic case, the value of existing flexibility is also a function of plant utilization. While the absolute total cost values as well as the relative cost savings are slightly different compared to the deterministic solutions, the overall trend is the same. In Cases 1 and 2, the values of additional flexibility are also similar to the cases with deterministic demand. While there are no investments made in Case 1, the second liquefier and one LN2 tank are installed in the first year in Case 2,

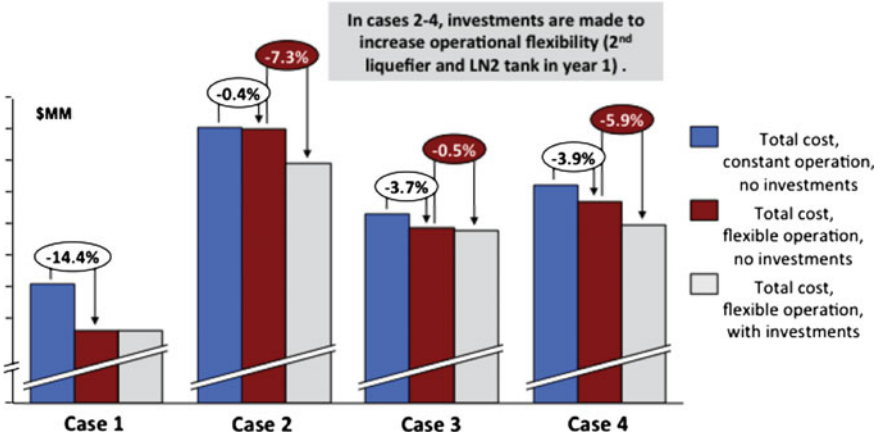

Fig. 18.17 Values of current and additional flexibility for the stochastic demand model

which leads to 7.3 % cost savings. However, in Cases 3 and 4, the stochastic solutions suggest a very different investment strategy compared to the deterministic solutions. In both cases, the second liquefier and one LN2 tank are installed in the first year. The associated values of additional flexibility are 0.5 % and 5.9 %, respectively.

In Cases 3 and 4, the underlying demand distributions have a high standard deviation from the expected value. Therefore, there are more scenarios with higher demand that potentially have higher costs due to external product purchases, if no investments are made. The additional liquefier provides sufficient production capacity to meet the demand in these scenarios, such that costs due to external product purchases are practically eliminated. Furthermore, the combination of the liquefier with the additional LN2 tanks allows a reduction in electricity cost by means of more flexible production. The cost difference is higher in Case 4 than in Case 3; this is mostly due to the skewed distribution, which generates scenarios with higher deviations from the expected demand, and due to the higher price for external product purchases.

18.6 Conclusions

In this chapter, we have addressed the problem of DSM for cryogenic air separation plants. Several systematic decision-making frameworks have been presented that consider various aspects of the problem by means of mathematical optimization approaches. The proposed modeling frameworks include:

- A computationally efficient MILP scheduling model that considers all critical operational constraints as well as various types of power contracts,
- A two-stage stochastic programming model that optimizes the integrated production scheduling and electricity procurement problem with respect to expected profit and risk (measured in terms of the CVaR),
- An affinely adjustable robust optimization model that considers uncertainty in load reduction demand when optimizing the provision of interruptible load,
- A two-stage strategic capacity planning model that relies on a multiscale time representation in which the seasonal behavior of electricity prices is exploited to simplify the model.

We have conducted case studies on real-world air separation plants with data provided by Praxair. The main insights from the case studies can be summarized as follows:

- As long as it does not have to primarily operate at full capacity, an air separation plant has sufficient operational flexibility to perform load shifting, leading to significant cost savings.
- In risk-neutral optimization, accounting for electricity price uncertainty does not result in any notable additional benefit. In contrast, in risk-averse optimization, modeling price uncertainty is crucial for obtaining good solutions.

- Significant financial benefit can be achieved by providing interruptible load; however, incorporating recourse in a robust optimization framework is crucial in order to guarantee feasible load reduction upon request.
- Capital investments in new equipment or retrofit are economically worthwhile if the operational flexibility of the plant increases such that it leads to considerable savings in operating cost.

We close by mentioning that industrial DSM remains an open area of research, with evolving opportunities as electricity markets are changing. It will also require significant additional effort to implement computational tools in practice such that they can be readily used in operations. The gap between research and practice is particularly large in optimization frameworks involving uncertainty, where computational efficiency and interpretation of the results are major challenges.

Acknowledgments The authors gratefully acknowledge the financial support from the National Science Foundation under Grant No. 1159443 and from Praxair.

References

Balas, E. (1985). Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM Journal on Algebraic Discrete Methods*, *6*(3), 466–486.

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust Optimization*. New Jersey: Princeton University Press.

Birge, J. R., & Louveaux, F. (2011). *Introduction to Stochastic Programming* (2nd ed.). Springer Science+Business Media.

Castle, W. F. (2002). Air separation and liquefaction: Recent developments and prospects for the beginning of the new millennium. *International Journal of Refrigeration*, *25*, 158–172.

Castro, P. M., Harjunkoski, I., & Grossmann, I. E. (2009). New continuous-time scheduling formulation for continuous plants under variable electricity cost. *Industrial & Engineering Chemistry Research*, *48*(14), 6701–6714.

Dupacova, J., Gröwe-Kuska, N., & Römisch, W. (2003). Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming Series A*, *95*, 493–511.

EnerNOC. (2014). PJM's Synchronized Reserve Market.

Farhangi, H. (2010). The path of the smart grid. *IEEE Power and Energy Magazine*, *8*(1), 18–28.

GAMS Development Corporation. (2015). GAMS version 24.4.1.

Gellings, C. W., Wikler, G., & Ghosh, D. (2006). Assessment of U.S. electric end-use energy efficiency potential. *Electricity Journal*, *19*(9), 55–69.

Ierapetritou, M. G., Wu, D., Vin, J., Sweeney, P., & Chigirinskiy, M. (2002). Cost minimization in an energy-intensive plant using mathematical programming approaches. *Industrial & Engineering Chemistry Research*, *41*(21), 5262–5277.

Karwan, M. H., & Keblis, M. F. (2007). Operations planning with real time pricing of a primary input. *Computers & Operations Research*, *34*(3), 848–867.

Merkert, L., Harjunkoski, I., Isaksson, A., Säynevirta, S., Saarela, A., & Sand, G. (2014). Scheduling and energy—Industrial challenges and opportunities. *Computers & Chemical Engineering*, *72*, 183–198.

Mitra, S., Grossmann, I. E., Pinto, J. M., & Arora, N. (2012). Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes. *Computers & Chemical Engineering*, *38*, 171–184.

Mitra, S., Sun, L., & Grossmann, I. E. (2013). Optimal scheduling of industrial combined heat and power plants under time-sensitive electricity prices. *Energy*, *54*, 194–211.

Mitra, S., Pinto, J. M., & Grossmann, I. E. (2014). Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. Part I: Modeling. *Computers & Chemical Engineering*, *65*, 89–101.

Paulus, M., & Borggrefe, F. (2011). The potential of demand-side management in energy-intensive industries for electricity markets in Germany. *Applied Energy*, *88*(2), 432–441.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk*, *2*, 21–42.

Samad, T., & Kiliccote, S. (2012). Smart grid technologies and applications for the industrial sector. *Computers & Chemical Engineering*, *47*, 76–84.

Zhang, Q., & Grossmann, I. E. (2016). Planning and Scheduling for Industrial Demand Side Management: Advances and Challenges. In M. Martin (Ed.), *Alternative Energy Sources and Technologies: Process Design and Operation*. Springer.

Zhang, Q., Grossmann, I. E., Heuberger, C. F., Sundaramoorthy, A., & Pinto, J. M. (2015a). Air separation with cryogenic energy storage: Optimal scheduling considering electric energy and reserve markets. *AIChE Journal*, *61*(5), 1547–1558.

Zhang, Q., Grossmann, I. E., Sundaramoorthy, A., & Pinto, J. M. (2015b). Data-driven construction of Convex Region Surrogate models. *Optimization and Engineering*.

Zhang, Q., Cremer, J. L., Grossmann, I. E., Sundaramoorthy, A., & Pinto, J. M. (2016a). Risk-based integrated production scheduling and electricity procurement for continuous power-intensive processes. *Computers & Chemical Engineering*, *86*, 90–105.

Zhang, Q., Morari, M. F., Grossmann, I. E., Sundaramoorthy, A., & Pinto, J. M. (2016b). An adjustable robust optimization approach to provision of interruptible load by continuous industrial processes. *Computers & Chemical Engineering*, *86*, 106–119.

Zhang, Q., Sundaramoorthy, A., Grossmann, I. E., & Pinto, J. M. (2016c). A discrete-time scheduling model for continuous power-intensive process networks with various power contracts. *Computers & Chemical Engineering*, *84*, 382–393.

Chapter 19
Operational and Maintenance Planning of Compressors Networks in Air Separation Plants

Georgios M. Kopanos, Dionysios P. Xenos, Matteo Cicciotti and Nina F. Thornhill

Abstract In this chapter, a general optimizasion-based approach for the integrated operational and maintenance planning of compressor networks in air separation facilities is presented. The proposed mathematical programming model considers operating constraints for compressors, performance degradation for compressors, several types of maintenance policies and other managerial aspects. The operating status, the power consumption, the startup and the shutdown costs for compressors, the compressor-to-header assignments, the timing and the type of necessary maintenance tasks as well as the outlet mass flow rates for compressed air and distillation products are optimized. The power consumption in the compressors is expressed by regression functions that have been derived using technical and historical data. The proposed optimization model can be readily used within a rolling horizon scheme to deal with uncertainty. Several case studies of the air separation plant of BASF SE in Ludwigshafen are solved. The results clearly demonstrate the considerable energy and total cost savings due to the simultaneous planning of operational and maintenance tasks.

Nomenclature

Indices/Sets

$e \in E$	distillation products (e.g., oxygen and nitrogen)
$i \in I$	compressors
$j \in J$	headers
$n \in N$	process plants
$t \in T$	time periods

G.M. Kopanos (✉)
School of Water, Energy and Environment, Cranfield University,
Bedfordshire MK43 0AL, UK
e-mail: g.kopanos@cranfield.ac.uk

D.P. Xenos · N.F. Thornhill
Department of Chemical Engineering, Centre for Process Systems Engineering,
London SW7 2AZ, UK

M. Cicciotti
BASF SE, Advanced Process Control, Automation Technology,
Ludwigshafen 67056, Germany

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_19

$u \in U$ distillation columns
$z \in Z$ storage tanks

Subsets

I^{dm} set of compressors that are subject to fixed maintenance
I^{fm} set of compressors that are subject to flexible time-window maintenance
$\tilde{I}_t^{dm}$ set of compressors that are under maintenance at the beginning of the current planning horizon (maintenance task started in the previous planning horizon)
I_j set of compressors that could serve header j
J_i set of headers that are connected to compressor i
J_n set of headers that are connected to process plant n
J_u set of headers that are connected to distillation column u
U_z set of distillation columns that are connected to storage tank z
Z_e set of storage tanks that can store product e

Superscripts

es earliest
ls latest
max maximum
min minimum

Parameters

α_j coefficient for the load curve of header j
β_j coefficient for the load curve of header j
$\gamma_{(e,u)}$ outlet mass flow rate of product e from distillation column u
$\delta_{(1\ldots3,i)}$ objective function coefficient factors for compressor i
ε_i penalty cost for re-assigning header compressor i during its operation
$\zeta_{(e,t)}$ demand for product e in time period t
η_t maximum number of simultaneous maintenance tasks in time period t
$\theta_{(n,t)}$ compressed air mass flow rate utility demand for process plant n during time period t
κ_t conversion factor of mass flow to aggregated mass amount in time period t (if needed)
λ_j problem-specific large number that could represent the capacity of header j
μ_t electricity price in time period t
ν_i duration of flexible time-window maintenance task in compressor $i \in I^{fm}$
$\xi_{(e,z)}$ storage capacity for product e in storage tank $z \in Z_e$
o_i maximum online time after the startup of compressor i (maximum run time)
π_i outlet compressed air pressure of compressor i
ρ_i outlet compressed air mass flow rate of compressor i
σ_e volumetric percentage of primary component of air e in the composition of air
τ_i starting time for maintenance task in compressor i in period t

ϕ_i shutdown cost for compressor i
χ_i startup cost for compressor i
ψ_i minimum offline time after the shutdown of compressor i (minimum shutdown time)
ω_i minimum online time after the startup of compressor i (minimum run time)
$\tilde{\beta}_{(e,z)}$ initial inventory of product e in storage tank $z \in Z_e$
$\tilde{\eta}_{(i,t)} = 1$ if compressor i is under pre-scheduled maintenance in time period t
$\tilde{\nu}_i$ total time that compressor i has been under maintenance (since the start of the maintenance task) at the end of the previous planning horizon
$\tilde{\varphi}_{(i,j)}$ active connection between compressor i and header j just before the beginning of the current planning horizon
$\tilde{\chi}_i$ operating status of compressor i just before the beginning of the current planning horizon
$\tilde{\psi}_i$ total number of time periods at the end of the past planning horizon that compressor i has been continuously offline since its last shutdown
$\tilde{\omega}_i$ total number of time periods at the end of the past planning horizon that compressor i has been continuously online since its last startup
Ω_i upper bound on the extra power consumption for compressor i
$\bar{\varepsilon}_i$ degradation performance rate of compressor i
$\bar{\rho}_i^{on}$ recovery rate of compressor i that undergoes online washing

Continuous Variables (Non-negative)

$A_{(e,z,t)}$ amount of product e extracted from storage tank $z \in Z_e$ at the end of time period t
$B_{(e,z,t)}$ inventory level of product e in storage tank $z \in Z_e$ at the end of time period t
$C_{(e,u,t)}$ mass flow rate of product e from distillation column u in time period t
$L_{(e,u,z,t)}$ amount of product e from distillation column u that is sent to storage tank $z \in Z_e$ in time period t
$M_{(i,j,t)}$ compressed air mass flow rate from compressor i supplied to header $j \in J_i$ in time period t
$\bar{M}_{(i,j,t)}$ total compressed air mass flow rate supplied to header $j \in J_i$ that is served by compressor i in time period t (auxiliary variable)
$O_{(e,t)}$ amount of product e acquired from external sources at the end of time period t
$P_{(i,j,t)}$ outlet pressure of compressor i that serves header $j \in J_i$ in time period t
$\bar{P}_{(i,t)}$ extra power a compressor i consumes in time period t because of its performance degradation
$\bar{T}_{(i,t)}$ cumulative operating time for compressor i in time period t

Binary Variables

$D_{(i,t)} = 1$ if compressor i changes header from time period $t-1$ to t
$F_{(i,t)} = 1$ if compressor i shutdowns at the beginning of time period t
$S_{(i,t)} = 1$ if compressor i startups at the beginning of time period t
$X_{(i,t)} = 1$ if compressor i is operating during time period t
$Y_{(i,j,t)} = 1$ if compressor i serves header $j \in J_i$ during time period t
$V_{(i,t)} = 1$ if online washing (condition-based maintenance) takes place in compressor i in time period t
$W_{(i,t)} = 1$ if a flexible maintenance task starts in compressor i at the beginning of time period t (also used for offline washing condition-based maintenance)

19.1 Introduction

In most process industries, compressed air, which is provided by compressors, is an indispensable utility for the main production processes. In industrial environments, several compressors are connected in series or in parallel, depending on the purpose of the system into which they are integrated. These networks of compressors can involve a number of compressor units that may differ in the type of drive and technical specifications (e.g., maximum load capacity, efficiency and operational range). Compressors are good targets for energy and cost savings because they are among the most energy-intensive parts of most industrial environments (Saidur et al. 2010; US Department of Energy 2003).

For the transfer of fluids, such as natural gas or ethylene, through long pipelines, several compressors or sub-networks of compressors are placed in series along the pipeline so as to provide the necessary pressure ratio and overcome the pressure drop due to friction losses. Our focus in this chapter is in compressors that are connected in parallel. Networks of parallel compressors are typically used to distribute a fluid material from an upstream process to other downstream processes. Few researchers have addressed the optimization of such networks of compressors. For example, van den Heever and Grossmann (2003) proposed a mathematical programming approach for the production planning and reactive scheduling problem of a hydrogen supply network. Emphasis was placed on the modeling of the pipeline network and not on the operation of the compressors. In particular, minimum run and shutdown times and costs for compressors were ignored. Camponogara et al. (2012) presented a real-time optimization framework for gas-lift compressors in oil fields, neglecting startup and shutdown decisions. Han et al. (2004) studied the optimization of the air and gas-supply network of a chemical plant. The optimization of compressors startup and shutdown actions was not considered. The previous works studied the optimization of network of compressors without emphasizing on their operational aspects, such as startups and shutdowns, and maintenance tasks.

There are several contributions in the literature describing process applications with features similar to those found in the optimal operation of compressors. Among them, Rong and Lahdelma (2005) presented a linear programming planning model and optimization algorithm for trigeneration. Thorin et al. (2005) proposed a mathematical programming model for the long-term planning of cogeneration systems in a competitive market environment. Kopanos et al. (2013) presented an optimization framework for the energy production planning of a network of combined heat and power generators. Zhuan and Xia (2013) proposed a dynamic programming algorithm for the operations scheduling of a pumping station with multiple pumps. Recently, Kopanos and Pistikopoulos (2014) introduced a reactive scheduling rolling horizon framework based on a state-space representation and multiparametric programming. The proposed approach was applied in a network of combined heat and power units. These articles offer a starting point for the work presented in this chapter. However, these papers do not address the simultaneous optimization of maintenance and operational tasks of the compressors, an aspect that is considered here.

This chapter is organized as follows. Section 19.2 provides a brief description of the modeling of the power consumption of multistage centrifugal compressors and the operational planning of cryogenic air separation facilities. The problem statement is formally defined in Sect. 19.3. The operational planning model for the problem of interest is presented in Sect. 19.4, and the maintenance planning model follows in Sect. 19.5. In Sect. 19.7, a number of problem instances, including industrial case studies from the major air separation plant of BASF SE, are solved by the proposed approach, and the results are presented and further discussed. Finally, some concluding remarks along with ongoing research directions are provided in Sect. 19.8.

19.2 Multistage Centrifugal Compressors and Air Separation

This section describes briefly the methodology used to derive the power consumption model of the motors of the compressors as a function of key process parameters and gives a short introduction in cryogenic air separation systems. More details can be found in Kopanos et al. (2015).

19.2.1 Modeling of the Power Consumption in Multistage Centrifugal Compressors

In this chapter, a black box method is used to represent the power consumption of the motor (PM^{el}) as a function of: (i) process variables such as mass flow rates (M) and outlet pressures (P), and (ii) parameters such as the mass flow rate of the cooling water (M^{water}) and the inlet pressure (P_{in}). Thereby, the output power of the motor

can be calculated by an expression of this type:

$$PM^{el}_{(i,j,t)} = f(P_{(i,j,t)}, M_{(i,j,t)}, M^{water}_{(i,t)}, P_{in}, \delta_{(1,i)}, \delta_{(2,i)}, \delta_{(2,i)})$$

This expression can describe compressors with Inlet Guide Vanes (IGVs) or inlet throttling control schemes. In the first case, the operating point of the compressor changes through the manipulation of the angle of the inlet guide vane that controls the outlet pressure. In the latter case, the operating point of the compressor changes by increasing the losses at the inlet of the compressor and in that way the inlet pressure can be modified. The industrial case does not provide the inlet pressure at the eye of the compressor, however the measurement of the inlet pressure provided from the industrial partner is at the filter of the inlet duct of the compressors which is approximately 1 bar for all compressors. This work does not focus on the estimation of the manipulated variables, i.e., angles of IGVs and throttling valves. Therefore, the expression presented above describes well both cases of compressors for the purpose of operational planning, since the variable that is taken into account in the optimization is the outlet pressure P. The efficiencies of the compressors and motors are embedded in the power consumption expression. The reason to use this method is that the efficiencies cannot be explicitly modeled in this case study because many measurements are not available.

It is important to note that the feasible window of operation of the compressor used for the optimization is not defined from the actual physical limits of the compressor, i.e. surge, choke, and minimum and maximum power. The feasible window of operation of the compressor is defined from the domain of the regression model (Brooks et al. 1988). The group of all the operating points of a collected data set derives the regression domain of the model. Therefore, a partial compressor map is captured in the regression models and this leads to a solution in a confined space. Nevertheless, the operation coming from the optimization using this operational space is expected more efficient than operation without optimization. The historical data reveal that the plant operators operated well within the physical limits (surge and choke) of the compressors during the past operation (Xenos et al. 2015).

19.2.2 Operational Planning of Cryogenic Air Separation Plants

The most common air separation method for large-scale production is by means of cryogenic distillation. The cryogenic air separation process starts with the intake of atmospheric air that is usually filtered to remove dust. This air is compressed and re-cooled with the use of an after-cooler. Then, the compressed air is purified through molecular sieves so as to remove carbon dioxide, gaseous hydrocarbons and water vapor. The compressed air is cooled to cryogenic temperatures in the main heat exchanger and then enters a distillation column which separates the air to its primary

components. The products of the distillation can be transferred via pipelines to interconnected industrial sites. For pipeline transfer, products should be vaporized into their gaseous form. For other longer-distance customers, it is more practical and economical to transfer them in a liquid form. In this case, the products of the distillation pass through liquefiers and then they are placed into product-dedicated storage tanks before their shipment by marine or land transportation means. The cryogenic separation of air is an energy-intensive process and the compressors network is one of the main sources of power consumption. Some representative works on the operational planning in air separation plants are discussed below.

Ierapetritou et al. (2002) presented a mixed integer programming formulation for the operational planning in an air separation plant under the objective to minimize the total operating cost. The operation of the plant was described by three different plant operation modes (regular, assisted, shutdown) that vary with respect to operational efficiency and energy requirements. An operating mode defines a convex set that characterizes the feasible operating space of a configuration. Binary variables were used to represent operating modes and switches among the different modes of operation. The model of Ierapetritou et al. (2002) can generate the schedule of process operation modes and production rates. Along the same lines, Karwan and Keblis (2007) proposed a MIP model that additionally considers product losses during configuration changes, while Mitra et al. (2012) presented a mathematical programming model that captures the transient behavior between different operating modes. Following a different approach, Zhu et al. (2011) presented a non-linear programming formulation that has as constraints a process model for the air separation process that accounts for mass and energy balances for the distillation columns, heat exchangers, and throttle valves. The above articles focused on the planning of the operational modes of the overall air separation plant, but they did not address the detailed operational planning of the compressors network. Maintenance tasks for the compressors network were not considered either from the previous works.

This chapter focuses on the detailed planning of operational and maintenance tasks in compressors networks that are part of an air separation facility. We present a general optimization-based approach that considers operating constraints for compressors, performance degradation and recovery for compressors, several types of maintenance policies and other managerial aspects. The operating status, the power consumption, the startup and the shutdown costs for compressors, the compressor-to-header assignments, the timing and the type of necessary maintenance tasks as well as the outlet mass flow rates for compressed air and distillation products are optimized. A rolling horizon modeling approach is adopted to deal with uncertainty, such as demand fluctuations and equipment breakdowns. A formal definition of the problem addressed in this chapter follows in the next section.

19.3 Problem Statement

The overall air separation system under study consists of a finite number of compressors, distillation columns, headers, and storage tanks. This chapter mainly focuses on the integrated operational and maintenance planning of the compressors network which is the major energy-consuming part of cryogenic air separation plants. The downstream air separation process is also considered but in much less detail. Figure 19.1 displays a representative layout of a cryogenic air separation plant. The resulting problem is formally defined in terms of the following items:

- A given planning horizon that is divided into a set of equal-length time periods $t \in T$.
- A set of different compressors $i \in I$ that have a maximum (minimum) outlet compressed air mass flow rate ρ_i^{max} (ρ_i^{min}) and pressure π_i^{max} (π_i^{min}). The minimum online time after the startup of compressor i (minimum runtime) ω_i and the minimum offline time after the shutdown of compressor i (minimum shutdown time) ψ_i, as well as the associated costs for startup (χ_i) and shutdown (ϕ_i) are also given. Compressors operate in parallel.
- A set of distillation columns $u \in U$ that separate air into its primary components $e \in E$. Maximum (minimum) outlet mass flow rates for the products of the distillation columns $\gamma_{(e,u)}^{max}$ ($\gamma_{(e,u)}^{min}$) are given.
- A number of local process plants $n \in N$ that are characterized by a given demand for compressed air for utilities $\theta_{(n,t)}$ per time period t.

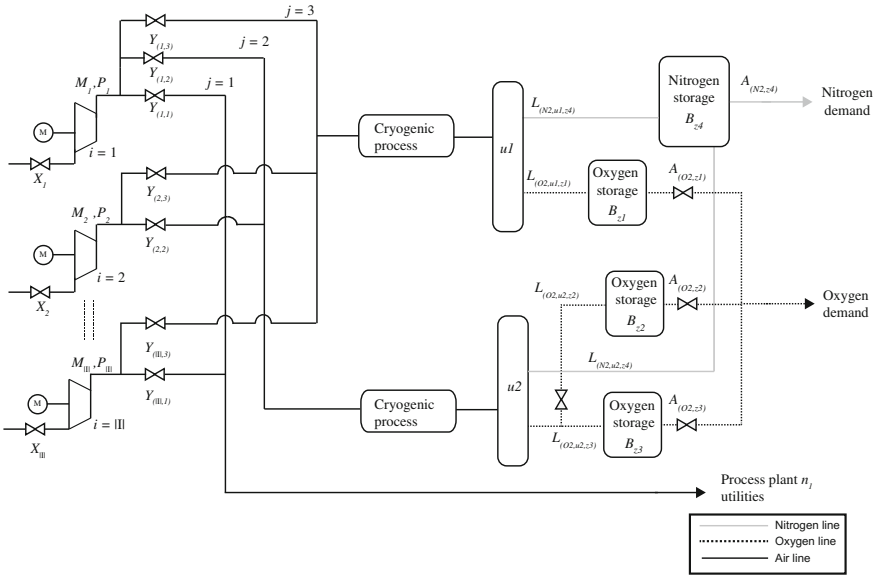

Fig. 19.1 Representative layout of a cryogenic air separation plant

- A set of headers $j \in J$ that receive compressed air from the compressors (J_i). The headers supply with compressed air the distillation columns (J_u) of the air separation network and the local process plants (J_n).
- A set of storage tanks $z \in Z$ which are connected with the outlets of the distillation columns (U_z), and can store specific products (Z_e). Every storage tank has a given maximum (minimum) storage capacity $\xi^{max}_{(e,z)}$ ($\xi^{min}_{(e,z)}$).
- A given demand for the products of the distillation $\zeta_{(e,t)}$ in time period t.
- Products e could be also acquired from external sources in a given cost $v_{(e,t)}$.
- A given electricity price μ_t for time period t, which can be fixed, or time-varying (e.g., following time-of-use rates, or real-time prices).

All parameters are assumed to be deterministic in principal. The discharge pressures of all compressors feeding the same header are assumed the same, which is reasonable because the connected pipes between the exits of the compressors and the main header are short. Each header is connected to a single distillation column or a local process plant (i.e., one destination point). The products of distillation columns are stored in product-dedicated storage tanks through which the demand for these products is satisfied. Each one of the storage tanks could be connected to multiple distillation columns. Also, a compressor can serve at most one header at a time, and during its continuous operation can change header. Generally speaking, the existence of storage tanks allows decoupling the operation of the plant from the demand for products, and gives the required flexibility for the operational planning of the air separation plant.

For every time period t, the key decisions to be made by the administrator of compressors network and the overall air separation plant are:

- The operating status (e.g., startup, in operation, shutdown, under maintenance) of every compressor i;
- The assignment of the operating compressors to headers;
- The compressed air mass flow rate from the operating compressors to each header;
- The total amount of products provided from each distillation column to which storage tanks; and
- The amount of air separation products acquired from external sources.

So as to minimize the total cost, which encompasses startup, shutdown and operating costs of compressors as well as costs for acquiring products from external sources. Maintenance costs are also included for condition-based maintenance tasks. The demand for all products should be fully satisfied.

19.4 Operational Planning Model

In this section, we present a mixed integer programming model for the operational planning of compressor networks in air separation plants. To facilitate the presentation of the proposed model, uppercase Latin letters for optimization variables and sets, and lowercase Greek letters for parameters have been used.

19.4.1 Minimum Run and Shutdown Time for Compressors

Minimum run and shutdown times are modeled by defining the following three sets of binary decision variables:

$$X_{(i,t)} = \begin{cases} 1, & \text{if compressor } i \text{ is in operation during time period } t \\ 0, & \text{otherwise.} \end{cases}$$

$$S_{(i,t)} = \begin{cases} 1, & \text{if compressor } i \text{ starts up at the beginning of time period } t \\ 0, & \text{otherwise.} \end{cases}$$

$$F_{(i,t)} = \begin{cases} 1, & \text{if compressor } i \text{ shuts down at the beginning of time period } t \\ 0, & \text{otherwise.} \end{cases}$$

Constraints (19.1) and (19.2) define $S_{(i,t)}$ and $F_{(i,t)}$ binary variables through binary variables $X_{(i,t)}$. Parameters $\tilde{\chi}_i$ denote the operating status (i.e., on or off) of compressor i just before the beginning of the current planning horizon and they can be calculated through parameters $\tilde{\varphi}_{(i,j)}$, which represent the active connection of compressor i to header j before the beginning of the current planning horizon. Since a compressor can supply compressed air to at most one header at a time: $\tilde{\chi}_i = \sum_{j \in J_i} \tilde{\varphi}_{(i,j)}$.

$$\begin{aligned} S_{(i,t)} - F_{(i,t)} &= X_{(i,t)} - \tilde{\chi}_i \qquad \forall i \in I, t \in T : t = 1 \\ S_{(i,t)} - F_{(i,t)} &= X_{(i,t)} - X_{(i,t-1)} \qquad \forall i \in I, t \in T : t > 1 \end{aligned} \tag{19.1}$$

$$S_{(i,t)} + F_{(i,t)} \leq 1 \qquad \forall i \in I, t \in T \tag{19.2}$$

Constraints (19.2) could be omitted, if startup and shutdown costs are part of the objective function. For every compressor, constraints (19.3) and (19.4) model the minimum run and shutdown time, respectively.

$$X_{(i,t)} \geq \sum_{t'=max\{1,t-\omega_i+1\}}^{t} S_{(i,t')} \qquad \forall i \in I, t \in T : \omega_i > 1 \tag{19.3}$$

$$1 - X_{(i,t)} \geq \sum_{t'=max\{1,t-\psi_i+1\}}^{t} F_{(i,t')} \qquad \forall i \in I, t \in T : \psi_i > 1 \tag{19.4}$$

19.4.2 Assignment of Compressors to Headers

An additional binary variable is introduced to denote the active connection of available compressors to headers:

$$Y_{(i,j,t)} = \begin{cases} 1, & \text{if compressor } i \text{ serves header } j \in J_i \text{ in time period t} \\ 0, & \text{otherwise.} \end{cases}$$

A compressor i can supply compressed air to at most one header $j \in J_i$ at a time, according to:

$$\sum_{j \in J_i} Y_{(i,j,t)} = X_{(i,t)} \qquad \forall i \in I, t \in T \tag{19.5}$$

19.4.3 Compressor-to-Header Assignment Changes

Any compressor i during its operation can change headers from one time period to another. Due to this fact, solutions that are characterized by many (and sometimes unnecessary) header changes may be obtained. In practice, compressor-to-header re-assignments are typically done manually through valve opening or closing by the technical personnel of the industry. For this reason, a low number of header changes is more desirable so as to: (i) reduce the utilization of personnel; (ii) decrease the probability for potential human errors; (iii) favor a more smooth operation (i.e., easier to implement, control, and revise) of the overall network of compressors; and (iv) avoid energy losses during the header change due to venting of the compressed air to the atmosphere. In order to optimize the total number of these compressor-to-header changes, we incorporated an associated penalty cost term (ε_t) in the objective function. To model compressor-to-header changes, the following set of binary variables has been introduced:

$$D_{(i,t)} = \begin{cases} 1, & \text{if compressor } i \text{ changes header from time period } t-1 \text{ to } t \\ 0, & \text{otherwise.} \end{cases}$$

In that way, the compressor-to-header changes could be modeled by:

$$\begin{aligned} D_{(i,t)} &\geq Y_{(i,j,t)} - \tilde{\varphi}_{(i,j)} - S_{(i,t)} \qquad \forall i \in I, j \in J_i, t \in T : t = 1 \\ D_{(i,t)} &\geq Y_{(i,j,t)} - Y_{(i,j,t-1)} - S_{(i,t)} \qquad \forall i \in I, j \in J_i, t \in T : t > 1 \end{aligned} \tag{19.6}$$

Due to the introduction of a penalty cost for header changes in the objective function, binary variables $D_{(i,t)}$ tend to zero.

19.4.4 Feasible Windows of Operation for Compressors

For any compressor i, constraints (19.7) and (19.8) denote the lower and upper bounds on the outlet compressed air mass flow rate and outlet pressure that the compressor could operate:

$$\rho_i^{min} Y_{(i,j,t)} \leq M_{(i,j,t)} \leq \rho_i^{max} Y_{(i,j,t)} \qquad \forall i \in I, j \in J_i, t \in T \tag{19.7}$$

$$\pi_i^{min} Y_{(i,j,t)} \leq P_{(i,j,t)} \leq \pi_i^{max} Y_{(i,j,t)} \qquad \forall i \in I, j \in J_i, t \in T \tag{19.8}$$

From constraints (19.8), if $Y_{(i,j,t)} = 0$ then $P_{(i,j,t)} = 0$. This zero pressure has no physical meaning, and actually it is just used for the calculation of the power consumption cost term in the objective function.

19.4.5 Demand for Compressed Air for Utilities

The network of compressors should meet the demand for compressed air for utilities $\theta_{(n,t)}$ for every process plant n connected to the compressors network in each time period t, according to:

$$\sum_{i \in I} \sum_{j \in (J_n \cap J_i)} M_{(i,j,t)} \geq \theta_{(n,t)} \qquad \forall n \in N, t \in T \tag{19.9}$$

Compressed air for utilities cannot be stored.

19.4.6 Downstream Process: Load Curve

The load curve of the downstream process, header and distillation column, represents the pressure drop of the header and the operational pressure at the inlet of the column. A linear relationship between mass flow, which is equal to the summation of the individual mass flows of the compressors operating to that header, provided in the header and outlet pressure of the compressor can represent the load curve by applying regression to available historical data. The outlet pressure of an operation compressor is then given by the corresponding load curve of the header that serves, according to:

$$P_{(i,j,t)} = \alpha_j \bar{M}_{(i,j,t)} + \beta_j Y_{(i,j,t)} \qquad \forall i \in I, j \in J_i, t \in T \tag{19.10}$$

Parameters α_j and β_j represent the coefficients for the load curve of header j. In order to avoid non-linearities in the load curve constraints (19.10), we have introduced auxiliary variables $\bar{M}_{(i,j,t)}$ that denote the total compressed air mass flow rate in the

header j which is served by compressor i during time period t. Non-negative variables $\bar{M}_{(i,j,t)}$ are modeled through the following set of big-M constraints:

$$\begin{aligned}
\bar{M}_{(i,j,t)} &\geq \sum_{i' \in I_j} M_{(i',j,t)} - \lambda_j (1 - Y_{(i,j,t)}) \qquad \forall i \in I, j \in J_i, t \in T \\
\bar{M}_{(i,j,t)} &\leq \sum_{i' \in I_j} M_{(i',j,t)} + \lambda_j (1 - Y_{(i,j,t)}) \qquad \forall i \in I, j \in J_i, t \in T \\
\bar{M}_{(i,j,t)} &\leq \lambda_j Y_{(i,j,t)} \qquad \forall i \in I, j \in J_i, t \in T
\end{aligned} \tag{19.11}$$

Parameters λ_j are problem-specific large numbers that can be usually calculated by considering the maximum capacity of header j. The load curves of the headers are estimated from industrial process data.

19.4.7 Downstream Process: Distillation Columns

The outlet mass flow rate of product e from each distillation column u during each time period t is given by:

$$C_{(e,u,t)} = \sigma_e \sum_{i \in I} \sum_{j \in (J_u \cap J_i)} M_{(i,j,t)} \qquad \forall u \in U, t \in T \tag{19.12}$$

Constraints (19.13) provide lower and upper bounds on the outlet mass flow rate of product e from each distillation column.

$$\gamma^{min}_{(e,u)} \leq C_{(e,u,t)} \leq \gamma^{max}_{(e,u)} \qquad \forall e \in E, u \in U, t \in T \tag{19.13}$$

19.4.8 Downstream Process: Mass Balances and Inventory Levels

The total amount of products e from the distillation columns u that is disposed to each storage tank $z \in Z_e$ during time period t is represented by variables $L_{(e,u,z,t)}$, and is given by:

$$\sum_{z \in Z_e} L_{(e,u,z,t)} = \kappa_t C_{(e,u,t)} \qquad \forall e \in E, u \in U, t \in T \tag{19.14}$$

Hence, the demand for any product e in each time period t is met by extracting the necessary amount of product ($A_{(e,z,t)}$) from the storage tanks $z \in Z_e$. The products could be also acquired from external sources ($O_{(e,t)}$), if the demand cannot be fully met by the internal production network.

$$\sum_{z \in Z_e} A_{(e,z,t)} + O_{(e,t)} = \zeta_{(e,t)} \qquad \forall e \in E, t \in T \tag{19.15}$$

Constraints (19.16) correspond to the mass balance in the product storage tanks under the complete satisfaction of the demand for products. Variables $B_{(e,z,t)}$ denote the amount of product e that is stored in storage tank $z \in Z_e$ at the end of time period t. Parameter $\tilde{\beta}_{(e,z)}$ represents the initial inventory of product e in storage tank $z \in Z_e$.

$$\begin{aligned} B_{(e,z,t)} &= \tilde{\beta}_{(e,z)} + \sum_{u \in U_z} L_{(e,u,z,t)} - A_{(e,z,t)} \qquad \forall e \in E, z \in Z_e, t \in T : t = 1 \\ B_{(e,z,t)} &= B_{(e,z,t-1)} + \sum_{u \in U_z} L_{(e,u,z,t)} - A_{(e,z,t)} \qquad \forall e \in E, z \in Z_e, t \in T : t > 1 \end{aligned} \tag{19.16}$$

Constraints (19.17) define the lower and the upper inventory bounds. The upper inventory bound represents the storage capacity of the storage tank. Safety stocks could be represented by the lower inventory bound.

$$\xi^{min}_{(e,z)} \leq B_{(e,z,t)} \leq \xi^{max}_{(e,z)} \qquad \forall e \in E, z \in Z_e, t \in T \tag{19.17}$$

The proposed optimization framework can be also used to cope with planning problems in multiple (i.e., multi-site) networks of compressors. The presented model can be directly applied in the multi-site case, if no interchange of products is allowed between the storage tanks of the different sites. Otherwise, the mass balance equations should be accordingly modified.

19.4.9 Initial State of the Compressors Network

In order to obtain feasible and optimal operational and maintenance plans the initial state of the compressors network must be identified and considered in the optimization. The initial inventory level of every storage tank $\tilde{\beta}_{(e,z)}$ (constraints (19.16)) and the initial operating status of every compressor $\tilde{\varphi}_{(i,j)}$ ($\rightarrow \tilde{\chi}_i$) (constraints (19.1)) partially describe the initial state of the network. Also needed is information on the startup and shutdown from the previous planning horizon. The following set of constraints allows past information regarding startup and shutdown to be carried over at the beginning of the current planning horizon:

$$X_{(i,t)} = 1 \qquad \forall i \in I, t = 1, \ldots, (\omega_i - \tilde{\omega}_i) : 0 < \tilde{\omega}_i < \omega_i \tag{19.18}$$

$$X_{(i,t)} = 0 \qquad \forall i \in I, t = 1, \ldots, (\psi_i - \tilde{\psi}_i) : 0 < \tilde{\psi}_i < \psi_i \tag{19.19}$$

Parameters $\tilde{\omega}_i$ ($\tilde{\psi}_i$) denote the total number of time periods at the end of the past planning horizon that compressor i has been operating (not operating) since its last

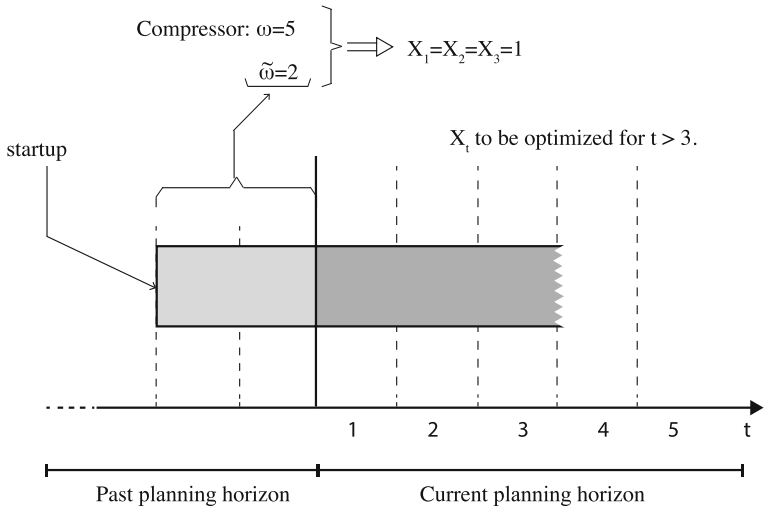

Fig. 19.2 Carryover of past startup information to model minimum run

startup (shutdown). Figure 19.2 shows an illustrative example of how the past startup information is captured from the above constraints.

19.4.10 Objective Function

The optimization goal in this study is to minimize the total startup, shutdown and power consumption costs of the compressors as well as the procurement cost of products from external sources. A compressor-to-header changes penalty cost term is also included. As already mentioned in Sect. 19.2.1, power consumption is a function of the outlet mass flow rate, the outlet pressure, the consumption of cooling water, and the ambient conditions. This function is usually represented well by a quadratic form, however a more simple linear form could be derived in some cases. The advantage of a linear objective function is the significant reduced computational effort to reach optimality, however a higher error in the results is usually expected. The linear total cost objective function considered here is given by:

$$\begin{aligned}\sum_{t\in T}\sum_{i\in I}(\varepsilon_i D_{(i,t)}) + \sum_{t\in T}\sum_{e\in E}(\upsilon_{(e,t)} O_{(e,t)}) + \sum_{t\in T}\sum_{i\in I}(\chi_i S_{(i,t)} + \phi_i F_{(i,t)}) \\ + \sum_{t\in T}\mu_t \sum_{i\in I}\sum_{j\in J_i}(\delta_{(1,i)} Y_{(i,j,t)} + \delta_{(2,i)} M_{(i,j,t)} + \delta_{(3,i)} P_{(i,j,t)})\end{aligned} \tag{19.20}$$

Parameters $\delta_{(1\ldots3,i)}$ correspond to power consumption normalized coefficient factors for every compressor i, and mainly depend on information related to cooling water consumption and ambient conditions.

19.5 Maintenance Planning Model

In this part, we present the maintenance planning model of the proposed optimization framework. In most relevant industrial environments, the maintenance plan of compressors is typically predefined before the operational planning of the compressors network. The duration ν_i of a maintenance task is generally known. Here, we consider five types of maintenance tasks: (i) runtime-based, (ii) fixed, (iii) flexible time-window, (iv) condition-based via online washing, and (v) condition-based via offline washing. The first type of maintenance tasks is related to the maximum runtime of the compressors. Fixed maintenance tasks take place in a predefined time horizon, such that the starting (completion) times of the maintenance tasks are known, while flexible time-window maintenance tasks generally take place within a predefined time window, and their exact starting (completion) times are additional decisions to be made. For the condition-based maintenance tasks, we consider a degradation and a recovery model for the performance of the compressor with respect to the cumulative time of its operation. The realization for these types of maintenance tasks depend on the performance of the compressor through its operation and they can be performed via online and offline washing. The description and the modeling of all maintenance tasks considered follows.

19.5.1 Runtime-Based Maintenance Tasks

The performance of a compressor depends on its current condition. Since the performance of the compressor deteriorates during its utilization, there is often a time limit on its continuous operation. After this so-called maximum runtime (o_i), the compressor should be switched off, and maintenance takes place to prevent mechanical damage and the energy-inefficient use of the compressor. The runtime-based maintenance policy can be considered as a simplified case of the condition-based maintenance policy (Li and Nilkitsaranont 2009). For the sake of simplicity here, the duration of these types of maintenance tasks are considered to be equal to the minimum shutdown time. For every compressor, the maximum runtime constraints are given by:

$$\sum_{t'=max\{1,t-o_i\}}^{t} X_{(i,t')} \leq o_i \qquad \forall i \in I, t \in T \tag{19.21}$$

Notice that if compressor i has been in operation at the end of the previous planning horizon (i.e., $\tilde{\omega}_i > 1$), then the total runtime from its last startup in the past horizon should be carried over to the current planning horizon so as to model successfully the maximum runtime limits. Constraints (19.22) describe the initial state of each compressor of the network in respect of its corresponding maximum runtime.

$$\sum_{t'=max\{1,t-(o_i-\tilde{\omega}_i)\}}^{t} X_{(i,t')} \leq (o_i - \tilde{\omega}_i) \qquad \forall i \in I, t = (o_i - \tilde{\omega}_i + 1) : \tilde{\omega}_i > 1 \quad (19.22)$$

If there is a maximum time that compressors could remain idle (i.e., maximum idle time), similar types of constraints could be easily derived.

19.5.2 Fixed Maintenance Tasks

The starting times τ_i of this type of maintenance tasks are known. For this reason, fixed maintenance tasks can be readily modeled by setting to zero (from the starting to the completion of the maintenance task) the operating binary variables $X_{(i,t)}$ of compressors $i \in I^{dm} \subseteq I$ that are subject to this type of maintenance:

$$X_{(i,t)} = 0 \qquad \forall i \in I^{dm}, t = \tau_i, \ldots, (\tau_i + \nu_i - 1)$$

These constraints generally hold, since the maintenance duration is typically at least equal to the minimum shutdown time.

Due to the presence of maintenance tasks, the initial maintenance state (i.e., under maintenance or not) of every compressor should be carried over from the previous planning horizon. To consider this properly, for the compressors $i \in \tilde{I}_t^{dm}$ that are under maintenance at the beginning of the current planning horizon and their maintenance tasks had already started in the previous planning horizon as well as for the compressors $i \in I^{dm}$ that are under fixed maintenance in the current planning horizon, the following set of constraints is included:

$$\begin{aligned}
\tilde{\eta}_{(i,t)} &= 1 \qquad \forall i \in (I^{dm} \setminus \tilde{I}^{dm}), t = \tau_i, \ldots, (\tau_i + \nu_i - 1) \\
X_{(i,t)} &= 1 - \tilde{\eta}_{(i,t)} \qquad \forall i \in (I^{dm} \setminus \tilde{I}^{dm}), t = \tau_i, \ldots, (\tau_i + \nu_i - 1) \\
\tilde{\eta}_{(i,t)} &= 1 \qquad \forall i \in \tilde{I}^{dm}, t = 1, \ldots, (\nu_i - \tilde{\nu}_i) \\
X_{(i,t)} &= 1 - \tilde{\eta}_{(i,t)} \qquad \forall i \in \tilde{I}^{dm}, t = 1, \ldots, (\nu_i - \tilde{\nu}_i)
\end{aligned} \quad (19.23)$$

Parameter $\tilde{\nu}_i$ denotes the total time that compressor i has been under maintenance (since the start of the maintenance task) at the end of the previous planning horizon. In other words, $(\nu_i - \tilde{\nu}_i)$ represents the remaining time that compressor i has to be under maintenance in the current planning horizon. Parameter $\tilde{\eta}_{(i,t)}$ denotes if compressor i is under fixed maintenance in time period t.

19.5.3 Flexible Maintenance Tasks Within a Given Time-Window

To model these types of maintenance tasks, the following set of binary variables has been introduced:

$$W_{(i,t)} = \begin{cases} 1, & \text{if the maintenance task in compressor } i \text{ begins in time period } t \\ 0, & \text{otherwise.} \end{cases}$$

For compressors $i \in I^{fm} \subseteq I$ that are subject to flexible maintenance, the corresponding maintenance tasks should start within a given time window $t \in [\tau_i^{es}, \tau_i^{ls}] \subseteq T$:

$$\sum_{t=\tau_i^{es}}^{\tau_i^{ls}} W_{(i,t)} = 1 \qquad \forall i \in I^{fm} \tag{19.24}$$

Notice that the latest starting time plus the duration of the maintenance task has to be lower or equal to the last time period of the planning horizon of interest (i.e., $(\tau_i^{ls} + \nu_i) \leq |T|$) in order to ensure that the maintenance task will be completed within the given planning horizon. Hence, constraints (19.25) ensure the unavailability of compressor i throughout the duration of the flexible maintenance task.

$$X_{(i,t)} + \sum_{t'=max\{\tau_i^{es}, t-\nu_i+1\}}^{min\{\tau_i^{ls}, t\}} W_{(i,t')} \leq 1 \qquad \forall i \in I^{fm}, t = \tau_i^{es}, \ldots, (\tau_i^{ls} + \nu_i - 1) \tag{19.25}$$

Fixed maintenance tasks can be modeled as flexible maintenance tasks, if one includes them in set I^{fm}, and then defines $\tau_i^{es} = \tau_i$ and $\tau_i^{ls} = (\tau_i + \nu_i - 1)$. Also, note that an incomplete flexible maintenance task that started in the previous planning horizon, in the current planning horizon is a fixed maintenance task that can be modeled by constraints (19.23).

19.5.3.1 Multiple Simultaneous Maintenance Tasks

In most industrial applications, the simultaneous maintenance of multiple compressors is expected to be undesired due to managerial or technical reasons. From a managerial point of view, having several compressors under maintenance at the same time significantly limits the operational flexibility in the presence of unexpected events (i.e., the total load capacity of the compressors network decreases). Moreover, maintenance resources may be limited, and therefore a maximum number of compressors (η_t) may be maintained simultaneously in every time period. These organizational aspects have been modeled by the following constraints:

$$\sum_{i\in(I^{dm}\cup\tilde{I}^{dm})}\tilde{\eta}_{(i,t)}+\sum_{\substack{i\in I^{fm}\\ \tau_i^{es}\le t\le(\tau_i^{ls}+\nu_i-1)}}\sum_{t'=max\{\tau_i^{es},t-\nu_i+1\}}^{min\{\tau_i^{ls},t\}}W_{(i,t')}\le\eta_t\qquad\forall t\in T\qquad(19.26)$$

19.5.4 Condition-Based Maintenance Tasks

The compressors that are placed in process gas applications mainly suffer from fouling (Aretakis et al. 2012), which is the depositions of particles of the fluid to the airfoil. These depositions increase the roughness of the surfaces of the internal mechanical components of the compressors (e.g., impeller and diffuser area) and restrict the passages areas of the fluid. The result of the fouling is a decrease in performance and an increase in power consumption for the same load compared to a non-fouled compressor.

There are two common strategies to deal with fouling, namely offline and online washing. The offline washing takes place when a compressor is not operating while the online washing cleans the compressor without interrupting its operation. Offline washing maintenance is completed after several cleaning steps while supplementary maintenance tasks, such as mechanical and electrical inspections, could be performed as well. For this reason, the total duration of this type of washing could increase up to a few days. During the online washing, a cleaning solution is injected inside the compressor while it is under operation. The advantage of this maintenance type is that less power is consumed without shutting down the compressor. However, the recovery of the performance is considerably smaller than that of the offline washing maintenance. Figure 19.3 displays representative trends of the efficiency of a compressor considering different types of washing methods.

To model online washing maintenance tasks, the following set of binary variables has been introduced:

$$V_{(i,t)}=\begin{cases}1, & \text{if there is an online washing in compressor } i \text{ in time period } t\\ 0, & \text{otherwise.}\end{cases}$$

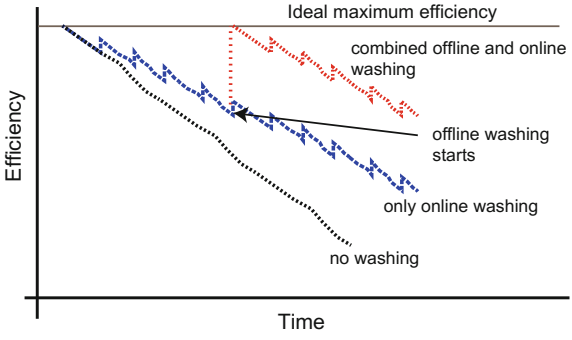

Fig. 19.3 Representative efficiency profiles for different types of washing methods

The duration of any online washing maintenance task is considered equal to a single time period. Having defined binary variables $V_{(i,t)}$ and by introducing the cumulative operating time variables $\bar{T}_{(i,t)}$, we can model the recovery of the performance of a compressor as a combination of the two different maintenance washing modes, according to:

$$\bar{T}_{(i,t)} = (\bar{T}_{(i,t-1)} + X_{(i,t)})\,(1 - W_{(i,t)})\,(1 - \bar{\rho}_i^{on} V_{(i,t)}) \qquad \forall i \in I, t \in I \tag{19.27}$$

The initial value (i.e., initial state) of the cumulative time variable for every compressor at the beginning of the prediction horizon is represented by $\bar{T}_{(i,0)}$.

$$\bar{P}_{(i,t)} = \bar{\varepsilon}_i \cdot \bar{T}_{(i,t)} \cdot X_{(i,t)} \qquad \forall i \in I, t \in T \tag{19.28}$$

$$\bar{P}_{(i,t)} \leq \Omega_i \qquad \forall i \in I, t \in T \tag{19.29}$$

Constraints (19.28) give the expression of the extra power a compressor i consumes because of its performance degradation. Here, we assume that the performance degradation is a linear function of the cumulative time $\bar{T}_{(i,t)}$, the degradation rate $\bar{\varepsilon}_i$ of the compressor and the operating binary variable $X_{(i,t)}$. Obviously, there is no extra power consumption if the compressor is not operating (i.e., $X_{(i,t)} = 0$). A way to linearize constraints (19.28) has been presented in Xenos et al. (2016). Constraints (19.29) give the maximum extra power consumption permitted $\Omega_{(i)}$ for every compressor due to performance degradation. Since we have assumed a linear degradation profile, this parameter can be derived from the relationship of the extra power and cumulative days, $\Omega_i = \bar{\varepsilon}_i o_i$, where o_i is the maximum runtime.

19.5.4.1 Remaining Constraints and Considerations

The injection of a cleaning solution into the compressor for performing online maintenance could have undesired side effects, such as corrosion of the blades, if done too frequently. For this reason, we consider a parameter for each compressor that represents the minimum number of time periods between two consecutive online washings, and we include corresponding constraints in our overall optimization framework that models this practical restriction. Parameters that define the initial state with respect to the minimum number of time periods between two consecutive online washings have been also introduced. In addition, our model includes constraints that: (i) do not allow an online maintenance task to take place if a compressor is not in operation, and (ii) limit the total number of online washings per compressor for the given prediction horizon. Finally, the associated cost terms for online and offline washing tasks must be included in the objective function, given by expression (19.20).

The main assumptions of the degradation and recovery model presented in this chapter are summarized below:

- The degradation rate of a compressor depends on the type of the compressor and on the cumulative time of operation after the last maintenance. A linear function between extra power consumption due to degradation and cumulative operational time is considered. In reality, this assumption holds true up to a particular level of fouling.
- If a compressor is switched off and remains offline without maintenance the condition of the compressor stays as it was before the shutdown. In reality, corrosion may continue between the fouling deposits and the blades, also during stand-by of the machine.
- If a compressor is washed offline, its efficiency is fully recovered. In other words, the extra power consumption due to degradation becomes zero immediately after the offline washing.
- The additional power consumed due to degradation is assumed constant during the time period t, and it does not depend on the operating point.

A detailed description and discussion of the additional constraints and these considerations has been provided by Xenos et al. (2016).

19.6 Rolling Horizon Approach

The proposed mathematical model has been formulated in such a way so as to be used within a rolling horizon framework, if needed. Figure 19.4 displays a representative algorithm for rolling horizon optimization. In brief, in the rolling horizon scheme, an optimization problem is solved for a planning horizon of certain length (i.e., prediction horizon), and then the solution for just a part of that planning horizon (i.e., control horizon) is applied. A number of optimization problems is solved iteratively by moving forward the time horizon after each iteration. The key part of any rolling horizon approach is to update properly the current state of the overall system before the optimization of the prediction horizon of interest. For a detailed discussion on issues regarding the scheduling via rolling horizon approaches kindly refer to Kopanos and Pistikopoulos (2014).

19.7 Case Studies

The air separation facility of BASF SE in Ludwigshafen Germany is considered in all case studies. The air compressor network of the facility consists of eleven multistage centrifugal compressors connected in parallel that supply three headers with compressed air. The first header $j1$ collects the compressed air for utilities in the chemical complex of BASF SE. The other two headers $j2$ and $j3$ are connected with two distillation columns $u1$ and $u2$, respectively. There are five small compressors with throttling valves ($i \in I_s = \{i1, i2, i3, i4, i5\}$) and six large compressors

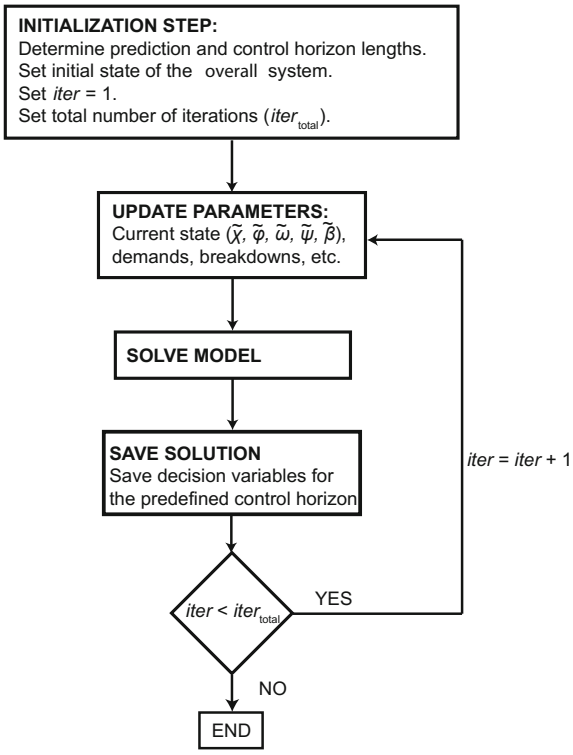

Fig. 19.4 A representative rolling horizon algorithm

with IGVs ($i \in I_b = \{i6, i7, i8, i9, i10, i11\}$). The main operating data for these compressors have been derived through the thorough analysis of historical data and can be seen in Fig. 19.5. The costs for startup and shutdown for the large compressors are approximately 60 % higher than those for the small compressors. Oxygen and compressed air for utilities are considered as the main products, and they cannot be acquired from external sources.

In this section, three industrial case studies are presented. The first case study considers deterministic demand profiles for products. The second case study deals with the reactive planning problem, where the demand profiles for products may vary from day to day, and therefore demand predictions are updated after every day period. In contrast to the first two case studies, the last case study considers explicitly the performance degradation and recovery of the compressors and optimizes the plan for online and offline maintenance tasks (i.e., a condition-based like approach). A total planning horizon of 30 days, divided into one-day periods, is studied in all case studies. Several maintenance policies are considered in order to highlight the important benefits of the simultaneous optimization of operational and maintenance tasks.

All given data and reported results are normalized and made dimensionless due to confidentiality reasons. All optimization problems have been solved in

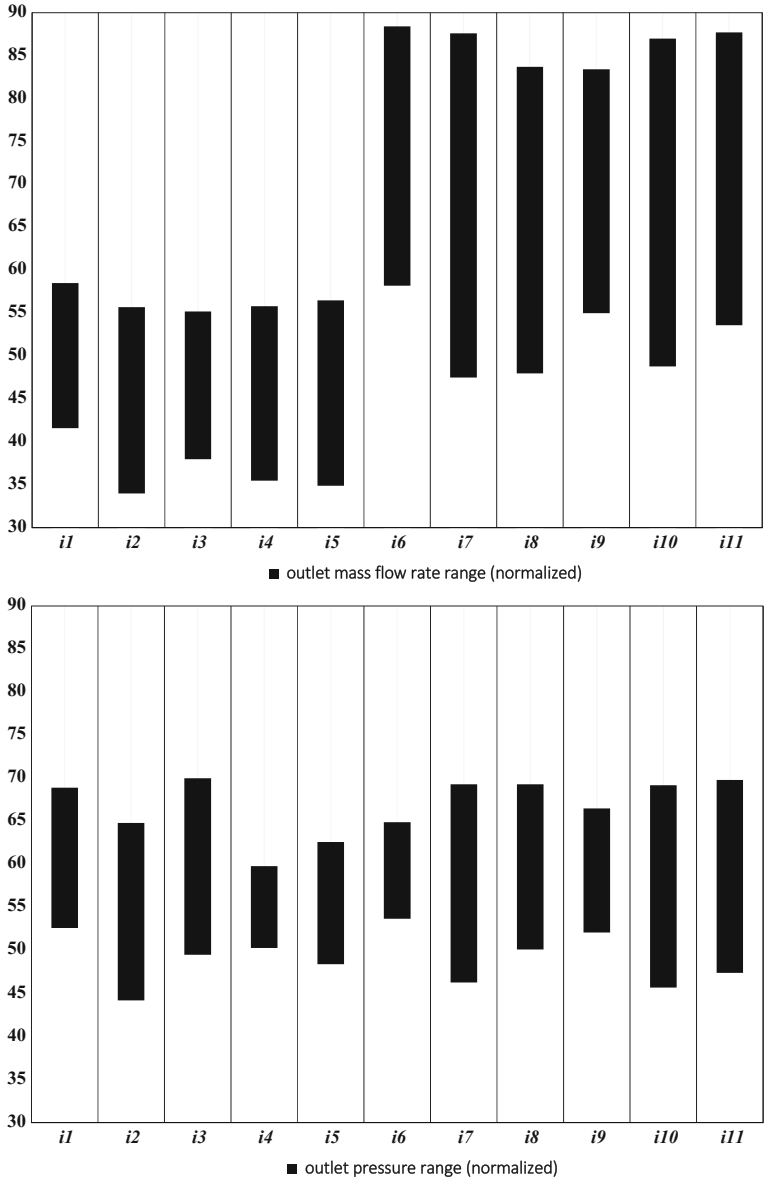

Fig. 19.5 Normalized operating bounds for compressors

GAMS/CPLEX 11.1, under default configurations, in an Intel(R) Core(TM) i7-2600CPU @3.4 GHz with 8 GB RAM. A zero optimality gap has been imposed in all problems instances. Problems instances have been solved to zero optimality gaps, if not otherwise stated.

Table 19.1 Case Study 1: information of maintenance tasks for the compressors

Problem instance	Parameter	$i1$	$i2$	$i3$	$i4$	$i5$	$i6$	$i7$	$i8$	$i9$	$i10$	$i11$
Problem 1.1	τ_i	26	–	–	7	15	2	–	–	22	–	–
Problem 1.1–1.3	ν_i	5	–	–	3	5	3	–	–	3	–	–
Problem 1.2–1.3	τ_i^{es}	5	–	–	5	5	5	–	–	5	–	–
Problem 1.2–1.3	τ_i^{ls}	25	–	–	25	25	25	–	–	25	–	–

19.7.1 Case Study 1: Deterministic Planning

Here, the proposed optimization-based framework is applied in a deterministic case study of the air separation plant of BASF SE. Three different planning policies for maintenance are considered: (i) fixed maintenance—Problem 1.1; (ii) flexible time-window maintenance through the simultaneous planning of operational and maintenance tasks—Problem 1.2; and (iii) flexible time-window maintenance considering maintenance resources limitations—Problem 1.3. Table 19.1 provides the given data for the maintenance policy for each problem instance. The normalized demand profiles for oxygen and compressed air for utilities are given in Fig. 19.6. The cost for changing the compressor-to-header assignment is equal to 250 m.u./change and purchases of products are not allowed. Small and large compressors could operate for maximum 20 and 60 consecutive days respectively, apart from $i5$ which can operate for a maximum of 50 consecutive days. Large compressors have a minimum runtime of six days while $i1$, $i2$, $i3$, $i4$, $i5$ have minimum run times equal to five, six, seven, six and five days, respectively.

All the related past data that fully describe the initial state of each compressor can be found in Table 19.2. More specifically, Table 19.2 contains information regarding: (i) the active compressor-to-header connection ($\tilde{\phi}_{(i,j)}$); (ii) the total duration that each compressor has been operating from its last startup ($\tilde{\omega}_i$); and (iii) the total duration that each compressor has not been operating since its last shutdown ($\tilde{\psi}_i$). For instance, at the beginning of the planning horizon, compressor $i2$ is connected to

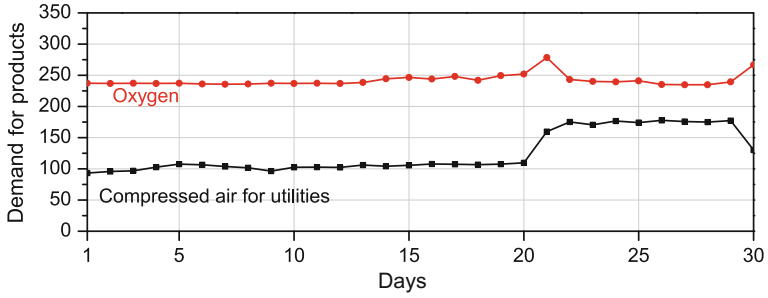

Fig. 19.6 Case Study 1: normalized demand for products

Table 19.2 Case Study 1: initial condition for all compressors (i.e., $t = 0$)

Parameter	$i1$	$i2$	$i3$	$i4$	$i5$	$i6$	$i7$	$i8$	$i9$	$i10$	$i11$
$\tilde{\varphi} = 1$	–	$j1$	–	–	$j2$	$j2$	–	$j1$	–	$j3$	$j3$
$\tilde{\omega}$	0	6	0	0	25	22	0	10	0	36	40
$\tilde{\psi}$	6	0	18	2	0	0	30	0	29	0	0

Table 19.3 Computational results for all problem instances of Case Study 1 and 3

Example	Equations	Bin vars	Cont vars	Nodes	CPU s	Obj val (m.u.)
Problem 1.1	11,545	2,310	3,150	1,704	139	55.61
Problem 1.2	11,683	2,444	3,150	58,626	7,721	54.65
Problem 1.3	11,817	2,594	3,150	82,517	14,000	54.65
Problem 3.1	14,638	3,297	4,050	749	137	1,022.00
Problem 3.2	16,601	3,627	4,707	522	85	996.00

header $j1$ and it has been operating for six time periods (not necessarily serving $j1$ in all of these periods) since its last startup. Also, at the beginning of the planning horizon, compressor $i7$ has not been operating for 30 time periods since its last shutdown.

Table 19.3 presents the computational results for all problem instances of this case study. As expected, the introduction of additional binary variables related to flexible maintenance tasks makes the resulting planning problem harder to solve compared to the fixed maintenance plan policy.

Problem 1.1 Figure 19.7a displays the optimal operational and maintenance plan for this problem instance. Nine compressors are used to cover the demand over the 30 days, and compressors $i1$ and $i3$ do not operate at any time period. No header changes are observed. As expected, the initial state of the compressors affect the obtained plan. For instance, compressor $i10$, which has been operating continuously for 36 days before the beginning of the planing horizon, shuts down at the beginning of day 25 because reaches its maximum runtime of 60 days at that time period. The normalized load distribution for each compressor per time period can be found in Fig. 19.8. It is observed that small compressors usually operate close to (or exactly at) their maximum mass flow rate, while the large compressors change their mass flow rate to a broader range to cover the demand fluctuations. This result was expected because small compressors are controlled with inlet throttling valves while large compressors with IGVs. In practice, the throttling of the small compressors makes them inefficient in lower mass flows. This output was observed from the past operation of the plant which shows a similar operational pattern. Therefore, despite the fact that the regression models used do not explicitly assess the performance of each compressor, they achieved to capture implicitly the efficiency of the compressors. Indeed, the small compressors operate close to their maximum load as validated by

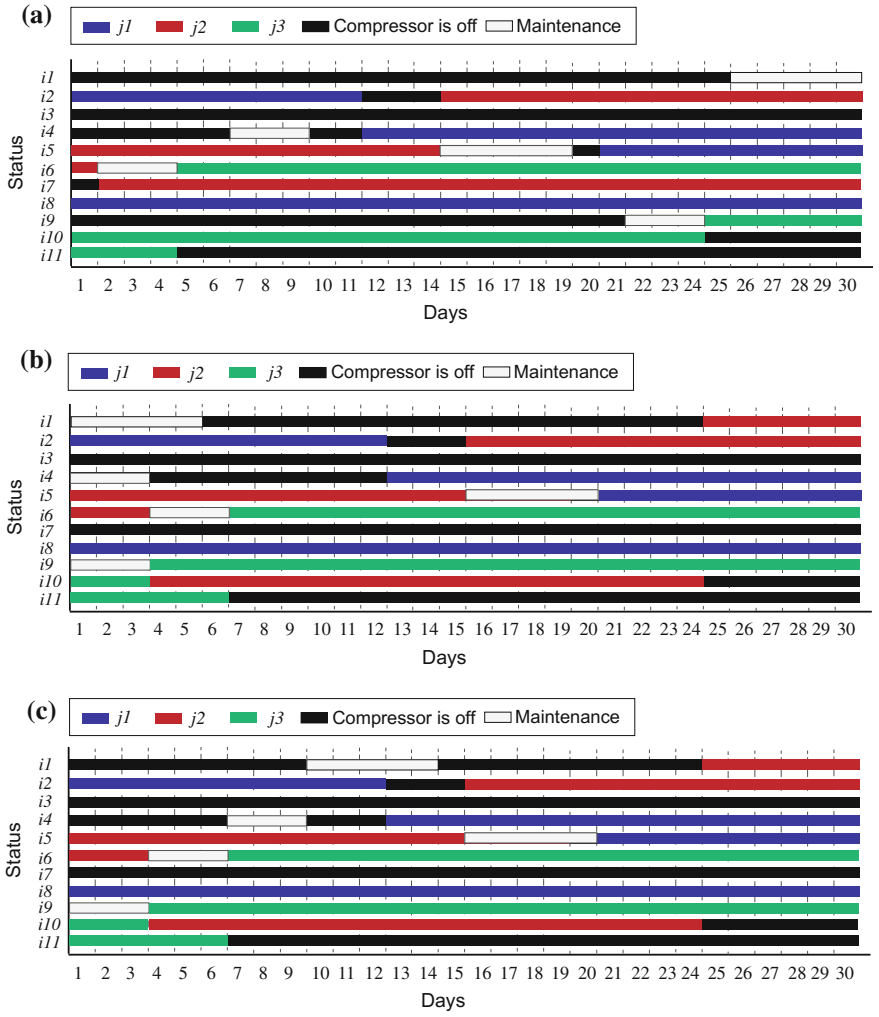

Fig. 19.7 Case Study 1: operational and maintenance plans for compressors. **a** Problem 1.1 (fixed maintenance). **b** Problem 1.2 (unconstrained flexible maintenance). **c** Problem 1.3 (constrained flexible maintenance $\eta_t = 1$)

the operators of the plant as well. In addition, a number of large compressors, such as $i6$ and $i10$, tend to operate at a fixed mass flow rate while other large compressors with better efficiency in a broader range cover the demand variations.

Problem 1.2 Figure 19.7b gives the operational and maintenance plan for the unconstrained flexible maintenance case. Due to the simultaneous optimization of maintenance and operational tasks, the maintenance tasks have been scheduled differently from Problem 1.1 in such a way that the total cost has been decreased by 1.8 %

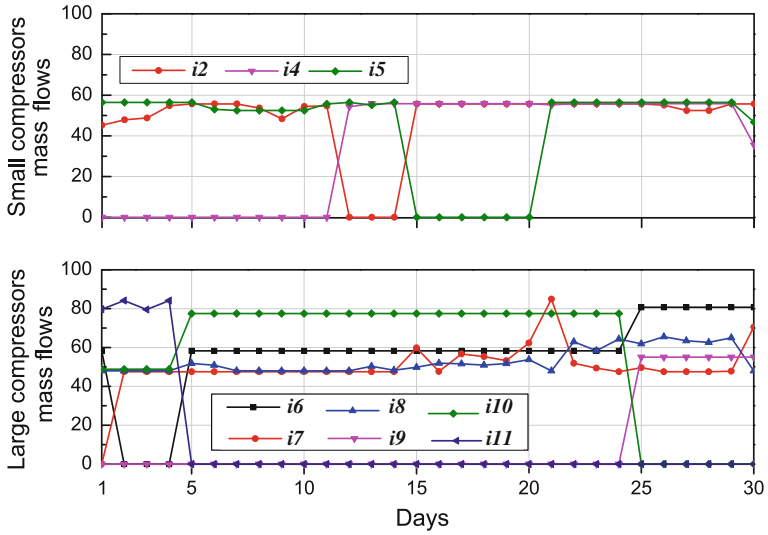

Fig. 19.8 Case Study 1: normalized outlet mass flow rates of compressors for Problem 1.1

in comparison to that in Problem 1.1 (see Table 19.3). Also, all maintenance tasks have been completed before day 21, where demand for products increases significantly. In this problem instance, there are several time periods where multiple maintenance tasks take place at the same time. For example, from day 1 to 3, three maintenance tasks are performed simultaneously, and in day 4 and 5, two maintenance tasks take place at the same time. As already discussed in Sect. 19.5.3.1, the simultaneous maintenance of multiple compressors may give rise to managerial or technical issues. For this reason, a constrained flexible maintenance policy is considered in the next problem instance in order to avoid simultaneous maintenance tasks.

Problem 1.3 In this constrained flexible maintenance case, at most one maintenance task can take place in every time period (i.e., $\eta_t = 1$). Figure 19.7c displays the optimal operational and maintenance of this problem instance. As it can be seen from Table 19.3, the total cost of Problem 1.3 is the same with that of Problem 1.2. This is because the resulting plan did not change much from Problem 1.2, since the only difference here is that the maintenance tasks in compressors $i1$ and $i4$ have been right-shifted in order to avoid multiple simultaneous maintenance tasks. These two changes did not affect the total cost since they do not add any cost, and the rest of the plan is the same as that in Problem 1.2. Although Problem 1.2 and Problem 1.3 generate the same total cost, Problem 1.3 gives a better operational and maintenance plan in terms of flexibility and managerial perspective, since maintenance tasks are distributed more uniformly over the planning horizon of interest.

Finally, Fig. 19.9a provides a comparison for the aggregated normalized total cost between Problem 1.1 and Problem 1.3. Figure 19.9b shows the normalized total cost breakdown for all problem instances (Problem 1.2 is exactly the same with

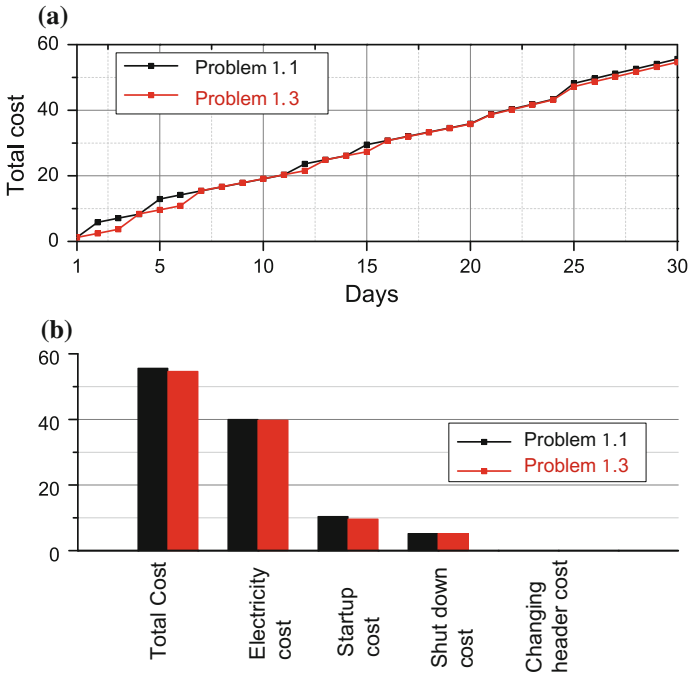

Fig. 19.9 Case Study 1. Problem 1.1 and Problem 1.3: **a** aggregated normalized objective value, and **b** normalized total cost breakdown

Problem 1.3). It is observed that startup and power consumption costs decrease under a flexible maintenance policy (i.e., Problem 1.2 and Problem 1.3). More specifically, in these cases power consumption cost reduces by 0.62 %, and startup cost declines by 7.70 % in comparison with those in Problem 1.1. Overall, it has been demonstrated that the integrated planning of maintenance and operational tasks could provide better solutions and decrease the total cost.

19.7.2 Case Study 2: Demand Uncertainty—Reactive Planning

In this problem instance, we show how the proposed optimization-based framework can cope successfully with an industrial case study under the presence of demand uncertainty. Similar to the previous case study, oxygen and compressed air are the products, change header cost is equal to 250 m.u./change, and purchases of products are not allowed. The demand for products fluctuates over time and for this reason the demand profiles are updated after each day period, through forecasting for a given prediction horizon. For the 30-day planning horizon considered, Fig. 19.10 shows the

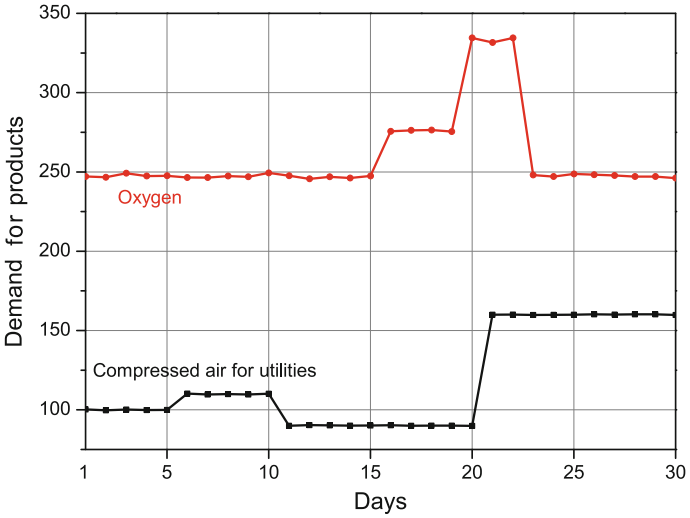

Fig. 19.10 Case Study 2: normalized demand for products (actual values)

actual normalized demand profile for oxygen and compressed air for utilities. The maximum runtime for compressors and the initial condition for each compressor are the same as in the previous case study (see Table 19.2). Here, some compressors follow a fixed maintenance policy and others a flexible time-window maintenance policy. Compressors $i9$ and $i11$ have been pre-scheduled for maintenance in days 20 and 15, respectively. Compressors $i2$ and $i6$ are under flexible time-window maintenance, and their corresponding maintenance tasks must start between days 1 and 15. The duration for each maintenance task is equal to three days.

The reactive planning problem has been solved via a rolling horizon approach, as the one displayed in Fig. 19.4. A time period is equal to one day. A prediction horizon equal to 21 time periods, and a single-period control horizon have been used. In other words, there is a new demand prediction for the following 21 days after every day. A total number of 30 iterations has been solved. An iteration represents a planning problem which employs a prediction horizon of 21 days. A time limit of 1,800 CPU s was set for each iteration. On average, solutions are obtained in low computational times. Negligible optimality gaps have been reported for iterations 20–22, where the imposed time limit was reached. As discussed in Kopanos and Pistikopoulos (2014), longer prediction horizons result in bigger mathematical model sizes, and as a consequence the necessary computational time will probably increase too.

At the beginning of each time period, a planning problem for the next 21 time periods (i.e., the prediction horizon) is solved with updated information regarding the current state of the overall system and the demand for products. The solution of the first time period of the current prediction horizon is only applied. In this problem instance, an unexpected breakdown of compressor $i5$ in the second time period is considered. Figure 19.11 displays an example of how the overall plan (for the 30-day

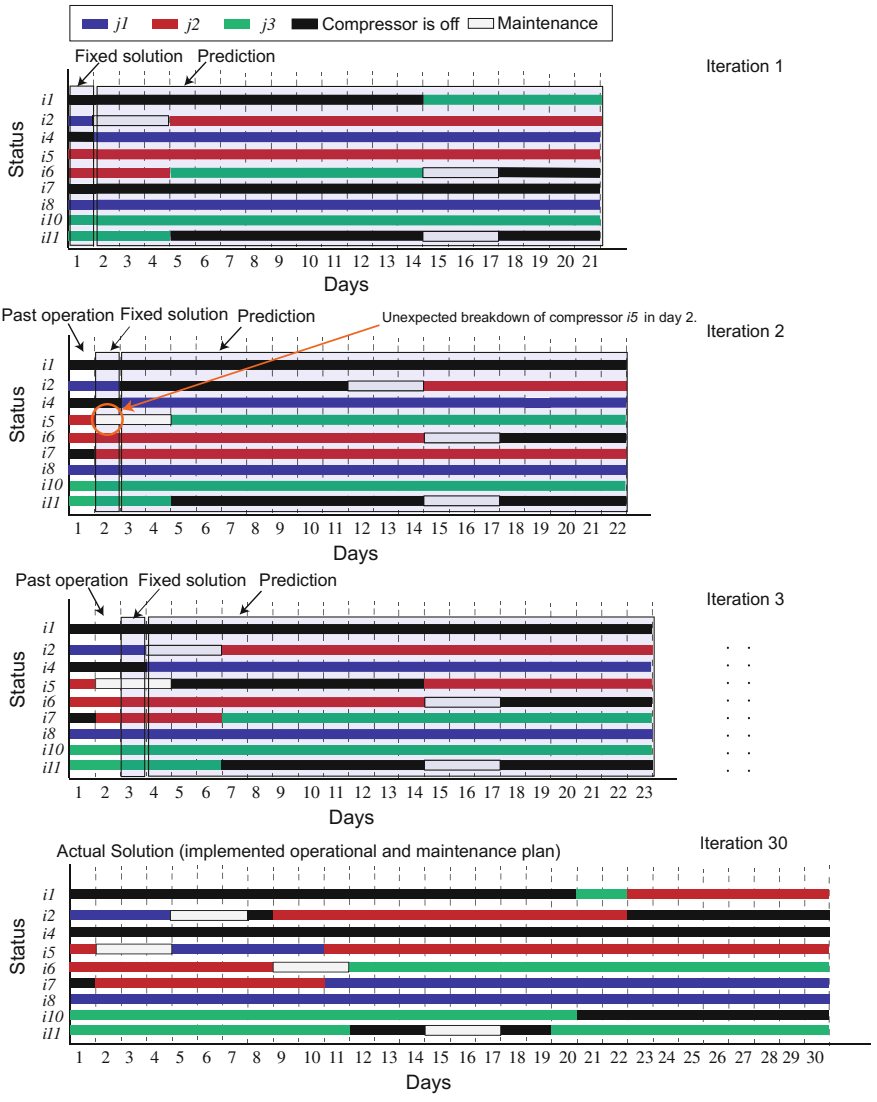

Fig. 19.11 Case Study 2: operational and maintenance plan via rolling horizon

horizon considered) is constructed through the solutions obtained in each iteration. The last Gantt chart in Fig. 19.11 gives the implemented operational and maintenance plan; compressors that do not operate in any iteration are not displayed.

The normalized load distribution of each compressor is given in Fig. 19.12. It can be observed that compressors *i*3, *i*4, and *i*9 remain idle during the 30-day planning horizon considered. It is known from the plant operators that compressor *i*3 is the most energy-inefficient compressor, and the results of all case studies give a

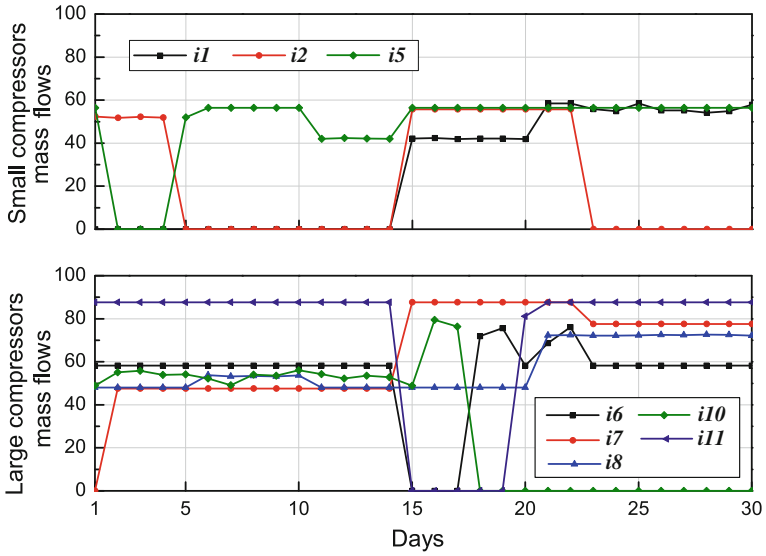

Fig. 19.12 Case Study 2: normalized outlet mass flow rates of compressors

clear evidence of it, since the obtained solutions preferred to keep it shutdown or operate it in limited capacity. At this point, if we consider the case that the demand for products for all 30 time periods is known with certainty at the beginning of the time horizon of interest (i.e., so-called perfect information case), one could solve the deterministic planning problem for the whole planning horizon at once. Actually, the perfect information solution is the best solution that one could obtain, however, in practice, this solution is impossible to be found due to the uncertainty in the demand forecasting. The aggregated total cost of the perfect information case and the solution derived by the rolling horizon approach are shown in Fig. 19.13. The results show that the rolling horizon solution is 11 % worse than that of the perfect information case. The obtained solution could be improved, if the forecasting accuracy is improved and the length of the prediction horizon increases. However, in practice as the length of the prediction horizon increases, the forecasting accuracy naturally decreases. Overall, through this problem instance it has been shown the applicability of the proposed optimization-based framework to deal with relevant planning problems in dynamic production environments.

19.7.3 Case Study 3: Condition-Based Maintenance

In this case study, we consider online and offline condition-based maintenance tasks. The normalized demand profile for products for this case study is displayed in Fig. 19.14. Nitrogen is considered in the product list here. The initial state of the

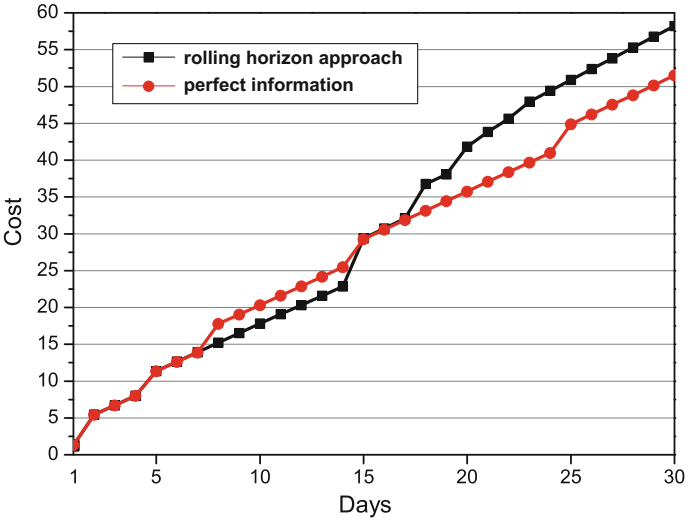

Fig. 19.13 Case Study 2: aggregated normalized objective value for the rolling horizon and the perfect information solution

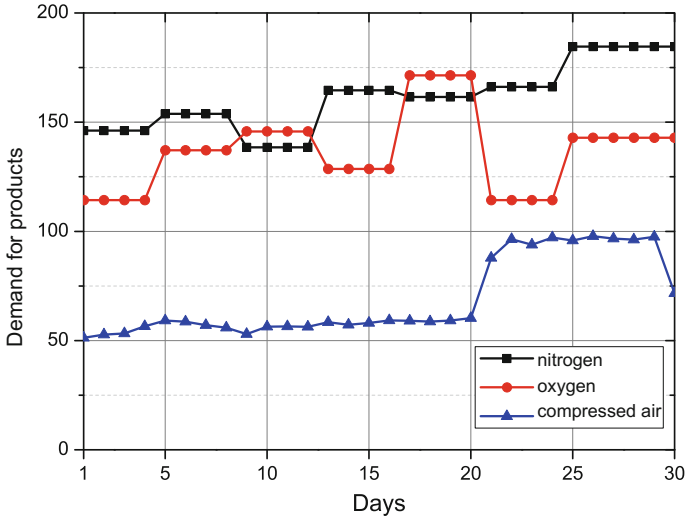

Fig. 19.14 Case Study 3: normalized demand for products

system under study as well as the other remaining input data can be found in Xenos et al. (2016). Two problem instances are considered here. Namely, a problem instance that only offline washing is allowed (Problem 3.1), and a problem instance where both online and offline washings are permitted (Problem 3.2). Table 19.3 presents the computational results for both problem instances. The solution of Problem 3.2, where

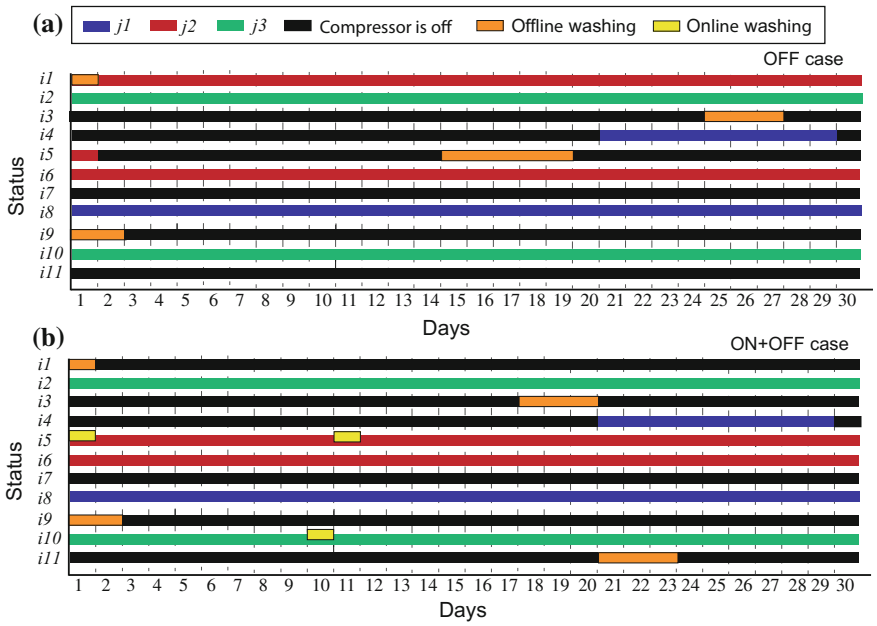

Fig. 19.15 Case Study 3: operational and maintenance plan considering **a** just offline washing, and **b** online and offline washing

online and offline washings are allowed, results in 2.5 % lower total cost than that of the more restricted Problem 3.1. The results of this case study clearly show that the maintenance strategy considerably affects the operational plan of the compressors network. The main reason that the operational and maintenance plan in Problem 3.2 (online and offline washing) results in lower total cost than that in Problem 3.1 (just offline washing allowed) is that the online washing complements to the offline resulting in a fewer number of compressors startups. The results of Problem 3.2 show that a compressor tends to operate continuously as much as possible, and this is actually in the same line with the industrial practice. For instance, Fig. 19.15a shows that compressor $i5$ shuts down after 41 days while Fig. 19.15b shows that the same compressor operates for 70 days.

As expected, the initial state of the system affects the optimal solution in both cases. The general trend of the solution of is that the compressors operate as much as possible providing compressed air to the headers with which the compressors are connected at the beginning of the planning horizon. However, not all active compressors-headers connections in the first time period is necessarily the same with those of the initial state. For instance, Fig. 19.15a shows that at the beginning of the planning horizon: (i) compressor $i11$ is switched off and it is not used at all, and (ii) compressor $i2$ has changed from header $j1$ to $j3$. The main reason for this is that compressor $i11$ has higher minimum limit of mass flow rate than compressor $i2$, therefore compressor $i11$ switches off and compressor $i2$ satisfies the load of header $j3$ along

with compressor *i*10. These changes support the further decisions which switch off compressor *i*5 at day 2 and startup compressor *i*1 connecting it to header *j*2. Finally, compressor *i*8 can satisfy the load of header *j*1 without the contribution of another compressor as the load of this header is relatively low compared to the load of the distillation columns connected with headers *j*2 and *j*3.

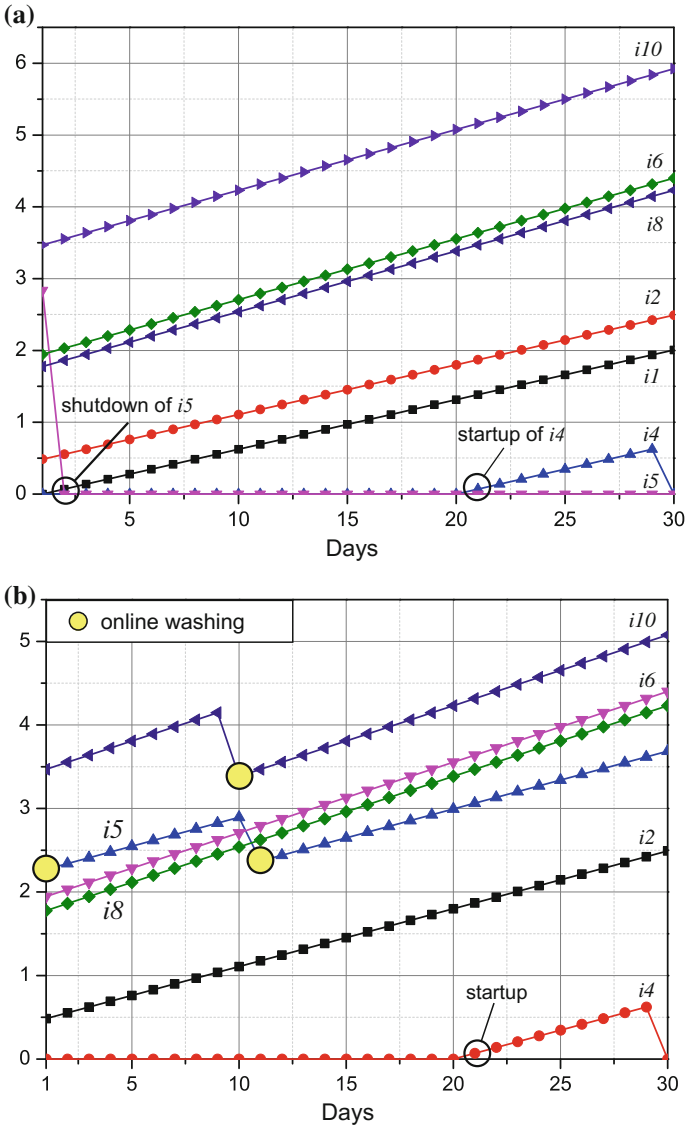

Fig. 19.16 Case Study 3: extra power consumption for compressors over time (scaled). **a** Problem 3.1 (considering just offline washing). **b** Problem 3.2 (considering online and offline washing)

In Problem 3.2, the results show that online washing is used along with offline washing, as it can be seen in Fig. 19.15b. Compressor $i11$ is not operating at all in both problem instances. In contrast to the solution of the previous problem instance, compressor $i5$ is not switched off but it is washed online and keeps operating throughout the whole planning horizon. Figure 19.16 displays the extra power consumption of each compressor in scaled units for both problem instances. In the last problem instance, compressor $i10$ undergoes online washing in day 10 in order to recover part of its performance, resulting in reduced extra energy consumption.

19.8 Conclusions

In this chapter, a general mathematical programming framework for the simultaneous optimization of maintenance and operational tasks of compressors networks in air separation plants has been presented. A distinctive feature of the proposed approach is that the power consumption in compressors is expressed by regression functions. The suggested approach considers operating constraints for compressors, several types of maintenance policies (i.e., runtime-based, fixed, flexible time-window, and condition-based) as well as managerial aspects regarding maintenance decisions. Power consumption, startup, shutdown and change header costs for the compressors in tandem with their maintenance tasks are optimized. The case studies solved have demonstrated that the simultaneous optimization of maintenance and operational tasks of compressors (i.e., flexible maintenance case) favor the generation of better solutions in terms of total costs in comparison to the predefined maintenance alternative (fixed maintenance case). According to the case studies, this is mainly due to the fact that the flexible maintenance policy results in reduced startup, shutdown and power consumption costs. The nature of the problem under consideration is combinatorial involving a large number of scenarios, and therefore it is impractical and probably impossible to identify the optimal patterns for the best decisions without the use of optimization. Even the most well-trained and experienced team of managers and operators could find it difficult to take the best decisions without using a systematic decision support tool, as the optimization model presented here. Ongoing research involves the more thorough study, modeling and optimization of condition-based maintenance tasks and their integration in the performance monitoring of the compressors.

Acknowledgments The financial support from the Marie Skłodowska-Curie Horizon 2020 project "PRONTO: PROcess NeTwork Optimization for efficient and sustainable operation of Europe's process industries taking machinery condition and process performance into account", Contract No: 675215 PRONTO is gratefully acknowledged. Financial support from the Marie Curie FP7-ITN project "Energy savings from smart operation of electrical, process and mechanical equipment–ENERGY-SMARTOPS", Contract No: PITN-GA-2010-264940 is gratefully acknowledged too. BASF SE Ludwigshafen is also acknowledged for providing the case studies.

References

Aretakis, N., Roumeliotis, I., Doumouras, G., & Mathioudakis, K. (2012). Compressor washing economic analysis and optimization for power generation. *Applied Energy*, *95*, 77–86.

Brooks, D. G., Carroll, S. S., & Verdini, W. A. (1988). Characterizing the domain of a regression-model. *American Statistician*, *42*(3), 187–190.

Camponogara, E., Nazari, L. F., & Meneses, C. N. (2012). A revised model for compressor design and scheduling in gas-lifted oil fields. *IIE Transactions*, *44*(5), 342–351.

Han, I., Han, C., & Chung, C. (2004). Optimization of the air and gas-supply network of a chemical plant. *Chemical Engineering Research and Design*, *82*(A10), 1337–1343.

Ierapetritou, M. G., Wu, D., Vin, J., Sweeney, P., & Chigirinskiy, M. (2002). Cost minimization in an energy-intensive plant using mathematical programming approaches. *Industrial and Engineering Chemistry Research*, *41*(21), 5262–5277.

Karwan, M. H., & Keblis, M. F. (2007). Operations planning with real time pricing of a primary input. *Computers and Operations Research*, *34*(13), 848–867.

Kopanos, G. M., Georgiadis, M. C., & Pistikopoulos, E. N. (2013). Energy production planning of a network of micro combined heat and power generators. *Applied Energy*, *102*, 1522–1534.

Kopanos, G. M., & Pistikopoulos, E. N. (2014). Reactive scheduling by a multiparametric programming rolling horizon framework: A case of a network of combined heat and power units. *Industrial and Engineering Chemistry Research*, *53*(11), 4366–4386.

Kopanos, G. M., Xenos, D. P., Cicciotti, M., Pistikopoulos, E. N., & Thornhill, N. F. (2015). Optimization of a network of compressors in parallel: Operational and maintenance planning—The air separation plant case. *Applied Energy*, *146*, 453–470.

Li, Y. G., & Nilkitsaranont, P. (2009). Gas turbine performance prognostic for condition-based maintenance. *Applied Energy*, *86*(10), 2152–2161.

Mitra, S., Grossmann, I. E., Pinto, J. M., & Arora, N. (2012). Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes. *Computers and Chemical Engineering*, *38*(12), 171–184.

Rong, A., & Lahdelma, R. (2005). An efficient linear programming model and optimization algorithm for trigeneration. *Applied Energy*, *82*(1), 40–63.

Saidur, R., Rahim, N. A., & Hasanuzzaman, M. (2010). A review on compressed-air energy use and energy savings. *Renewable and Sustainable Energy Reviews*, *14*(4), 1135–1153.

Thorin, E., Brand, H., & Weber, C. (2005). Long-term optimization of cogeneration systems in a competitive market environment. *Applied Energy*, *81*(2), 152–169.

US Department of Energy, 2003. Improving compressed air system performance.

van den Heever, S. A., & Grossmann, I. E. (2003). A strategy for the integration of production planning and reactive scheduling in the optimization of a hydrogen supply network. *Computers and Chemical Engineering*, *27*(12), 1813–1839.

Xenos, D. P., Cicciotti, M., Kopanos, G. M., Bouaswaig, A. E. F., Kahrs, O., Martinez-Botas, R., et al. (2015). Optimization of a network of compressors in parallel: Real Time Optimization (RTO) of Compressors in chemical plants—An industrial case study. *Applied Energy*, *114*, 51–63.

Xenos, D. P., Kopanos, G. M., Cicciotti, M., & Thornhill, N. F. (2016). Operational optimization of networks of compressors considering condition-based maintenance. *Computers and Chemical Engineering*, *84*, 117–131.

Zhu, Y., Legg, S., & Laird, C. D. (2011). Optimal operation of cryogenic air separation systems with demand uncertainty and contractual obligations. *Chemical Engineering Science*, *66*(5), 953–963.

Zhuan, X. T., & Xia, X. H. (2013). Optimal operation scheduling of a pumping station with multiple pumps. *Applied Energy*, *104*, 250–257.

Chapter 20
Integrating Energy Optimization and Production Scheduling in Energy-Intensive Industries

Lennart Merkert and Iiro Harjunkoski

Abstract The share of volatile renewable energy generation is rapidly increasing in many countries around the world. As in electric power grids the supply always needs to equal the demand, the increasing volatility of energy supply imposes a major challenge to the stability of power grids. Demand response actions offer a very cost-efficient way to cope with this challenge. Especially energy-intensive industries such as metals, cement or pulp and paper offer a high potential to adjust their energy consumption towards the power grid in the form of large controllable loads. In this chapter we look into how this potential could be used without affecting the production volume. Advanced scheduling algorithms allow to efficiently plan the production at industrial sites. Enabling such scheduling algorithms for energy-aware production planning as well as the integration of scheduling and energy optimization solutions allows to leverage a high potential for shifting energy consumption from times with low to times with high renewable energy generation. In addition, this approach allows plant owners to significantly reduce their energy cost.

20.1 Introduction

Increasing the share of renewable energy generation (wind, solar, biomass …) is a major political goal in many countries. In the United States, this is reflected e.g. in the US NSF program "Energy for Sustainability". In the European Union, the frame program "Horizon 2020" supports the increase in renewable energy generation. Many renewable power sources such as solar or wind power generation heavily depend on volatile weather conditions. As these forms of volatile generation are becoming more and more important in power grids this imposes major challenges to the stability of the electric grid. The volatility of renewable generation needs to be

L. Merkert · I. Harjunkoski (✉)
ABB AG Corporate Research, Ladenburg, Germany
e-mail: iiro.harjunkoski@de.abb.com

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_20

compensated by energy storage, controllable generation or demand-side management.

Short-term production planning—or scheduling—has in the last decades evolved both on the modeling as well as on the algorithmic front. Better and more compact models can today address problems of industrial relevance owing to improved mathematical representations and new process related features among the multitude of modeling options (e.g. State-Task-Network, Resource-Task-Network, event- and precedence based continuous-time methods), see e.g. Harjunkoski et al. (2014). In the light of the recent developments, it only is natural that the energy consumption of production processes also becomes part of the operational scheduling problem, especially since with emerging renewables and smart grids the exact time of consumption may significantly affect the price for the electricity. It is also today technically possible to already in the production scheduling phase take into account the varying electricity prices for the next days (Merkert et al. 2015). The joint consideration of earlier isolated problem aspects requires new ways of formulating and solving the integrated problems.

In this context, the main challenge is to ensure that the production scheduling model encapsulates the immediate effect of a scheduling decision on the energy consumption as part of the optimization problem. This resembles the challenge of integration between scheduling and control (Engell and Harjunkoski 2012; Chu and You 2012; Zhuge and Ierapetritou 2012; Baldea and Harjunkoski 2014), where different dynamics including the inertia of a process must be optimally combined. There are also a number of contributions on heat integration and scheduling (Georgiadis and Papageorgiou 2001; Halim and Srinivasan 2009; Seid and Majozi 2014) where also energy storage and water re-use are considered. Many of the contributions in heat integration are in fact related to process design, whereas here the focus is on the operational scheduling. For a review of the major contributions in energy handling, see e.g. Fernández et al. (2012). Most of the earlier contributions do not consider continuously varying energy or electricity prices, which is a major topic and challenge for the energy-intensive industry of today.

20.2 Energy Management in Industry

With the industrialization of many developing countries the global energy need is expected to be growing continuously (Conti and Holtberg 2011). Industry accounts for about one third of overall energy consumption in Europe (http://www.eea.europa.eu/data-and-maps/figures/sector-share-final-energy-consumption) and in the US (http://www.eia.gov/beta/MER/index.cfm?tbl=T02.01#/?f=A&start=200001). As in many countries increased energy efficiency and reduction of CO_2 emissions is on the political agenda (European Union 2011), the most energy intensive industries (metals, pulp and paper, chemicals, cement and air separation) are studied for efficiency potentials. For those industries the energy cost is one of their major

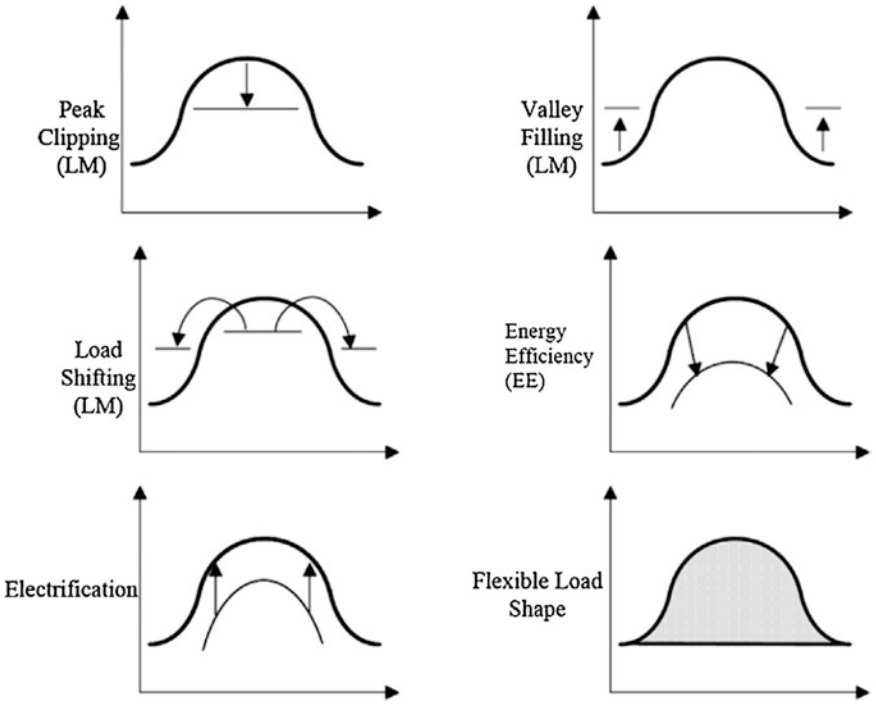

Fig. 20.1 Load shapes (Charles River Associates 2005)

operational cost factors, thus improved efficiency and other ways to reduce energy cost directly influence their profits.

There are different ways how electricity consumers can shape their load profiles. Figure 20.1 shows an overview on these options allowing to categorize load management actions by consumers.

20.2.1 Energy Efficiency

Producing more energy efficiently means that the same amount of product is produced with less energy consumed—or that more is produced with the same energy consumption. Increasing the energy efficiency is very interesting for industry as besides reduced CO_2 emissions consuming less energy directly reduces the energy cost.

There are different well established ways of how to increase the energy efficiency. The following list only shows some often implemented actions:

- Installation of more efficient equipment (e.g. replacing light bulbs with LEDs).
- Replacing pumps with throttle and other electric machines not running at full capacity with variable speed drives.

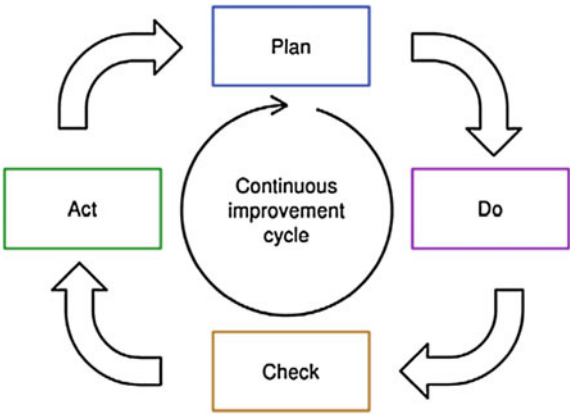

Fig. 20.2 The Plan–Do–Check–Act cycle for continuous improvement (Olsen et al. 2012)

- Improve insulation of thermal systems to reduce thermal losses.
- Improve operational strategy (e.g. shutdown of equipment that is not needed during certain time periods like production breaks).
- Early detection and elimination of inefficient operations (e.g. leakages in compressed air distribution systems).

In order to justify investments in energy efficiency it is important to understand and track the energy efficiency gains. Energy management systems allow to measure and monitor current and historic energy consumption. Using such a system allows to justify investments in energy efficiency as it enables return of invest calculations as well as to prove the success of energy efficiency actions. However monitoring the consumption alone does not achieve important energy savings. In order to continuously reduce the energy consumption, processes like the Plan-Do-Check-Act cycle (Fig. 20.2) should be implemented, in order to ensure that corrective actions against inefficiencies are implemented.

The international standard ISO 50001 describes the implementation of energy management systems in an organization. Continuous improvement of energy efficiency is one of its key elements. As the implementation of an energy management system is the foundation for increased energy efficiency, countries e.g. Germany support the implementation of ISO 50001 compliant energy management systems in industry and other organizations with reductions of energy taxes.

20.2.2 Demand Side Management

With the rising share of renewable energy generation, the volatility of electric power generation has grown importantly. Hence the balancing of load and demand in the electric grid has become more and more difficult. Demand-side management (the adaption of a load profile to power grid needs) allows to cope with this

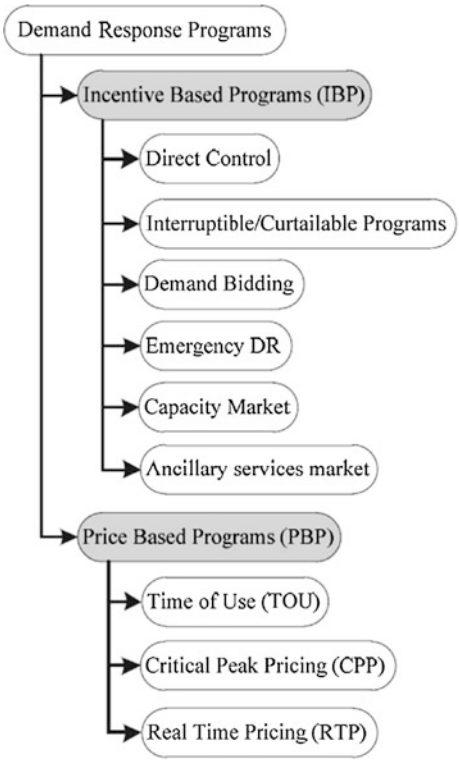

Fig. 20.3 Classification of demand response programs (Baboli et al. 2011)

challenge in a cost-efficient way. As Fig. 20.1 shows, there are several ways how load profiles can be influenced. In practice, there are two major ways of demand-side management besides energy efficiency:

- Load shedding or peak clipping (shut down non-essential loads when a, typically from the outside initiated, trigger occurs).
- Load shifting (shift energy consumption from peak to valley hours).

To motivate energy consumers to shed or shift loads they need to get a remuneration. These remunerations are mostly organized in demand response programs. Figure 20.3 provides an overview of different categories of such demand response programs.

Incentive based programs mostly implement a remuneration scheme for load shedding allowing e.g. grid operators to shut down big electric consumers in industrial facilities. Some plants already participate very actively in this market. For example, the aluminum company Alcoa participates with 70 MW in direct load control and with 75 MW as interruptible load at their site in Warrick, US (Todd 2011). Hence, the aluminum smelters operate based on power grid needs, allowing Alcoa to generate additional revenues.

Price based programs normally address a longer time scale than incentive based programs that allow direct control actions. For price based programs, the load shaping effects are more indirect as the electricity price profile should incentivize consumers to change their consumption profile. Reacting to such price profiles could be done by integration of energy management with production scheduling. When energy intensive production steps are shifted into lower price periods this can enable significant operational cost savings.

20.3 Integrating Energy Management and Production Scheduling

20.3.1 Scheduling in Industry

Scheduling is a decision-making process that plays an important role in most manufacturing and service industries (Pinedo and Chao 1999). Scheduling problems arise in almost any type of industrial production such as Chemicals, Food and Beverages, Pulp and Paper, Metals, Refineries, Pharmaceuticals, Transportation, Service, etc., where given tasks need to be processed on specified resources. The most common decisions are:

- The sequence of executing selected tasks
- The location (e.g. equipment) for processing the tasks
- The timing for executing the tasks

In short, for almost any process, the production must be appropriately planned such that the equipment, material, utilities, personnel and other resources are available at the plant exactly when needed. An illustration of the decisions is shown in Fig. 20.4. In this chapter, the main focus is on energy that is normally categorized as a utility. Production scheduling by itself is a complex task comprising the activity of planning the production of e.g. customer orders in detail on a given production facility. The complexity often occurs from the amount of possibilities, each of which may result into a different cost. Recent reviews on production scheduling and its industrial relevance can be found in Ribas et al. (2010) and

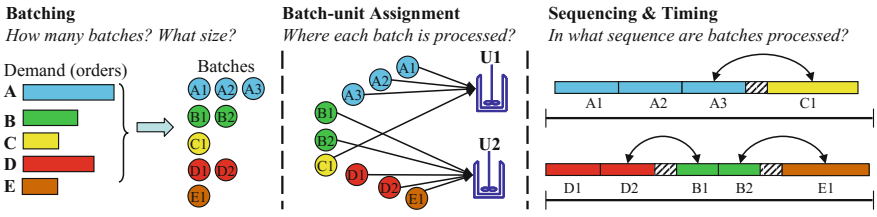

Fig. 20.4 Main scheduling decisions (Harjunkoski et al. 2014)

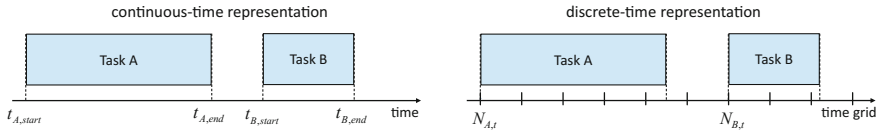

Fig. 20.5 Continuous- and discrete-time representations

Harjunkoski et al. (2014), where also the main references to academic methodological overviews are collected.

Here, we only shortly address some of the major characteristics of scheduling models without going into details. One of the most important decisions is how to represent the time. The natural way is to use a continuous-time representation where an event can start at any time and exact schedules can be produced. Here, so called precedence models are widely used. For example, if the start times of batches A and B are represented by the continuous variables $t_{A,\,start}, t_{B,\,start}$ the inequality $t_{A,\,start} < t_{B,\,start}$

would imply that batch B can start only after batch A has started. There are several continuous-time based scheduling formulations and many of them suffer from two main problems: Firstly, for a generic scheduling problem the above hard-limitation needs to be relaxed to also account for the opposite sequence (process B before A), which results in a large number of combinations. Typically, the relaxation is done using so called big-M constraints that require additional binary (0-1) variables, y_{AB}, which when true (= 1) state that A should be processed before B. The second problem arises from the possible need to track material and energy balances in which the standard continuous-time approach does not work. This requires some kind of balance points, around which a mass balance can be built. Some of these aspects are discussed in detail e.g. in Floudas and Lin (2004).

In Fig. 20.5 two tasks on a timeline is shown, both using continuous-time (left) and discrete-time (right) representations. This reveals some fundamental differences between the two approaches.

- The decision variable for a continuous-time representation is the exact time (t) when a task should start, whereas in discrete-time approach the decision is to select the most suitable time slot (N) that is pre-assigned to a given point of time.
- Following this, in discrete-time a task may only start at the specified grid points. Normally the decision variables do not include the task end times in discrete-time formulations.
- In discrete-time approaches the equipment assignment is often coupled with the timing decision, whereas this is a separate decision in most continuous-time approaches. Material balances can be easily tracked at the grid points.

Due to the mismatch of the selected grid points and "real" continuous time, part of the true equipment capacity may remain unplanned if a task ends before the next

grid point. Very well received and generic discrete-time approaches are State-Task Network (STN, Kondili et al. 1993) and Resource-Task Network (RTN, Pantelides 1994) formulations. In e.g. the RTN approach the entire discrete-time scheduling formulation is built around a unified resource balance constraints, which makes it very attractive an approach for various problems.

$$R_{r,t} = R_r^0|_{t=1} + R_{r,t-1} + \sum_i \sum_{\theta=0}^{\tau_i} \mu_{r,i,\theta} N_{i,t-\theta} + \Pi_{r,t} \; \forall r, t \tag{20.1}$$

Here in Eq. (20.1), $R_{r,t}$ refers to the amount of resource r at time t. The initial status is marked with a superscript 0 and applies to the first time grid point. In all other cases the amount of resources changes from the earlier $(t-1)$ state depending on the tasks i and their duration τ_i. The binary variable $N_{i,t}$ defines at which time point t the task i begins. The key parameter $\mu_{r,i,\theta}$ tracks the need of resource r during task i. The parameter can be positive or negative and the value may vary along the task duration τ_i. By setting the parameter value to -1 at the beginning of a task and to $+1$ at the end of a task, it can be used to occupy and release equipment or other discrete resources. The last term $\Pi_{r,t}$ enables interaction with the environment, e.g. demand fulfillment where a given amount of product is shipped away.

While the material tracking can be seen as a centerpiece of discrete-time formulations, it is not such a natural fit for continuous-time formulations, as already stated above. Typical objective of a continuous-time formulation is to minimize the total completion time (make span) at which the last task ends or to maximize the throughput. The main constraints include equipment assignment for each production stage, sequencing between the tasks for avoiding overlapping activities on restricted units, timing relation between successive stages and the make span definition. These are discussed in detail in the next section.

20.3.2 Holistic Approach for Scheduling with Energy Management

There are several ideas on how to deal with the energy management aspect and production scheduling (Merkert et al. 2015; Gahm et al. 2016). The most intuitive and appealing approach is to extend an existing scheduling formulation with the energy-related variables, constraints and objective components. This result in a holistic and valid problem formulation, which is also guaranteed to result in a globally optimal solution. Additionally, this approach should fully follow the production scheduling rules and constraints. In this context it is not possible to provide the full mathematical models but rather highlight the main features of selected approaches.

20.3.2.1 Main Principles

Building a holistic approach to combine scheduling and energy considerations requires model extensions for most existing scheduling formulations. One of the main extensions is to be able to consider the time intervals that are used for energy pricing and tracking. This is often country-specific and typically either 15 or 60 min. This means in practice that we need to have a grid point to account for energy balances every 15 or 60 min. In principle, the standard STN/RTN discrete-time approaches could work off-the-shelf if this time grid also matches with the process requirements. In practice, many batch processes must be planned more exactly than e.g. 60 min, whereas for continuous processes this time interval may be sufficient as there are fewer planned changes needed. For continuous-time formulations the challenge is to add the grid points and match them to the planning timeline. There are some existing methods using global or unit-specific event points and more technical insights are discussed in a review paper on discrete and continuous-time approaches by Floudas and Lin (2004).

Another required change is to add the necessary variables and parameters to ensure that the resulting electricity price can be calculated and considered by the optimization. Electricity price can e.g. be the purchasing price from markets with hourly varying tariffs or prices based on the time-of-use (recall Fig. 20.3). Furthermore, there may be penalties involved that stem from energy consumption or load deviations from an earlier committed plan. These must also be tracked hourly (or every 15 min) and may constitute a significant part of the electricity bill.

20.3.2.2 Using Mathematical Programming

The models can be built using mixed integer linear programming (MILP). Here we focus on batch processes and the main model components are highlighted.

RTN Approach

An example on how to model the scheduling of a steel plant using the discrete-time RTN is discussed in detail in Castro et al. (2013), where three modeling alternatives are presented. The discrete-time formulation is linked to a single uniform time grid featuring T slots of width δ (min) that span over 24 h. The parameter δ can be chosen by the modeler, effectively setting the approximation level of the problem data. Consequently, the real processing times d_i are rounded to a multiple of δ, leading to the calculation of parameters $\tau_i = \lceil d_i/\delta \rceil$ that give the duration of task i in number of time slots. In other words, the discrete-time representation cannot, in general, accurately model events occurring at time instances that are not multiples of δ, particularly when considering changeover or transfer times. Overall, wherever accuracy becomes a serious issue, continuous-time models should be tried despite being potentially more complex. Here, electricity is actually an aggregate representation of two resources: power (MW) $r \in R^{PW}$ and energy (MWh) $r \in R^{EN}$. Nonnegative continuous variables $\Pi_{r,t}$ are used to calculate the

power/energy requirements in slot t. Typically, a process states its power demand and tracking is done over the consumed energy.

Here, we slightly modify Eq. (20.1). The $R_{r,t-1}$ term is not written for power and energy resources to avoid them from propagating from one slot to the next, such that the external variables $\Pi_{r,t}$ come into play at the right time. The last term is unchanged and is negative for electricity consumption.

$$R_{r,t} = R_r^0|_{t=1} + R_{r,t-1}|_{r \notin (R^{PW} \bigcup R^{EN})} + \Pi_{r,t}|_{r \in (R^{PW} \bigcup R^{EN})} + \pi_{r,t}|_{r \in R^{EQ}} + \sum_i \sum_{\theta=0}^{\tau_i} \mu_{r,i,\theta} N_{i,t-\theta} \ \forall r,\ t \tag{20.2}$$

Additional constraints are needed. Assume that subset $I_{h,u}$ holds the task corresponding to the processing of heat h in unit u and that subset U_k comprises the units able to process the stage k. Equation (20.3) ensures that all heats are processed exactly once per each stage k.

$$\sum_{u \in U_k} \sum_{i \in I_{h,u}} \sum_t N_{i,t} = 1 \ \forall h,\ k \tag{20.3}$$

For power and energy we also need to prevent existence of resources linked to power and energy, resulting in that the consumption (last term in Eq. (20.2)) is always balanced with the external variable $\Pi_{r,t}$.

$$R_{r,t} = 0 \, \forall \, r \in \left(R^{PW} \cup R^{EN}\right), \forall t \tag{20.4}$$

Further constraints may define upper bounds on the maximum power and energy consumption, as well as, compute the deviations for energy resources and penalties for over- and under-consumption. The entire model and the computational results are reported in Castro et al. (2013). It turns out that an accuracy of more than 15 min for a 24 h planning horizon becomes very hard to solve within reasonable time (few minutes) and for many batch problems, the desired "exactness" should be at minimum 5 min.

General Precedence Approach

Since the above RTN approach does not always provide a sufficient accuracy, we shortly discuss how a general precedence continuous-time approach can be expanded to cover the typical demand-side management needs. The complete models and explanations are available in Hadera et al. (2015). Here we simplify the models significantly in order to be able to illustrate the most relevant aspects more clearly. The standard scheduling model is based on precedence- and assignment variables that determine the sequence of production and on which of the parallel machines shall be used at each stage. Following the nomenclature in Hadera et al. (2015), the binary general precedence variable $V_{st,p,p'}$ is true (= 1) if a product p is processed before a product p' on stage st. The binary assignment variable $X_{p,m}$ is true if product p is processed on machine m. Equation (20.5) states that exactly one

machine should process a product per each stage. The subset $SM_{st,m}$ contains the machines m that are suitable for stage st.

$$\sum_{m \in SM_{st,m}} X_{p,m} = 1 \; \forall p \in P, \; st \; ST \tag{20.5}$$

Equation (20.6) relates the finishing time $t^f_{p,m}$ with the starting time $t^s_{p,m}$ considering the processing time $\theta_{p,m}$ on the selected machine m.

$$t^f_{p,m} = t^s_{p,m} + X_{p,m} \cdot \theta_{p,m} \quad \forall m \in M, \; p \in P \tag{20.6}$$

The precedence constraint in (20.7) restricts that a next heat should be processed only after the previous one has finished plus a setup time.

$$\begin{aligned} t^s_{p',m} \geq t^f_{p,m} + t^{setup}_m - \left(M + t^{setup}_m\right)\left(3 - V_{st,p,p'} - X_{p,m} - X_{p',m}\right) \\ \forall p, \; p' \in P, \; st \in ST, \; m \in SM_{st,m}, \; p \neq p' \end{aligned} \tag{20.7}$$

Here only the key constraints have been shown and the main addition to the standard scheduling model is how to map the grid points for tracking the electricity consumption to the model. One option from Hadera et al. (2015) is to define event binaries that are shown in Fig. 20.6.

Here the timing variables $t^s_{p,m}$ and $t^f_{p,m}$ are mapped with the starting time of each time slot through using event binary variables $Y^s_{p,st,s}$ and $Y^f_{p,st,s}$ denoting if the stage st of product p starts and ends within the slot s. In Fig. 20.6, the corresponding processing times of the task are shown and this can be used to summarize the

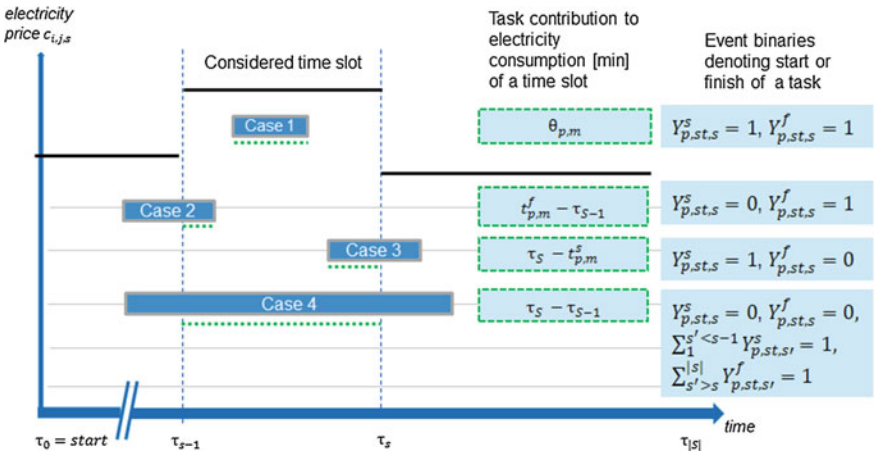

Fig. 20.6 Event binaries model to describe the consumption of electric energy in time slots of the price (Hadera et al. 2015)

electricity prices and deviations from committed loads over each time slot, resulting in the total energy cost.

Both approaches work and can theoretically deliver the global optimal solution but they also form very large optimization problems that are difficult to solve efficiently. Due to this reason it is not always possible to apply monolithic optimization models to the energy-aware scheduling of real industrial processes.

20.3.3 Iterative Approach

When implementing new features in industrial plants the existing environment needs to be taken into account. In many cases there are already separate solutions for production planning and/or energy management. Replacing the existing highly integrated solutions with one large system for energy management and production planning is often not wanted. A suitable alternative to this is a collaborative approach where production planning and energy management are integrated in an iterative way.

A combined production planning and energy management problem can always be formulated as a monolithic model with two nested optimization problems: production planning problem (PP) and energy flow network problem (EFN) (Ait-Ali 2015). The cost function of such problem can be described as follows:

$$F = C_{EFN}^T f + C_{PP}^T y \tag{20.8}$$

With respect to the functional constraints

$$A_{EFN} f + D_{EFN} x - b_{EFN} \leq 0 \tag{20.9}$$

$$A_{PP} y + D_{PP} q - b_{PP} \leq 0 \tag{20.10}$$

$$Af - Iq = 0 \tag{20.11}$$

$$f \geq 0;\ x \in X;\ y \in Y;\ q \geq 0 \tag{20.12}$$

The objective cost function (20.8) consists of the cost of the production planning—PP (production cost) and the cost of the energy (EFN). The inequality constraint (20.9) represents the functional constraints of the flow network or the energy management whereas (20.10) describes those of the production planning. The equality constraint (20.11) links the production demand of electricity to the energy optimization problem. The inequalities and inclusions in (20.12) represent the bounds as well as the type (continuous or discrete) of variables.

There are several decomposition methods which allow to separate this problem. With Primal decomposition as introduced by Benders (Benders 1962) the problem

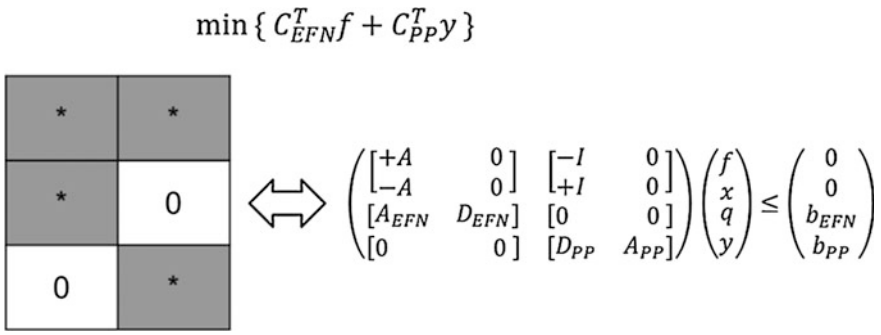

Fig. 20.7 Primal block-angular structure of the monolithic problem

is reformulated in a primal block-angular structure having the form shown in Fig. 20.7.

The problem is divided into two sub-problems which are linked by some constraints containing variables of both sub-problems (Eq. (20.11) in our case). Primal decomposition divides a complex problem into two problems: the primal sub-problem, which gives an upper bound and the primal master problem, which provides a lower bound. Normally both are iterated until an optimal solution is reached. However, the primal master problem can be even more complicated than the monolithic problem (Ait-Ali 2015).

Dual decomposition as proposed by Wolfe and Danzig (Wolfe and Danzig 1960) is using a different block-angular structure (Fig. 20.8).

For dual decomposition the different sub-problems are linked with a complicating variable (q in our case). This method uses Lagrangean relaxation to eliminate the complicating constraints. As for primal decomposition dual decomposition introduced two problems: the dual sub-problem giving a lower bound and the dual master problem giving an upper bound for the optimal solution. However, as for the primal decomposition the master problem has to be solved, which may be more difficult than solving the original monolithic problem (Ait-Ali 2015).

Roy (1983) introduced a method called cross decomposition which iterates between the primal sub-problem of the primal decomposition (Bender's decomposition) and the dual sub-problem of the dual decomposition (Danzig-Wolfe's

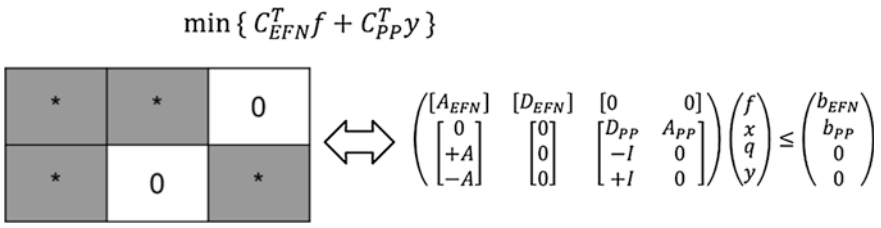

Fig. 20.8 Dual block-angular structure of the monolithic problem

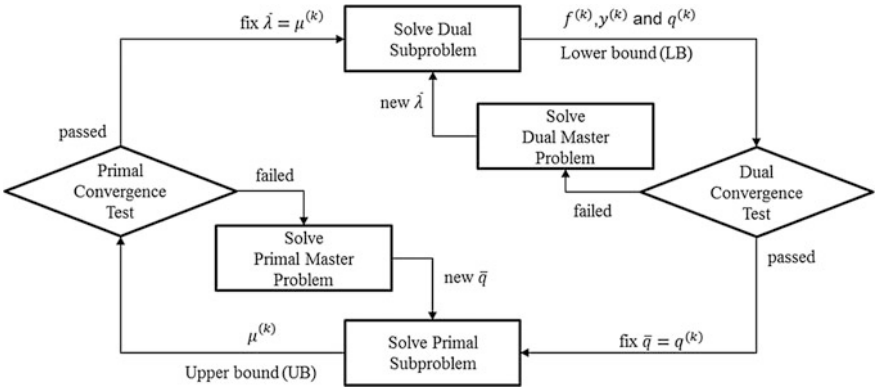

Fig. 20.9 Cross decomposition algorithm, based on (Hadera et al. 2015)

decomposition). In each iteration a convergence test is performed. If it fails, the according master problem needs to be solved (Fig. 20.9).

This method is having the advantage that only the sub problems need to be solved in most iterations. However the master problems still need to be generated which are complicated to solve (Ait-Ali 2015).

Holmberg (1992) introduced the mean-value cross decomposition which was proven to converge to optimality for linear programs (Holmberg 1994). Mean-value cross decomposition only iterates between the primal and dual sub problems without the need to solve any master problem in case of a failing convergence test. It uses the mean value of all past iteration solutions as input for the next iteration.

$$\mu^k = \frac{1}{k}\tilde{\mu}^{k-1} + \frac{k-1}{k}\mu^{k-1} \tag{20.13}$$

Here the signal input in the previous iteration is μ^{k-1} and $\tilde{\mu}^{k-1}$ is the previous solution from the dual or primal sub-problem and μ^k is the signal input in the current iteration.

For non-linear convex problems a method call weighted-mean-value cross decomposition has proven convergence to optimality (Holmberg and Kiwiel 2006). It uses a weighted mean to calculate the following input signal:

$$\mu^k = \delta_k\tilde{\mu}^{k-1} + (1-\delta_k)\mu^{k-1} \tag{20.14}$$

where $\delta_i = \frac{\beta+\gamma}{\beta i+\gamma}$.

In the case of production scheduling and energy monitoring, using mean values might result in infeasible sub-problems. One sided mean-value cross decomposition is a way to cope with this problem. Here, only one signal is constructed as mean value of all past input signals while the other input signal is calculated directly from the current solution of the sub problem. Despite giving good results in several

industrial size case studies, there is no convergence proof for this strategy (Ait-Ali 2015).

20.3.4 Use Cases

Especially energy intensive industrial processes such as air separation, cement production, electric steel production or thermo-mechanical pulping are interesting for the purpose of load shifting. On the one hand they offer as large energy consumers relatively easy control of a huge electric load. On the other hand, energy cost is a major share of operational costs such that already a small percentage of savings in energy costs can cause significant financial gains. Hence, not only technical feasibility but also economic feasibility is within reach.

20.3.4.1 Cement

Cement production is one of the core industries where energy is an important cost factor. Reducing this cost is of high interest, especially since energy costs are rising due to increasing energy demand and changes in the generation mix and supply. The entire process is shown in Fig. 20.10 where the grinding step is one of the main consumers.

Since cement is mainly a continuous process including storage constraints, it is a good candidate for applying RTN-based approaches. Various approaches are reported in Castro et al. (2011), including using a discrete-time grid with variable interval lengths. This approach can consider rather complex electricity cost/availability profiles and is rigorous for unrestricted power availability. Further

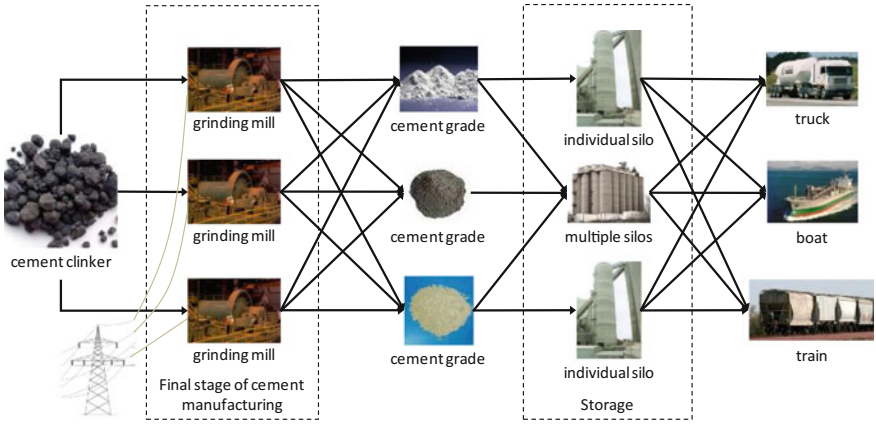

Fig. 20.10 Cement production logistics (Merkert et al. 2015)

tested approaches included the modeling of discrete events that occur at predetermined points in time with a continuous-time scheduling formulation, naturally including variable electricity costs. The conclusion was that only problems of small size can be handled effectively.

In the study it was claimed that under restricted power availability the discrete-time formulation is the best approach. The formulation is capable of finding very good solutions with few computational resources even though it cannot fully close the optimality gap. This highlights the fact that in discrete-time models discrete events can be handled in a more natural and straightforward way. The main question is whether one can use a sufficiently fine time grid to represent the problem data accurately. State-of-the-art scheduling formulations have the potential to achieve major savings when compared to procedures that are mostly focused on feasibility. The computational results showed potential cost savings of around 20 %.

20.3.4.2 Steel

Steel making is a multi-stage process, which contains a number of critical production constraints (Fig. 20.11). This makes it difficult to schedule even without electricity constraints. Here we assume that some production decisions such as caster campaign planning have been taken a priori. The most energy-relevant decisions e.g. assignment of the batches/heats to units and the timing of these operations must be determined by the optimization. It is assumed that all heats are already assigned to a particular caster.

For the steel case study the considered electricity purchasing contracts include:

- Long-term contract (base contract or base load) with a constant price and amount of electricity delivered over time.
- Short-term contract (Time-of-Use or TOU)—two price levels (on- and off-peak).
- Spot market (day-ahead)—hourly-varying prices, assumed to be known 24 h ahead.
- Onsite generation—constant price with additional start-up costs.

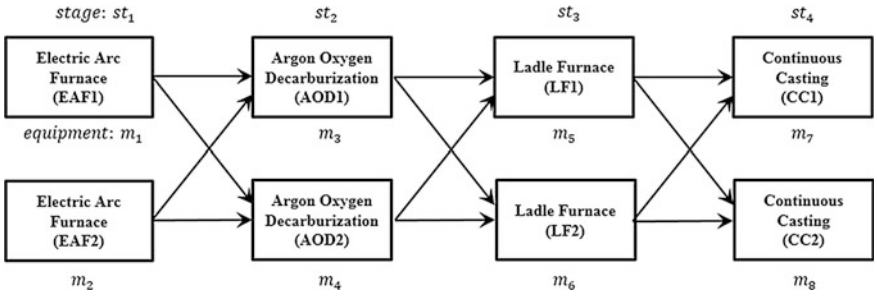

Fig. 20.11 Stainless-steel production process (melt shop section) from Hadera et al. (2015)

By using the monolithic model it was not possible to solve the problems containing 16–20 products (heats) to optimality and after 1 h (3600 s) of optimization the optimality gap (potential improvement) was still between 12 and 39 %. This resulted in net electricity costs between 83 and 146 kEUR per 24 h, which on one hand shows the significance of the energy optimization for steel processes. By using a bi-level heuristic, which splits the problem into lower and upper level problems, in most cases better results could be found in 10 minutes, the best improvements being 22 %.

Similar experiences were collected by using an RTN-based scheduling model. In comparison with a schedule, which does not take into account the electricity costs, an energy-aware model was able to reduce the electricity costs by up to 15 % for higher capacity utilization and up to 50 % for lower capacity utilization, where there are more flexibility to shift the production. This clearly indicates that fluctuating electricity prices present an important driving force for the dynamic operational behavior of the steel plant. However, these theoretical studies are not verified with real production where skilled people already perform the scheduling taking into account many of the discussed aspects. In a verified energy-aware industrial environment, an optimization approach can be expected to reduce the energy costs by 2–5 %, which is a significant cost savings for a typical steel plant.

20.3.4.3 Pulp

There are several ways how to produce pulp, the raw product for paper products. For industrial demand-side management especially the thermo-mechanical pulping process is of interest as it is very electricity-intensive.

In the thermo-mechanical pulping process wood chips are processed using heat and multiple mechanical refiners which are usually powered with electric machines. In the first production step, the wood chips are crushed and grinded in electric refiners creating heat and separating the wood fibers. Then the produced pulp is screened and clumps are reprocessed in a second refiner line. After that some less energy intensive processing steps like bleaching follow. The final pulp is then stored in a tank and can be consumed by paper machines which run continuously. Figure 20.12 gives an overview on different process steps of the thermo-mechanical pulping process.

The most energy intensive process steps in pulping process are the refiners as one refiner easily consumes several MW of electric power. In order to be prepared for outages in production most plants have tanks of different sizes between some production steps of the pulping process. These tanks, as well as, the tank for final pulp offer flexibility for demand-side management. The storage allows to decide when to run the energy-intensive refining processes in order to reduce the overall energy cost.

In a case study using a real plant setup of a pulp and paper mill in northern Europe with multiple refiner lines for pulp refiners and reject pulp refiners and electricity prices from Nord Pool Spot from 2014 and 2015, significant cost savings

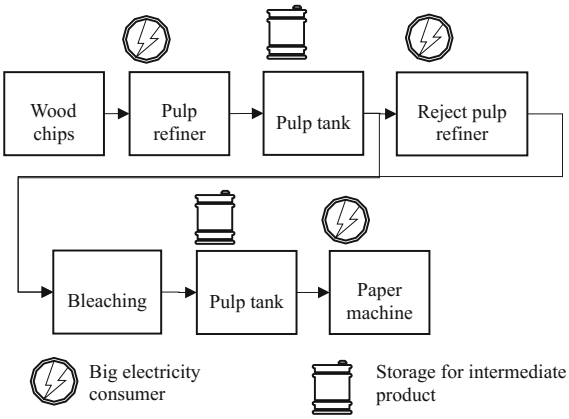

Fig. 20.12 The thermo-mechanical pulping process

have been proven. The production scheduling problem was modeled using an energy-aware Resource Task Network (see RTN approach) and an optimal production plan was calculated for different weeks in summer and winter, using a monolithic model, as well as, the iterative approach discussed above. With more smart operations of the pulp refiners, energy cost savings from 5–20 % per week have been possible compared to traditional operation. These savings were achieved using the existing pulp tanks as flexibility buffer without changing the paper machine throughput.

20.4 Conclusions

Integration of energy optimization and scheduling offers interesting potentials for load shifting. Especially in energy-intensive industries such as metals, cement and pulp and paper, energy-aware production planning can have a strong impact as energy cost can be reduced significantly. For such industries where the energy cost hold an important share of operational costs, already savings in a small percentage range can sum up to important financial gains. In this chapter, we have discussed different approaches how to leverage the cost savings potential in an efficient way. Hence, the presented work not only reflects a technical potential but also promises a high economic potential in many energy markets. All in all, these concepts can help to ensure the stability of power grids despite the growth of volatile renewable generation. From the methodological point-of-view, there are still a number of technical improvements needed for making energy-aware scheduling everyday life in the industry.

References

Ait-Ali, A. (2015). Integration of production scheduling and energy management. Master's thesis, KTH Stockholm. http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-160136.

Baboli, P. T., Moghaddam, M. P., & Eghbal, M. (2011). Present status and future trends in enabling demand response programs. *IEEE Power and Energy Society General Meeting,* 1–6.

Baldea, M., & Harjunkoski, I. (2014). Integrated production scheduling and process control: A systematic review. *Computers and Chemical Engineering, 71*, 377–390.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik, 4*, 238–252.

Castro, P. M., Harjunkoski, I., & Grossmann, I. E. (2011). Optimal scheduling of continuous plants with energy constraints. *Computers and Chemical Engineering, 35*, 372–387.

Castro, P., Sun, L., & Harjunkoski, I. (2013). Resource-task network formulations for industrial demand side management of a steel plant. *Industrial and Engineering Chemistry Research, 52* (36), 13046–13058.

Charles River Associates. (2005). Primer on demand-side management with an emphasis on price-responsive programs. Prepared for The World Bank.

Chu, Y., & You, F. (2012). Integration of scheduling and control with online closed-loop implementation: fast computational strategy and large-scale global optimization algorithm. *Computers and Chemical Engineering, 47*, 248–268.

Conti, J., & Holtberg, P. (2011). International energy outlook 2011. US Energy Information Administration.

Engell, S., & Harjunkoski, I. (2012). Optimal operation: scheduling, advanced control and their integration. *Computers and Chemical Engineering, 47*, 121–133.

European Union. (2011). Energy 2020. A strategy for competitive sustainable and secure energy.

Fernández, I., Renedo, C. J., Pérez, S. F., Ortiz, A., & Ma˜nana, M. (2012). A review: energy recovery inbatch processes. *Renewable and Sustainable Energy Reviews, 16*(4), 2260–2277.

Floudas, C. A., & Lin, X. (2004). Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. *Computers and Chemical Engineering, 28*, 2109–2129.

Gahm, C., Denz, F., Dirr, M., & Tuma, A. (2016). Energy-efficient scheduling in manufacturing companies: A review and research framework. *European Journal of Operational Research, 248*(3), 744–757.

Georgiadis, M. C., & Papageorgiou, L. G. (2001). Optimal scheduling of heat-integrated multipurpose plants under fouling conditions. *Applied Thermal Engineering, 21*(16), 1675–1697.

Hadera, H., Harjunkoski, I., Sand, G., Grossmann, I. E., & Engell, S. (2015). Optimization of steel production scheduling with complex time-sensitive electricity cost. *Computers and Chemical Engineering, 76*, 117–136.

Halim, I., & Srinivasan, R. (2009). Sequential methodology for scheduling of heat-integrated batch plants. *Industrial and Engineering Chemistry Research, 48*(18), 8551–8565.

Harjunkoski, I., Maravelias, C. T., Bongers, P., Castro, P. M., Engell, S., Grossmann, I. E., et al. (2014). Scope for industrial applications of production scheduling models and solution methods. *Computers and Chemical Engineering, 62*, 161–193.

Holmberg, K. (1992). Linear mean value cross decomposition: A generalization of the Kornai-Liptak method. *European Journal of Operational Research, 62*, 55–73.

Holmberg, K. (1994). A convergence proof for linear mean value cross decomposition. *ZOR—Methods and Models of Operations Research, 39*, 157–186.

Holmberg, K., & Kiwiel, K. C. (2006). Mean value cross decomposition for nonlinear convex problems. *Optimization Methods and Software, 21*(3), 401–417.

Kondili, E., Pantelides, C. C., & Sargent, W. H. (1993). A general algorithm for short-term scheduling of batch operations—IMILP formulation. *Computers and Chemical Engineering, 2*, 211–227.

Merkert, L., Harjunkoski, I., Isaksson, A., Säynevirta, S., Saarela, A., & Sand, G. (2015). Scheduling and energy—industrial challenges and opportunities. *Computers and Chemical Engineering, 72*, 183–198.

Olsen, D., Goli, S., & McKane, A. (2012). Examining synergies between energy management and demand response: A case study at two California industrial facilities. Lawrence Berkeley National Laboratory.

Pantelides, C. C. (1994). Unified frameworks for optimal process planning and scheduling. IN *Foundations of computer-aided process operations* (pp. 253–274). New York: CACHE Publications.

Pinedo, M., & Chao, X. (1999). *Operations scheduling with applications in manufacturing and services*. Boston: Irwin/McGraw-Hill. ISBN 0-07-289779-1.

Ribas, I., Leisten, R., & Framiñan, J. M. (2010). Review and classification of hybrid flowshop scheduling problems from a production system and a solutions procedure perspective. *Computers and Operations Research, 37*(8), 1439–1454.

Roy, V. T. J. (1983). Cross decomposition for mixed integer programming. *Mathematical Programming, 25*, 46–63.

Seid, E. R., & Majozi, T. (2014). Optimization of energy and water use in multipurpose batch plants using an improved mathematical formulation. *Chemical Engineering Science, 111*, 335–349.

Todd, D. (2011). Alcoa—dynamic demand response DOE workshop—10/25–10/26. In *Presented at 2011 Department of Energy Load Participation in Ancillary Services Workshop*, Washington, D.C. October 2011. https://www1.eere.energy.gov/analysis/pdfs/alcoa_dewayne_todd.pdf.

Wolfe, P., & Dantzig, G. B. (1960). Decomposition principle for linear programs. *Operations Research,* 101–111.

Zhuge, J., & Ierapetritou, M. G. (2012). Integration of scheduling and control with closed loop implementation. *Industrial and Engineering Chemistry Research, 51*(25), 8550–8565.

Chapter 21
Heat Integration Across Plants Considering Distance Factor

Yufei Wang and Xiao Feng

Abstract Heat integration across plants is an extension of conventional heat integration in a single plant for further improving energy efficiency. This chapter addresses the application of both Pinch Analysis and Mathematical Programming on solving heat integration problems across plants. For heat integration across plants, the required pipelines between plants is much longer than heat integration within a single plant, so more attentions must be paid on distance factor as it incurs more expense. A number of factors can affect the final design of pipelines between plants, for example, direct and indirect heat integration, the connection patterns between plants, the selection of intermediate fluid, etc. In this chapter, three connection patterns (series, split, parallel) for interconnectivity of individual plants in an area are presented. Each connection pattern has different performance on energy saving and pipeline length. To determine the energy target for the three connection patterns, a graphical methodology is presented. In addition, Mathematical Programming is used to determine the optimal design considering both direct and indirect heat integration. Parameters of intermediate fluid can be also optimized if indirect heat integration is applied. Some case studies are illustrated to demonstrate the capabilities of the presented models and graphic tool.

Keywords Heat integration · Multi-plants · Pipeline · Optimization

Y. Wang
State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing 102249, China

X. Feng (✉)
School of Chemical Engineering & Technology, Xi'an Jiaotong University, Xi'an 710049, China
e-mail: xfeng@xjtu.edu.cn

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_21

Nomenclature

Sets

HPS Set of hot process steam
STH Set of stage in plant with surplus heat
CPS Set of cold process stream
STC Set of stage in plant required heat
NP Set of intermediate fluid loops between plants

Parameters

a	$/y	Cost factor for heat exchanger area
b		Cost exponent factor for heat exchanger area
c	$/y	Fixed charge for heat exchangers
C_{cu}	$/kW · y	Utility cost coefficient for cold utility
C_{hu}	$/kW · y	Utility cost coefficient for hot utility
D_{in}	m	Inner diameter of pipe
D_{out}	m	Outer diameter of pipe
cp	J/kg · °C	Specific heat capacity
CP	kW/°C	Heat capacity flow rate
h	kW/m^2 · °C	Heat transfer coefficient
L	m	Distance between plants
$Pipe$	$/m	Pipeline cost per meter
T	°C	Extremity temperature of streams
u	m/s	Flow velocity
Wt_{pipe}	kg/m	Weight of pipe per meter
ρ	kg/m^3	Fluid density
Ω	kW	Upper bound for heat exchangers loads
Γ	°C	Upper bound for temperature difference

Variables

A	m^2	Heat transfer area of heat exchangers
dt	°C	Temperature approach
F_n	kW/°C	Heat capacity flow rate of intermediate fluid
f	kW/°C	Heat capacity flow rate for the streams after split
t	°C	Temperature
$LMTD$	°C	Log mean temperature difference
q	kW	Heat duty
P_{cul}	$/m · y	Pipeline cost
Z		Binary variable indicating the existence of the match

Subscript

i Index for hot process steam
j Index for cold process steam
n Index for intermediate fluid
k Index for stage

Superscript

H Plant with surplus heat
C Plant required heat
in Inlet
out Outlet

21.1 Introduction

Heat Integration across plants provides a number of heat recovery opportunities outside the boundary of single plants. Heat integration across plants is initially achieved through different levels of steam. High-grade heat is used to generate steam, and then the steam is transported to other plants to supply heat. A steam system is a conventional utility system widely used in industries, and the pipelines for different levels of steam are well established. For this reason, it is relatively simple to achieve heat integration across plants based on the existing steam system. The early researches on heat integration across plants were almost based on a steam system.

Total site (Dhole and Linnhoff 1993) is defined to describe a set of processes serviced by and linked through a common utility system. In this work, the ideal utility profiles can provide the balance between the net steam supply and demand, regardless of the "pockets" heat from the grand Composite Curves. A Total Site Heat Integration methodology is proposed to determine the heat transfer between processes by using different levels of steam (Hu and Ahmad 1994). In this work, the heat integration method using intermediate fluid or steam is defined as indirect heat integration and heat integration method using process streams is defined as direct heat integration. After that, Klemes et al. (1997) further developed the Total Site Profile and the Site Utility Grand Composite Curve to evaluate Total Site heat recovery potential.

Although the steam system is the most common form of recovering heat across plants, because of the physical property of steam, it cannot be used as an intermediate fluid for low-grade heat recovery. Another drawback for steam is that steam transfer latent heat at fixed temperatures and some energy saving opportunities may lose. For utilizing low-grade heat, Hammond and Norman (2014) explored the opportunities for recovering all levels of surplus heat based on a large

number of data from UK. Many options for heat recovery have been analyzed, such as on-site heat integration, off-site heat integration, heat upgrading by heat pump, using Rankine cycles to generate electrical energy and using absorption chillers to fulfill chilling demand. In their work, they pointed out that the greatest potential for reusing surplus heat was on-site and off-site heat integration. Kapil et al. (2012) mentioned that applying low-grade heat had to be with the aid of simultaneous consideration of site utility systems optimization. A wide range of low-grade heat recovery technologies, including heat pumping, organic Rankine cycles, energy recovery from exhaust gases, absorption refrigeration and boiler feed water heating are considered. From their results, boiler feed water heating is the most attractive option which is a special case of heat integration across plants. Some methodologies considering steam and hot water loops can simultaneously recover heat in both high and low level. Matsuda et al. (2009) studied Kashima industrial area in Japan by using both hot water and steam. The results show that the consumption of low pressure steam was further reduced by using hot water loops. Hackl et al. (2011) analyzed the Sweden's largest chemical cluster. Total Site Pinch methodology is used to explore the opportunities for using hot water as intermediate fluid to further reduce energy consumptions. From this research, heat integration across plants can explore more energy saving opportunities.

Other contributions expand heat integration across plants through mathematical programming approach. Rodera and Bagajewicz (1999) established energy targets for both direct and indirect heat integration based on a two plants integration case. The economy analysis showed that indirect integration using intermediate fluid circles was more economically beneficial, especially when the distance was long. They (Bagajewicz and Rodera 2000) continued their previous work and studied heat integration within a number of plants. A systematic procedure based on LP and MILP models was used to find the optimal location of the intermediate fluid circles. The cascade diagrams for each plant allowed for the detection of assisted and unassisted heat integration, which showed that both the heat transfer between pinch points and the external regions could lead to effective energy saving. After that, they (Rodera and Bagajewicz 2001) presented a method to synthesize the multipurpose heat exchanger networks with minimum number of heat exchanger units.

As mentioned (Chew et al. 2013), very long distances between heat sources and sinks are a critical feature of total site heat integration. The distance factor has been considered in many energy systems related works. For example, in heat exchanger network synthesis, Suaysompol and Wood (1993) mentioned that actual capital cost is also influenced by the network topology as this will influence the cost of pipework. In their work, a network cost estimation procedure considering pipe cost is proposed. Jiang and Chang (2013) considered the influence of different pipeline layout on total capital cost in a flexible multi-period problem. Although many works have mentioned distance factor in heat integration across plants, most works consider this factor as a fixed value, it is not involved in optimization.

To fully consider the distance factors involved in heat integration across plants, this chapter presented some works in detail. Firstly, Pinch Analysis is used to analyze the energy saving potential and distance factors for different connection patterns. Then mathematical models are presented to optimize the heat integration problem across plants.

21.2 Direct and Indirect Heat Integration

Direct and indirect heat integration are two different methods that feature different performance in heat integration across plants. For direct heat integration, process streams exchange heat with each other directly, so that the heat integration process only experience heat transfer once. This feature means that the heat recovery can be larger due to a smaller temperature difference during heat transfer between two plants. However, because process streams cannot be mixed with each other, when the number of process streams participated in heat integration is large, a number of heat transfer loops are required. Since the large distance between plants, each loop boosts the investment of pipeline. From the view of economics, if the number of process streams participated in heat integration is large, it is not beneficial for using direct integration method. Contrarily, for indirect heat integration, heat is firstly transferred from hot streams to the intermediate fluid in one plant, and then transferred from intermediate fluid to cold streams in another plant. Because of the multiple heat transfer, the total temperature difference has to be higher than direct heat integration, resulting in a lower energy recovery. Moreover, the number of heat exchangers is normally higher. The benefit for using indirect heat integration is that an intermediate fluid can be split and mixed, so that normally heat can be transported from one plant to another by one intermediate fluid loop. Pipeline investment can be largely cut down when the number of process streams participated in Heat Integration is higher.

21.3 Graphical Technique for Indirect Heat Integration Across Plant

Both direct and indirect heat integration can be solved by graphical technique such as Pinch Analysis. For direct heat integration, process streams exchange heat with each other directly, so the energy target methodology is similar with heat integration within one single plant. In this section, only graphical technique for indirect heat integration across plants is introduced, and both energy and distance factors are considered.

21.3.1 Three Basic Connection Patterns

When the distance between plants is accounted for in the heat integration across plants, it is necessary to consider the connection pattern between each heat sink and source. In this section, the situation of heat integration between three plants is considered, when the number of plants is more than three, the presented connection patterns can be used as basic connection patterns to analyze the problem (Wang et al. 2014). Figure 21.1 illustrates three possible connection patterns when there are three plants. In Fig. 21.1a, a heat source exchanges heat with two heat sinks separately, and this connection pattern allows the heat source to provide intermediate fluids with different temperatures and flow rates to the two heat sinks. However, this connection pattern requires the longest pipe length. The connection shown in Fig. 21.1b indicates a connection pattern with a medium pipe length. In this connection pattern, one intermediate fluid stream flowing out of a heat source is split into two to serve two heat sinks. With this connection pattern, the inlet temperatures of the intermediate fluid to the two heat sinks are the same, and the flow rate can be varied. The connection in Fig. 21.1c gives a connection pattern that requires the shortest pipeline. In this connection pattern, a full intermediate fluid flows from a heat source to two heat sinks consecutively. With this connection pattern, the flow rates of the intermediate fluid going to the two sinks are the same, and the intermediate fluid inlet temperature of the second heat sink almost equals the intermediate fluid outlet temperature of the first heat sink.

The three connection patterns shown in Fig. 21.1a, b and c are defined as parallel, split and series connection pattern, respectively. The potential heat sinks and sources are taken from Grand Composite Curves to consist a new Composite Curves (Linnhoff and Hindmarsh 1983) which are used to illustrate the heat feature in the three connection patterns, as shown in Fig. 21.2.

In Fig. 21.2, the gray thick lines indicate the intermediate fluid. In Fig. 21.2b, the black thick lines denote the two split intermediate fluids. From Fig. 21.2a, it can be deduced that in the parallel pattern, because of the two independent intermediate fluid loops, the two intermediate fluid lines can be combined into one intermediate fluid curve. Because the intermediate fluid curve is more likely to fit the profile of the heat source Composite Curve, it is relatively difficult to get pinched with the heat source curve, so this connection pattern can normally recover more heat than the other two patterns.

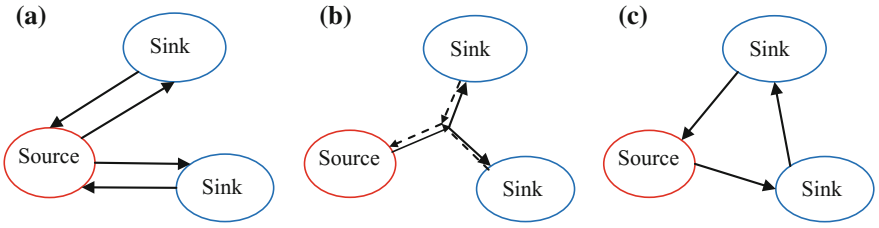

Fig. 21.1 Different connection patterns between one heat source and two heat sinks

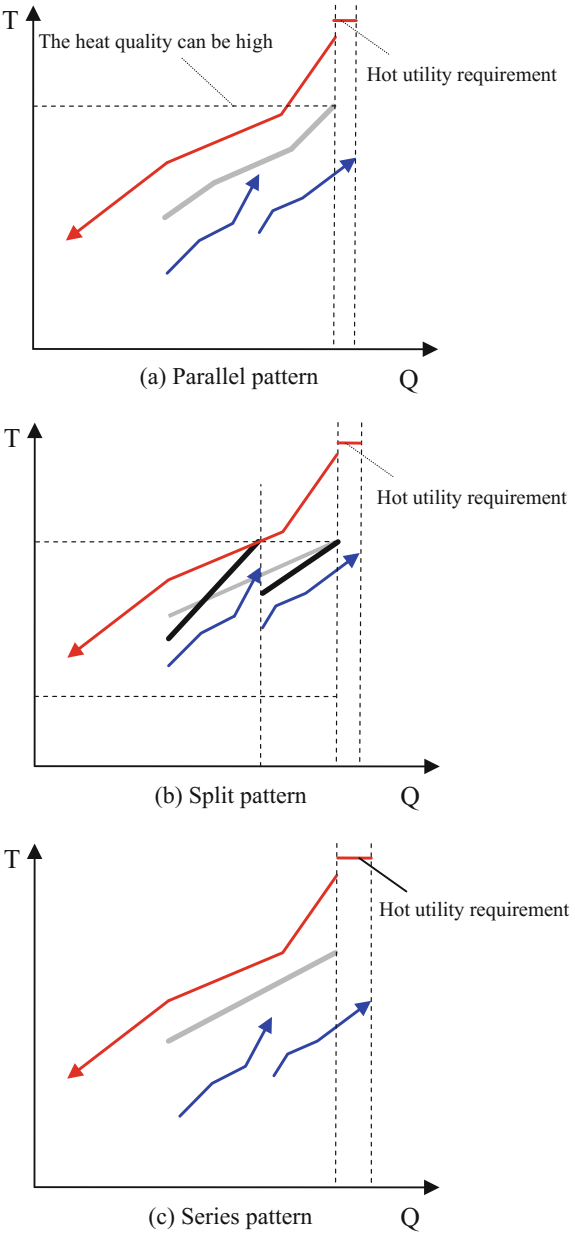

Fig. 21.2 Heat recovery in the three connection patterns

In the split pattern, the intermediate fluid is split after leaving the heat source plant. Therefore, only one intermediate fluid exchanges heat with the heat source, so that in the Composite Curve, as shown in Fig. 21.2b, there is only one intermediate fluid line. From the figure, it is obvious that a straight line is more difficult to fit the profile of the heat source Composite Curve compared with the intermediate fluid curve in the parallel pattern. In other words, it is easier to get pinched with the heat source Composite Curve compared with the situation in the parallel pattern. Therefore, the heat recovered will be in a relatively low quality and quantity range compared with that of the parallel pattern. In this pattern, because the split ratio of the main intermediate fluid can be changed, the distribution of flow rate to the two heat sinks can be adjusted according to the demand of the heat sinks. However, due to the same temperature supplied to the two sinks, some heat exchange may violate cascade utilization of heat, resulting in a waste of high quality heat.

In the series pattern, there is also only one intermediate fluid that exchanges heat with the heat source, so that the recovered heat will be in a relatively low quality and quantity range. In this pattern, the flow rate of the intermediate fluid going to the two sinks is the same, and the outlet intermediate fluid temperature of the heat sink with higher temperature almost equals the inlet intermediate fluid temperature of the heat sink with lower temperature. From the figure, it can be seen that the heat sink Composite Curves and intermediate fluid Composite Curve are more likely to get pinched compared with the other two patterns, resulting in a lower heat recovery.

Figure 21.2 shows that the two heat sink Composite Curves are not combined. For a conventional Composite Curve in a single plant, it is assumed that all the cold or hot streams in the same temperature interval can be mixed to form one stream so that within one network all the hot or cold streams can be combined into one Composite Curve. However, in heat integration across plants, this assumption cannot be used since long distances between plants do not permit streams from different plants in the same temperature interval to be combined.

21.3.2 Determination Energy Target and the Flow Rate of Intermediate Fluid in Heat Integration Between Two Plants

The difference between an intermediate fluid (hot water or hot oil) and steam on determining the energy target is that on the Composite Curve the line for an intermediate fluid is an oblique line while the line for steam is a horizontal line. Therefore, in heat integration across plants using an intermediate fluid, it is important to determine the temperature and flow rate of the intermediate fluid in order to determine the energy target. To determine the flow rate, for simplicity, a site containing only one heat sink plant and one heat source plant is analyzed. In indirect total site integration, twice heat transfer occurs: heat is taken from heat sources and

sent to heat sinks by the intermediate fluid. In this work, ΔT_{min} is defined as the minimum approach temperature between the intermediate fluid and a heat sink/source. Thus, the minimum approach temperature between a heat sink and a heat source should be twice the T_{min} value due to the twice heat transfer. Once T_{min} is defined, the temperature difference between the intermediate fluid and heat sinks as well as that between the intermediate fluid and heat sources must not violate T_{min}.

In pinch technology the pinch point indicates the place where the temperature difference between hot and cold streams equals the minimum approach temperature ($2\times T_{min}$). Therefore, it can be deduced that the intermediate fluid line must go through the pinch point to ensure that the minimum temperature difference between the intermediate fluid and heat sinks/sources is not smaller than T_{min}.

Figure 21.3 is used as an example to show how the flow rate of an intermediate fluid can be determined. In a Composite Curve plot, the gradient of a line represents the heat capacity flow rate (CP) which is equal to the product of specific heat and flow rate. The flow rate of an intermediate fluid can thus be determined from the gradient of the line on a Composite Curve. Figure 21.3 shows a plot of Shifted Composite Curves which are obtained by subtracting T_{min} from the temperatures of all the hot streams, adding T_{min} to the temperatures of all the cold streams, and keeping the temperature of the intermediate fluid unchanged. In Fig. 21.3, the two thick curves indicate the heat source Composite Curve (upper one) and the heat sink Composite Curve (lower one), while the thin gray line denotes the intermediate fluid line. Heat recovery region is defined as the region within which heat can be transferred from a heat source to a heat sink on the Composite Curve. Because no heat transfer occurs outside the heat recovery region, the intermediate fluid line should not extend beyond the boundaries of the region.

Normally, a pinch point is formed by a convex point on one particular curve touching a straight line on the other. As shown in Fig. 21.3a, the pinch point is formed by the convex point on the heat sink curve touching a straight line on the

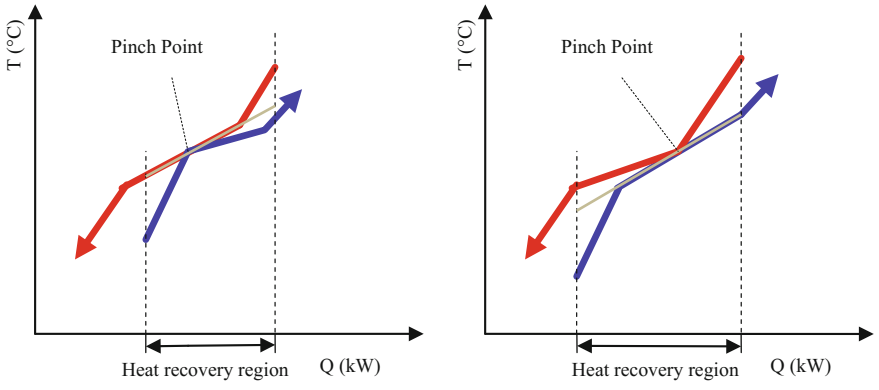

Fig. 21.3 Intermediate fluid flow rate determination through **a** hot stream **b** cold stream

heat source curve. In Fig. 21.3a, b, it is obvious that when the *CP* of the intermediate fluid line equals the *CP* of the straight line forming the pinch point, maximum heat recovery can be achieved. However, maximum heat recovery is also possible under two situations in which a range of *CP* values of the intermediate fluid exist:

(1) The pinch point is formed by a convex point on one curve touching another convex point on the other.
(2) The problem is a threshold problem.

In the situations described in Fig. 21.3, it is straightforward to determine the *CP* of the intermediate fluid, and the maximum heat recovery equals the energy target achieved by the pinch approach. However, in some complicated situations, the *CP* of the intermediate fluid that can achieve maximum heat recovery may be somewhat difficult to determine and the maximum heat recovery may not equal the energy target achieved by the Pinch Approach. One such situation is shown in Fig. 21.4. In this figure, the gray line denotes the *CP* of the intermediate fluid that can achieve maximum heat recovery. It is evident that the *CP* of the intermediate fluid line no longer coincides with the straight line forming the pinch point. Therefore, the intermediate fluid CP cannot be determined simply through the method used in Fig. 21.3. Nonetheless, it gives us a clue for developing a feasible method, as described below.

In this chapter, it is assumed that only one intermediate fluid is used to transfer heat from a heat source plant to a heat sink plant. Therefore, the intermediate fluid line must be a straight line and not a curve on Composite Curves. So normally, complicated Composite Curve profiles can be simplified. For the heat source plant, the simplified curve is a curve connecting three points on the heat source Composite Curve that make the entire heat source Composite Curve above it within the heat transfer region. For the heat sink plant, the simplified curve is a curve connecting three points on the heat sink Composite Curve that make the entire heat sink Composite Curve below it within the heat transfer region, as shown in Fig. 21.5a.

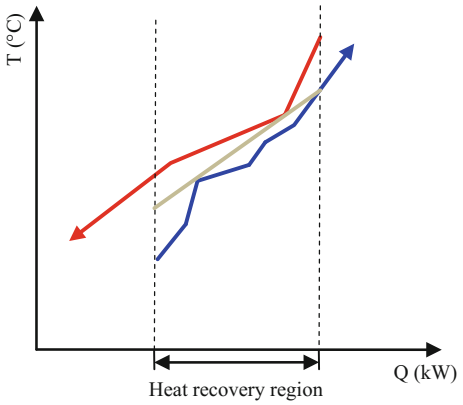

Fig. 21.4 Determination of intermediate fluid heat capacity flow rate in a complicated situation

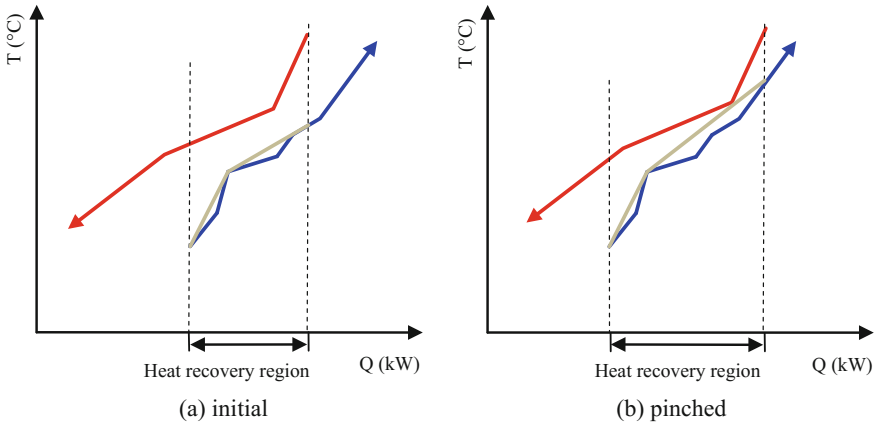

Fig. 21.5 The determination procedure of CP in simplified Composite Curves

The gray curves in Fig. 21.5a are the simplified heat sink Composite Curves. These three points are always the convex points or the end points on the Composite Curve. Through this simplification, the Composite Curve can be substituted by the simplified Composite Curve. The simplified curve only has one turning point on it. It can guarantee the temperature difference between the simplified Composite Curve and the other Composite Curve is always larger than the minimum approach temperature when the simplified Composite Curve gets pinched with the other Composite Curve.

It is noted that during the pinching movement the heat recovery region becomes larger, so that the three points defining the simplified Composite Curve may be changed in the region, as shown in Fig. 21.5b. Therefore, the gradient of the simplified Composite Curve will change during pinching. When the simplified Composite Curve is pinched with the other Composite Curve, the method used in Fig. 21.3 is applicable and the *CP* of the pinched line is the *CP* of the intermediate fluid that can achieve maximum heat recovery, as shown by the situation in Fig. 21.5.

From the discussion above, we can formulate a graphically based heuristic methodology for determining the maximum heat recovery and the *CP* of the intermediate fluid that can achieve maximum heat recovery in indirect total site integration between two plants. The salient features of the methodology are given below.

(1) Find the pinch point, and then find the straight line and the convex point that form the pinch point.
(2) Substitute the simplified Composite Curve for the Composite Curve containing the straight line that forms the pinch point.
(3) Re-pinch the simplified Composite Curve and the other Composite Curve to find the new pinch point. The maximum heat recovery can then be determined

and the *CP* of the intermediate fluid equals the *CP* of the line forming the pinch point on the simplified Composite Curve.

21.3.3 Determination of Maximum Heat Recovery in Indirect Heat Integration for Different Connection Patterns

In this section, the situation of three plants is analyzed. It is assumed that the site comprises two heat sink plants and one heat source plant. Note that a site comprising one heat sink and two heat sources can be analyzed in the same way. The three connection patterns for three plants described in Fig. 21.1 are considered.

21.3.3.1 Determination of the Energy Target for Parallel Connection Pattern

To determine the energy target of the total site containing three plants connected in the parallel mode (Fig. 21.1a), firstly, the simplified Composite Curves are applied to both heat sink plants. The simplified Composite Curve for each heat sink plant is moved toward the Composite Curve of the heat source plant to locate the pinch point for each heat sink plant. Two intermediate fluid lines can then be found according to the energy target method developed for the case of two plants.

Next, the two individual intermediate fluid lines are used to generate the intermediate fluid curve according to the method used to generate Composite Curves in conventional Pinch Technology. Finally, the energy target can be found by pinching the intermediate fluid Composite Curve and the heat source Composite Curve. The procedure described above is shown graphically in Fig. 21.6.

21.3.3.2 Determination of the Energy Target for Split Connection Pattern

The split connection pattern (Fig. 21.1b) is a special case of the parallel connection pattern. A major difference between them is that in the case of the split pattern the temperature of intermediate fluids supplied to the two heat sink plants are the same. Also, the two intermediate fluids leaving the two heat sink plants are re-mixed to form a single stream that flows back to the heat source plant. In determining the intermediate fluid line step, for the heat sink plant with lower temperature (sink 1), the hot end temperature is adjusted to be equal to the hot end temperature of the

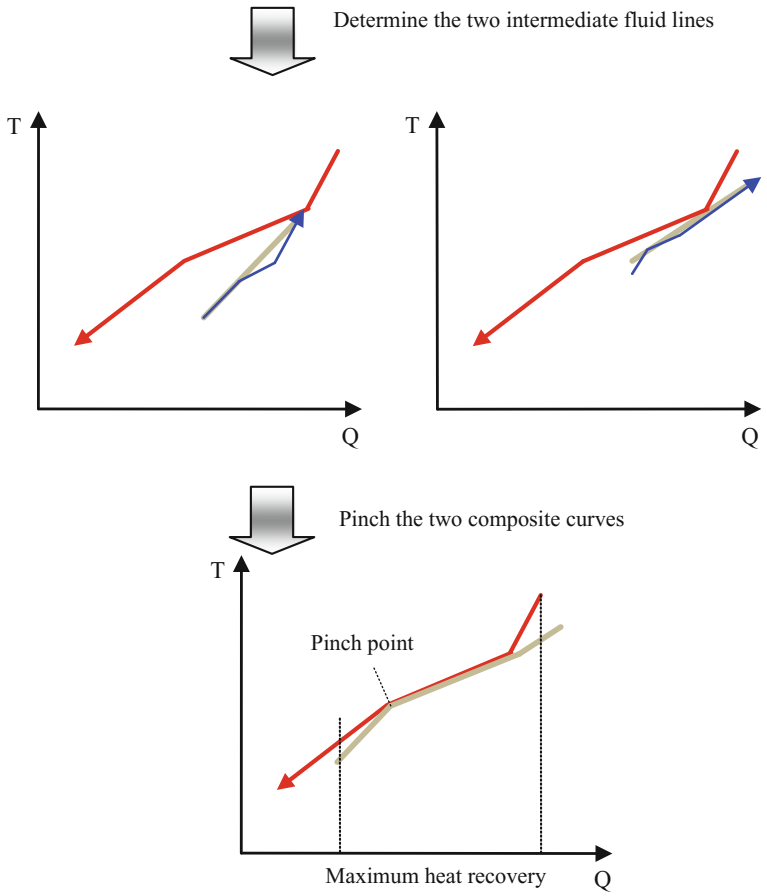

Fig. 21.6 Procedure for determining the energy target for parallel connection pattern

heat sink plant with higher temperature (sink 2). In Fig. 21.7a, this step illustrates as from dashed gray line to solid gray line. By doing this, the temperatures of intermediate fluids supplied to the two heat sink plants are the same and the supply temperature is high enough to satisfy the two heat sink plants. The two intermediate fluid lines (two thin lines in Fig. 21.7c) are then mixed to generate the intermediate fluid line (the thick line in Fig. 21.7c) for the split connection pattern. Finally, the energy target can be obtained by pinching the intermediate fluid line and heat source Composite Curve. The procedure is shown graphically in Fig. 21.7.

The main differences between the split and parallel connection patterns in determining the energy target on Composite Curves are summarized below.

(1) In the parallel pattern, the intermediate fluid is represented as a curve but in the split pattern on Composite Curves, it is represented as a line.

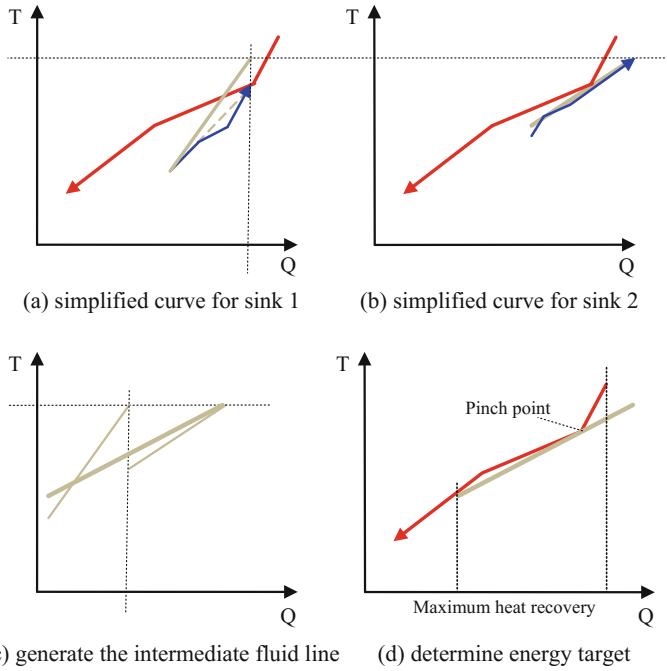

Fig. 21.7 Procedure for determining the energy target for split connection pattern

(2) The split pattern requires more high temperature heat for heat sink 1. As a result, it is easier for the intermediate fluid to get pinched with the heat source Composite Curve.

21.3.3.3 Determination of the Energy Target for Series Connection Pattern

In the series connection pattern (Fig. 21.1c), the two heat sink Composite Curves share one intermediate fluid. Therefore, one simplified Composite Curve can be used to represent the two heat sink Composite Curves, as shown by the gray line in Fig. 21.8.

After generating the simplified Composite Curve for the two heat sink plants, the energy target can be obtained by pinching the simplified Composite Curve and the heat source Composite Curve according to pinch technology. However, it is found that if a part of the pinched stream for heat sink plant 1 is moved out of the heat recovery region (this part of the heat sink is heated by utility), the total site can recover more heat from heat sink plant 2, as shown in Fig. 21.8a. Therefore, using only pinch technology will not find the real energy target for the series connection

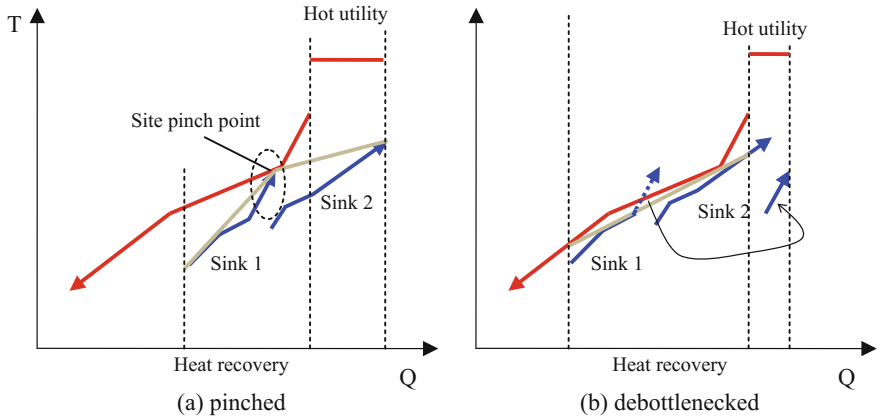

Fig. 21.8 Determination of the energy target for series connection pattern and the existence of site pinch

pattern in total site heat integration. This pinch point that cannot determine the real energy target is defined as site pinch point in this work.

The failure of the site pinch point to determine the energy target is due to the fact that in the Composite Curve the two heat sink Composite Curves are not combined and that they share one heat source curve. As a result, in some regions in the heat source curve the low temperature part in the heat source curve needs to exchange heat with the high temperature part in the heat sink curve, leading to a local heat transfer bottleneck. In other words, the heat recovery in heat sink plant 2 is constrained by the heat recovery in heat sink plant 1. So the site pinch point is always located on the heat sink with lower temperature (sink 1). The concept of site pinch is similar to the network pinch proposed by Asante and Zhu (1996) in heat exchanger network retrofit. In the network pinch approach, the stream lines in an

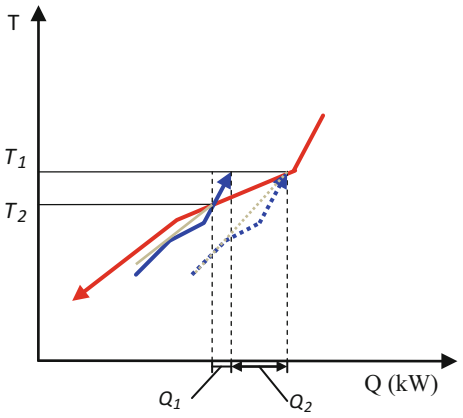

Fig. 21.9 The benefit brought by debottlenecking site pinch

existing network cannot be combined to form one Composite Curve on Composite Curve, resulting in local heat transfer bottlenecks.

Figure 21.9 shows how the site pinch point can be debottlenecked by moving part of the Composite Curve of heat sink 1 out of the heat recovery region. In this figure, only heat sink 1 is drawn. As shown in Fig. 21.9, the Composite Curve of the heat sink is moved from the dashed curve position to the solid curve position. The gray line is the simplified Composite Curve for the heat sink plant. From the figure, it is obvious that the part of the Composite Curve requiring hot utility (Q_1) is shorter than the distance covered by the movement of the heat sink Composite Curve to the left (Q_2). Also, it can be deduced from the figure that:

$$\begin{cases} Q_1 = (T_1 - T_2) \cdot CP_c \\ Q_1 + Q_2 = (T_1 - T_2) \cdot CP_h \end{cases} \tag{1}$$

where CP_c and CP_h are heat capacity flow rate of the cold and hot streams that get crossed on composite curve, and T_1 and T_2 are the temperature at original site pinch point and temperature at the crossed point of two curves, respectively. From Eq. 1, if a larger heat recovery is desired after debottlenecking, Q_1 must be smaller than Q_2. Equation 1 indicates that the relation between CP_h and CP_c must satisfy the following condition.

$$CP_h > 2 \cdot CP_c \tag{2}$$

As mentioned, the site pinch must be on the composite curve of the heat sink with lower temperature (sink 1). Accordingly, for the series connection pattern, if the most convex point is not on the composite curve of heat sink 1, there is no site pinch point.

The above analysis suggests that heat recovery of total site for the case of series connection pattern can be increased by moving part of the heat sink composite curve out until one of the conditions listed below is met.

(1) Equation 2 is not satisfied.
(2) All the heat requirement of heat sink 2 is satisfied by the heat source (the problem becomes a threshold problem).
(3) The most convex point does not lie on the composite curve of heat sink 1.

Site pinch point indicates the unreasonable heat transfer between heat sinks and heat sources. After debottlenecking the site pinch, the maximum heat recovery can be readily obtained.

The three connection patterns are the basic connection patterns that are used in indirect total site heat integration. For a total site containing more than three plants, the method used for these connection patterns can also be used to determine the energy target.

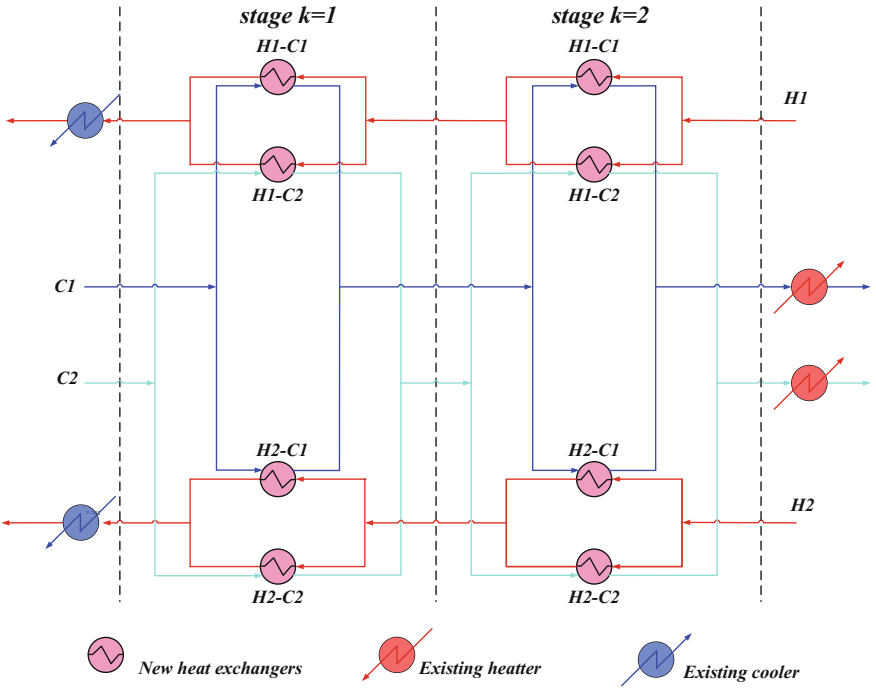

Fig. 21.10 The heat exchanger network superstructure proposed by Yee and Grossmann (1990)

21.4 Mathematical Programming Methodology for Direct and Indirect Heat Integration Across Plants

For direct and indirect heat integration across plants, Mathematical Programming can be used to make a detailed trade-off between energy cost and capital cost. A two plants heat integration model considering distance factors are presented (Wang et al. 2015).

For direct heat integration, process streams exchange heat without using intermediate fluids. It is assumed that the inner heat exchanger network in each plant has been well established, and so only the streams with surplus heat in heat source plant and the streams that required heat in heat sink plant are considered to be integrated. Therefore, the problem can be considered as a heat exchanger network with a distance between hot streams in one plant and cold streams in the other. The heat exchanger network superstructure proposed by Yee and Grossmann (1990) is used, as shown in Fig. 21.10. To distinguish the difference between heat exchanger network and heat integration between plants, the cost of pipeline is counted in. Moreover, as the plants are existed, the cost of utility heat exchanger does not need to be considered. Therefore, the objective function for direct integration can be written as below:

$$TAC = \min\left(C_{cu} \cdot \sum_{i \in HPS} q_{cui} + C_{hu} \cdot \sum_{j \in CPS} q_{huj} + a \sum_{i \in HPS} \sum_{j \in CPS} \sum_{k \in ST} z_{ijk} + b \sum_{i \in HPS} \sum_{j \in CPS} \sum_{k \in ST} A_{ijk}^{c} + 2L \cdot \sum_{i \in HPS} Pipe_i\right) \tag{3}$$

In Eq. 3, *HPS* and *CPS* are two sets for hot streams and cold streams, i and j are two index for hot and cold streams, *ST* is a set for stage, k is index for stage, C_{cu} and C_{hu} are the price of hot and cold utility, q is heat duty, q_{cu} and q_{hu} are the duty of cold utility heat exchangers and hot utility heat exchangers, a, b and c are cost constants for heat exchanger area cost, z is a binary variable indicate the existence of heat exchanger, A is the area for heat exchanger, *Pipe* is the pipe cost per meter including capital cost and installation cost, and L is the distance between plants.

From Eq. 3, it can be seen that the cost of pipeline is involved, and the cost of pipeline is calculated through the weight of pipe, which is determined by inner diameter of pipes. The inner diameter of pipes is calculated by Eq. 4.

$$D_{in} = \sqrt{\frac{4CP}{\pi \cdot \rho \cdot cp \cdot u}} \tag{4}$$

where cp is specific heat capacity, ρ is fluid density, and u is flow velocity.

In direct heat integration, the flowrate of each process stream is known. However, in indirect heat integration, the flowrate of the intermediate fluid is not known until the final design is determined. It can be known that the flowrate of the intermediate fluid is an important variable because it is correlated with diameter of pipeline and heat capacity flowrate. With large flowrate, the energy recovery can be larger with a higher pipeline cost (Chang et al. 2015). Therefore, a tradeoff between pipeline investment and energy cost should be involved. It is considered in this mathematical model.

For indirect heat integration for heat integration across plants, because of the use of intermediate fluids between plants, the superstructure for the problem is different, as shown in Fig. 21.11. In the figure, H1 and H2 indicate two hot streams, C1 and C2 indicate two cold streams, and N1 and N2 indicate two intermediate fluids, respectively. In the figure, the stage is further divided into two: k^c and k^h. The stage k^c indicates the heat integration between cold streams and intermediate streams and stage k^h indicates the heat integration between hot streams and intermediate fluid. The mathematical model for indirect heat integration includes all the possible connections. In the figure, it can be seen that two intermediate fluids exchange heat with each stream in each stage. Different from normal superstructure for indirect heat integration, this superstructure can involve more than one intermediate fluid loops.

Based on the superstructure, the overall heat balance shown in Eqs. 5 and 6 are used to ensure sufficient heating or cooling for each process steam to get its target temperature.

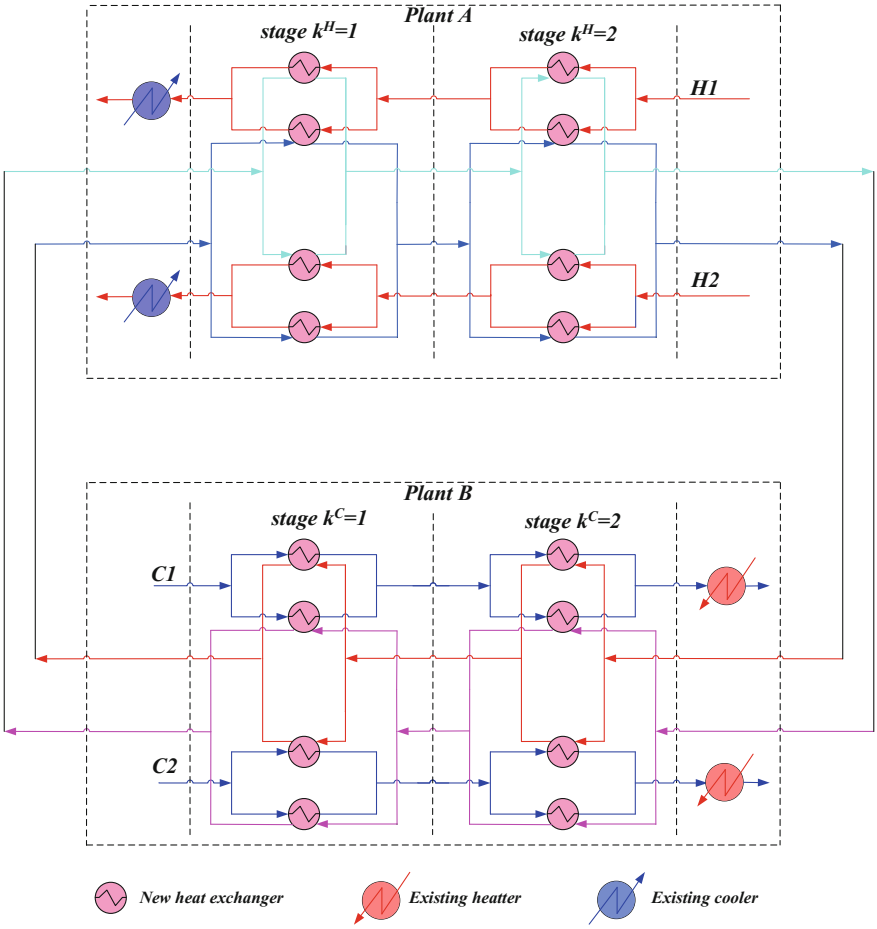

Fig. 21.11 The superstructure for indirect heat integration

$$\sum_{k^H \in STH} \sum_{n \in NP} q^H_{ink^H} + q_{cu_i} = CP^H_i \cdot \left(T^{in}_i - T^{out}_i\right) \quad i \in HPS \tag{5}$$

$$\sum_{k^C \in STC} \sum_{n \in NP} q^C_{njk^C} + q_{hu_j} = CP^C_j \cdot \left(T^{out}_j - T^{in}_j\right) \quad j \in CPS \tag{6}$$

For each split, mass balances are used to determine the heat capacity flow rates of each heat exchanger, as shown in Eqs. 7–10.

$$\sum_{n \in NP} f^H_{ink^H} = CP^H_i \quad i \in HPS, k^H \in STH \tag{7}$$

$$\sum_{i \in HPS} f^N_{ink^H} = CP_n \quad n \in NP, k^H \in STH \tag{8}$$

$$\sum_{j \in CPS} f^N_{njk^C} = CP_n \quad n \in NP, k^C \in STC \tag{9}$$

$$\sum_{n \in NP} f^C_{njk^C} = CP^C_j \quad j \in CPS, k^C \in STC \tag{10}$$

where f is a variable indicates the heat capacity flow rate in each branch, for example, $f^H_{ink^H}$ indicates the heat capacity flow rate between hot stream i and intermediate fluid n in the stage k^H. In the equations, STH is the set for stage in the plant providing heat and STC indicates the set for stage in plant receiving heat, k^H and k^C are two index for sets STH and STC, NP is a set for intermediate fluids, n is the index for NP, T^{in}_i and T^{out}_i are parameters that indicate inlet and outlet temperatures of hot stream i, T^{in}_j and T^{out}_j are parameters that indicate inlet and outlet temperatures of cold stream j, and $q^H_{ink^H}$ means the heat duty between hot stream i and intermediate fluid n in stage k^H, respectively.

Energy balance around each additional heat exchanger is performed in order to determine the outlet temperature of the heat exchanger, which leads to equations with bilinear terms, as shown in Eqs. 11–14. In these equations, t is a variable indicates temperature.

$$q^H_{ink^H} = f^H_{ink^H} \cdot \left(t_{ik^H} - t^{out}_{ink^H}\right) \quad i \in HPS, n \in NP, k^H \in STH \tag{11}$$

$$q^H_{ink^H} = f^N_{ink^H} \cdot \left(t^{out}_{ik^H} - t_{k^H+1}\right) \quad i \in HPS, n \in NP, k^H \in STH \tag{12}$$

$$q^C_{njk^C} = f^N_{njk^C} \cdot \left(t_{nk^C} - t^{out}_{njk^C}\right) \quad n \in NP, j \in CPS, k^C \in STC \tag{13}$$

$$q^C_{njk^C} = f^C_{njk^C} \cdot \left(t^{out}_{njk^C} - t_{nk^C+1}\right) \quad n \in NP, j \in CPS, k^C \in STC \tag{14}$$

Energy balance around each mixer determines the inlet temperatures of stages, which also leads to equations with bilinear terms, as shown in Eqs. 15–18.

$$CP^H_i \cdot t_{ik^H+1} = \sum_{n \in NP} f^H_{ink^H} \cdot t^{out}_{ink^H} \quad i \in HPS, k^H \in STH \tag{15}$$

$$CP_n \cdot t^N_{nk^h} = \sum_{i \in HPS} f^N_{ink^H} \cdot t^{out}_{ink^H} \quad n \in NP, k^H \in STH \tag{16}$$

$$CP_n \cdot t^N_{nk^C+1} = \sum_{j \in CPS} f^N_{njk^C} \cdot t^{out}_{njk^c} \quad n \in NP, k^C \in STC \tag{17}$$

$$CP^C_j \cdot t_{jk^C} = \sum_{n \in NP} f^C_{njk^C} \cdot t^{out}_{njk^C} \quad j \in CPS, k^C \in STC \tag{18}$$

According to the superstructure, the assignment of the inlet temperatures is as follows:

$$t^{H}_{ik^H} = T^{in}_{i} \quad i \in HPS, k^H \in Firsth \tag{19}$$

$$t^{N}_{nk^H} = T^{H,out}_{n} \quad n \in NP, k^H \in Firsth \tag{20}$$

$$t^{N}_{nk^H} = T^{H,in}_{n} \quad n \in NP, k^H \in Lasth \tag{21}$$

$$t^{N}_{nk^C} = T^{C,out}_{n} \quad n \in NP, k^C \in Lastc \tag{22}$$

$$t^{N}_{nk^C} = T^{C,in}_{n} \quad n \in NP, k^C \in Firstc \tag{23}$$

$$t^{C}_{jk^c} = T^{in}_{j} \quad j \in CPS, k^C \in Lastc \tag{24}$$

where *Firsth* and *Lasth* indicate the first and last stage in the plant that provides heat, and *Firstc* and *Lastc* indicate the first and last stage in the plant that requires heat.

Energy balances for final utility units determine the utility loads, and the equations for solving utility heat load are shown below:

$$q_{cu_j} = CP^{H}_{i} \cdot \left(t^{H}_{ik^H} - T^{out}_{i}\right) \quad i \in HPS, k^H \in Lasth \tag{25}$$

$$q_{hu_j} = CP^{C}_{j} \cdot \left(T^{out}_{j} - t^{C}_{jk^C}\right) \quad j \in CPS, k^C \in Firstc \tag{26}$$

The temperature constraints should ensure feasibility of temperatures, and they specify monotonic decreases in the temperatures along the stages:

$$t_{ik^H+1} \le t_{ik^H} \quad i \in HPS, k^H \in STH \tag{27}$$

$$t_{nk^H+1} \le t_{nk^H} \quad n \in NP, k^H \in STH \tag{28}$$

$$t_{ik^H} \ge T^{out}_{i} \quad i \in HPS, k^H \in Lasth \tag{29}$$

$$t_{nk^C+1} \le t_{nk^C} \quad n \in NP, k^C \in STC \tag{30}$$

$$t_{jk^C+1} \le t_{jk^C} \quad j \in CPS, k^C \in STC \tag{31}$$

$$t_{jk^c} \le T^{out}_{j} \quad j \in CPS, k^C \in Firstc \tag{32}$$

Upper bound constrains are needed to relate the heat duty q with the binary variables z:

$$q^H_{ink^H} - \Omega^H_{ink^H} \cdot z^H_{ink^H} \le 0 \quad i \in HPS, n \in NP, k^H \in STH \tag{33}$$

$$q^C_{njk^C} - \Omega^C_{njk^C} \cdot z^C_{njk^C} \le 0 \quad n \in NP, j \in CPS, k^C \in STC \tag{34}$$

where Ω is the upper bound for heat exchangers duty, and z is a binary variable that determines the existence of the match.

Logical constrains are needed to ensure that the temperature difference can be calculated if the heat exchanger exists. The parameter Γ is an upper bound for the temperature difference, and dt is temperature difference.

$$dt^H_{ink^H} \le t_{ik^H} - t^{out}_{ink^H} + \Gamma^H_{ink^H} \cdot \left(1 - z^H_{ink^H}\right) \quad i \in HPS, n \in NP, k^H \in STH \tag{35}$$

$$dt^H_{ink^h+1} \le t^{out}_{ink^H} - t_{ink^H+1} + \Gamma^H_{ink^H} \cdot \left(1 - z^H_{ink^H}\right) \quad i \in HPS, n \in NP, k^H \in STH \tag{36}$$

$$dt^C_{njk^C} \le t_{jk^C} - t^{out}_{njk^C} + \Gamma^C_{njk^C} \cdot \left(1 - z^C_{njk^C}\right) \quad n \in NP, j \in CPS, k^C \in STC \tag{37}$$

$$dt^C_{njk^C+1} \le t^{out}_{njk^C} - t_{jk^C+1} + \Gamma^C_{njk^C} \cdot \left(1 - z^C_{njk^C}\right) \quad n \in NP, j \in CPS, k^C \in STC \tag{38}$$

The trade-off between investment cost and operating cost are considered by adding the following constraints:

$$dt^H_{ink^H} \ge \Delta T_{min} \quad i \in HPS, n \in NP, k^H \in STH \tag{39}$$

$$dt^C_{njk^C} \ge \Delta T_{min} \quad n \in NP, j \in CPS, k^C \in STC \tag{40}$$

The driving forces are calculated by the logarithmic mean temperature difference (*LMTD*) for each heat exchanger:

$$LMTD^H_{ijk^H} = \left[dt^H_{ijk^H} \cdot dt^H_{ijk^H+1} \cdot \frac{dt^H_{ijk^H} + dt^H_{ijk^H+1}}{2} \right]^{\frac{1}{3}} \quad i \in HPS, n \in NP, k^H \in STH \tag{41}$$

$$LMTD^C_{njk^C} = \left[dt^C_{njk^C} \cdot dt^C_{njk^C+1} \cdot \frac{dt^C_{njk^C} + dt^C_{njk^C+1}}{2} \right]^{\frac{1}{3}} \quad n \in NP, j \in CPS, k^C \in STC \tag{42}$$

The area of process heat exchanges can be calculated by the following equations:

$$A_{ink^H} = q^H_{ink^H} \cdot \left(\frac{1}{h^H_i} + \frac{1}{h_n} \right) / LMTD^H_{ink^H} \quad i \in HPS, n \in NP, k^H \in STH \tag{43}$$

$$A_{njk^C}^{C} = q_{njk^C}^{C}\left(\frac{1}{h_j^C}+\frac{1}{h_n}\right)/LMTD_{njk^C}^{C} \quad n \in NP, j \in CPS, k^C \in STC \tag{44}$$

where A is the area of a heat exchanger, and h is heat transfer coefficient.

Finally, the objective function is the total annual cost (TAC) and is shown in Eq. 43.

$$TAC = \min\left(\begin{array}{l} C_{cu}\cdot\sum\limits_{i\in HPS} q_{cui} + a\sum\limits_{i\in HPS}\sum\limits_{n\in NP}\sum\limits_{k\in STH} z_{nk^H}^{H} + b\sum\limits_{i\in HPS}\sum\limits_{n\in NP}\sum\limits_{k\in STH} A_{ink^H}^{H} + 2L\cdot\sum\limits_{n\in NP} Pipe_n \\ \quad + C_{hu}\cdot\sum\limits_{j\in CPS} q_{huj} + a\sum\limits_{n\in NP}\sum\limits_{j\in CPS}\sum\limits_{k^C\in STC} z_{jk}^{C} + b\sum\limits_{n\in NP}\sum\limits_{j\in CPS}\sum\limits_{k^C\in STC} A_{njk^C}^{c} \end{array}\right) \tag{45}$$

By solving the mathematical model above, the optimal heat integration design can be obtained.

21.5 Case Study

The case study is a heat integration project between a refinery and a rubber plant. The refinery is the plant that supplies heat and the rubber plant is the plant that requires heat. The stream data are shown in Table 21.1. The distance between the two plants is 650 m. In this case, the prices for cold and hot utilities are 20 \$ $kW^{-1}y^{-1}$ and 190 $\$\cdot kW^{-1}y^{-1}$ (Hipólito-Valencia et al. 2014), respectively. The equation for calculating exchanger area cost is shown in Eq. 46 (Kovač Kralj et al. 2005), the cost factor in the equation is updated to fit current price of heat exchangers. The pipe line cost is calculated through the weight of 40 sch (a type of carbon steel) pipeline, and the equations are shown in Eqs. 47–48 (Stijepovic and Linke 2011). In Eq. 49, D_{out} and D_{in} are outer diameter and inner diameter of pipe,

Table 21.1 Stream data for the case study

Streams	Tin/°C	Tout/°C	ΔH/kW	F/(kW · °C^{-1})	h/(W · m^{-2} · °C^{-1})
H1 (refinery)	148	122	6,000	230.8	969
H2 (refinery)	130	100	2,300	76.7	599
H3 (refinery)	105	75	1,600	53.3	621
H4 (refinery)	95	60	1,200	34.3	657
C1 (rubber plant)	112	137	1,500	60	667
C2 (rubber plant)	111	136	1,500	60	665
C3 (rubber plant)	110	136	1,500	57.7	668
C4 (rubber plant)	109	134	1,500	60	667
C5 (rubber plant)	90	120	2,000	66.7	772
C6 (rubber plant)	55	98	2,100	48.8	651
C7 (rubber plant)	50	75	2,200	88	598

and Wt_{pipe} is the weight of pipe per meter in Eq. 48. In Eq. 49, I indicates the depreciation factor (20 %), y is the service life of pipe. In this case, y is 3 years. The intermediate fluid used in this case study is hot water. The density of water is 950 kg/m^3. The flow velocity u for hot water is set to 1.2 m/s. The ΔT_{min} in this work is set to 8 °C. The case study is solved less than 2 min of CPU time on a desktop PC (Inter (R) Core (TM) i5 CPU 3.33 GHz, with 4.00 GB of RAM) using GAMS24.21. The MINLP solver is Dicopt, while the NLP solver is Conopt.

$$Cost\ of\ Heat\ Exchangers\left(\$ \cdot y^{-1}\right) = 3000 + 150\left[Area\left(m^2\right)\right]^1 \tag{46}$$

$$D_{out}(m) = 1.052D_{in} + 0.005251 \tag{47}$$

$$Wt_{pipe}\left(kg \cdot m^{-1}\right) = 644.3D_{in}^2 + 72.5D_{in} + 0.4611 \tag{48}$$

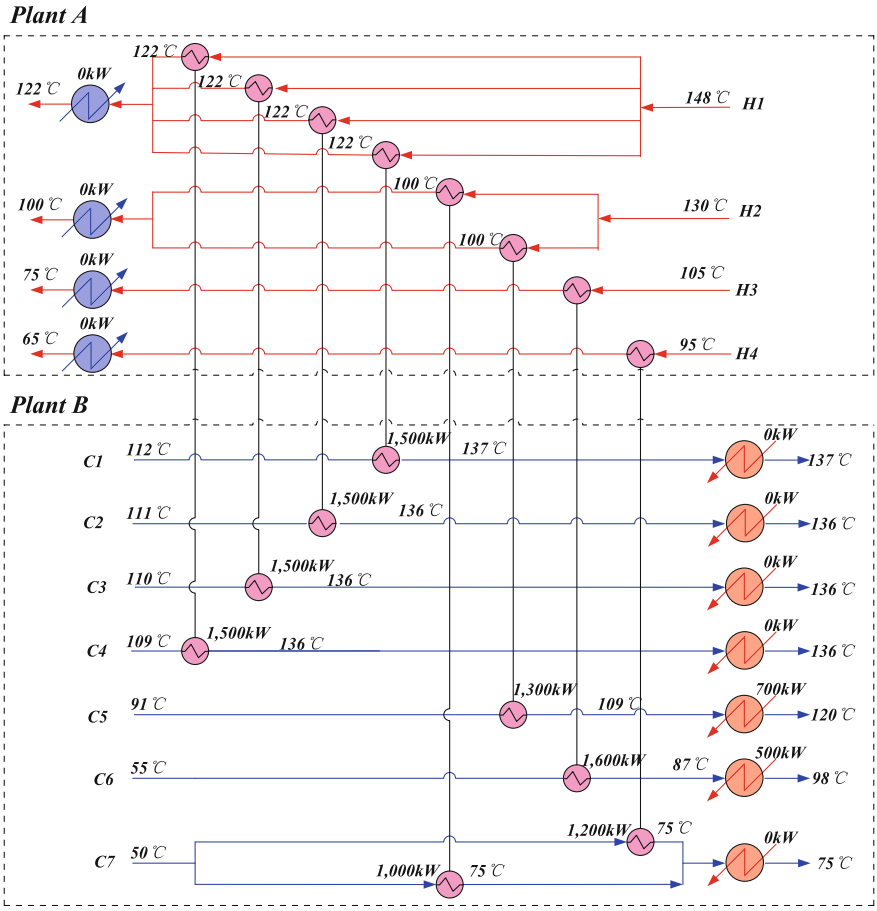

Fig. 21.12 The optimal heat integration grid diagram for direct heat integration

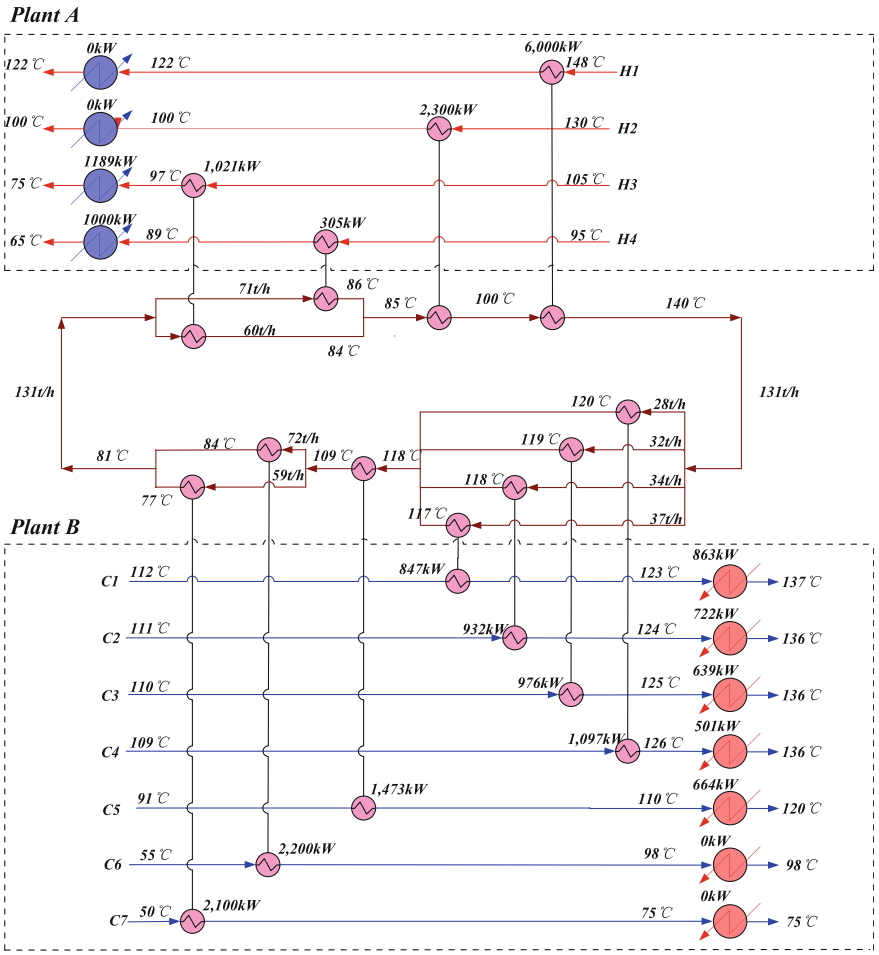

Fig. 21.13 The optimal heat integration grid diagram for indirect heat integration

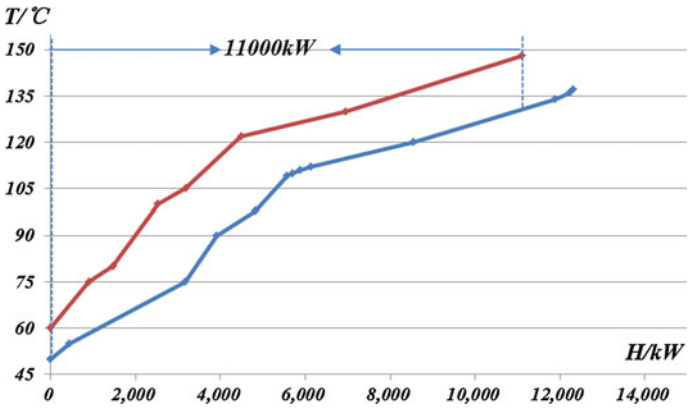

Fig. 21.14 The Composite Curves for direct heat integration

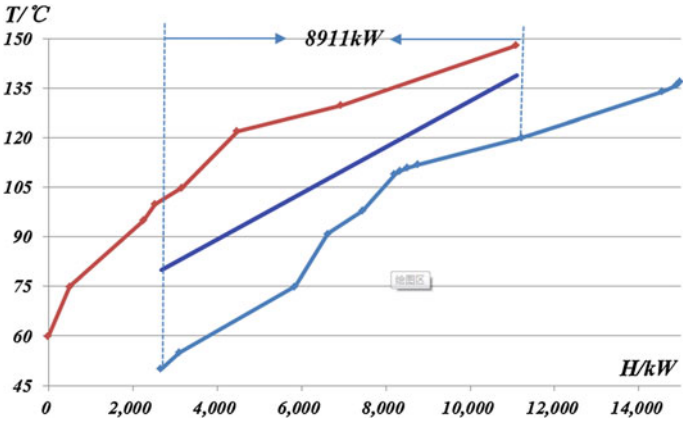

Fig. 21.15 The Composite Curves for indirect heat integration

$$Pipe\left(\$\cdot m^{-1}\right)=\frac{I\cdot(1+I)^{y}}{(1+I)^{y}-1}\left(0.82Wt_{pipe}+185D_{out}^{0.48}+6.8+265D_{out}\right) \tag{49}$$

By solving the presented mathematical models, the optimal heat integration designs for direct and indirect integration can be obtained. The optimal structures for direct and indirect heat integration are shown in Figs. 21.12 and 21.13. Also, the Composite Curves for direct and indirect integration are shown in Figs. 21.14 and 21.15. From the structure, it can be seen that 4 (H1, H2, C6, C7) loops and 9 heat exchangers are used to transfer heat in the direct heat integration. For the indirect integration, only 1 intermediate fluid loop is used and 11 heat exchangers are used to transfer heat. From the view of investment, the direct integration requires more loops and less heat exchangers, indirect integration requires fewer loops and more heat exchangers. It is noted that all the process streams have a utility heat exchanger even most of the exchangers have no heat duty. This is because they are needed when one of the plants is shut down.

From the Composite Curves, it can be found that for the direct integration, the two curves are close to each other, which indicates a higher energy recovery. For the indirect integration with an intermediate fluid loop, in the Composite Curves, the intermediate fluid loop is illustrated by a straight line. Because a straight line

Table 21.2 The economic performance for the two integration patterns

Project items	Pipeline loops	Additional heat exchangers	Heat exchangers area	Pipeline cost ($\$\cdot y^{-1}$)	Heat exchangers cost ($\$\cdot y^{-1}$)	Heat cost ($\$\cdot y^{-1}$)	TAC ($\$\cdot y^{-1}$)
Direct integration	4	8	2,069	632,220	334,465	228,000	1,194,686
Indirect integration	1	11	2,132	166,369	352,920	687,590	1,206,879

cannot match the profile of two curves well, the gap between the two curves is larger, indicating a lower energy recovery. The optimization results are listed in Table 21.2.

In the results, each pipeline loop is 1300 m long, which is twice of the distance between the plants. From the results, for the direct integration, because of the number of pipeline loops is high, the pipeline cost is much higher than the indirect integration. For the indirect integration, it can be seen that the energy recovery is much lower than the direct integration.

21.6 Conclusion

Direct and indirect heat integration are two basic methods for heat integration across plants. When distance factors are involved in heat integration across plants, it is found that direct heat integration is not economic since every stream involved in heat integration requires an independent pipeline. In indirect heat integration, the parameters of the intermediate fluid should be optimized because it correlated energy saving and pipeline investment. An optimization methodology is presented in this chapter to optimize the design of direct and indirect heat integration across plants. From the results of case study, it can be found that if the distance is long, and the number of process streams involved in heat integration is large, indirect heat integration is much more beneficial. Only when the distance is short and the number of process streams involved in heat integration is low, direct heat integration is more economic.

For indirect heat integration, when distance factors are involved, it is also important to consider the interconnection pattern between plants. Three connection patterns—parallel, split, and series—are presented in this chapter, which feature different energy performances and pipeline length requirements.

Among the three connection patterns, the parallel connection pattern can always recover more heat than the other two, but it requires the longest pipeline which will incur high investment as well as operation costs. When the heat quality requirements of the two heat sink plants are similar, the split connection pattern becomes attractive when compared to the parallel connection pattern, because its pipeline length requirement is noticeably smaller while its energy requirement is not significantly larger. When the heat quality requirements of the two heat sink plants are very different, the series connection pattern performs well because it offers the shortest pipeline length requirement. However, its energy requirement is always higher than that of the parallel connection pattern.

Acknowledgments Financial support from the National Basic Research Program of China (973 Program: 2012CB720500), the National Natural Science Foundation of China under Grant No. 21476256 are gratefully acknowledged.

References

Asante, N. D. K., & Zhu, X. X. (1996). An automated approach for heat exchanger network retrofit featuring minimal topology modifications. *Computers & Chemical Engineering, 20*, S7–S12.

Bagajewicz, M., & Rodera, H. (2000). Energy savings in the total site heat integration across many plants. *Computers & Chemical Engineering, 24*, 1237–1242.

Chang, C., Wang, Y., & Feng, X. (2015). Indirect heat integration across plants using hot water circles. *Chinese Journal of Chemical Engineering, 23*, 992–997.

Chew, K. H., Klemeš, J. J., WAN ALWI, S. R., & ABDUL MANAN, Z. (2013). Industrial implementation issues of total site heat integration. *Applied Thermal Engineering, 61*, 17–25.

Dhole, V. R., & Linnhoff, B. (1993). Total site targets for fuel, co-generation, emissions, and cooling. *Computers & Chemical Engineering, 17*((supp.1)), S101–S109.

Hackl, R., Andersson, E., & Harvey, S. (2011). Targeting for energy efficiency and improved energy collaboration between different companies using total site analysis (TSA). *Energy, 36*, 4609–4615.

Hammond, G. P., & Norman, J. B. (2014). Heat recovery opportunities in UK industry. *Applied Energy, 116*, 387–397.

HIP Lito-Valencia, B. J., RUBIO-CASTRO, E., PONCE-ORTEGA, J. M., SERNA-GONZ LEZ, M., N POLES-RIVERA, F., & EL-HALWAGI, M. M. (2014). Optimal design of inter-plant waste energy integration. *Applied Thermal Engineering, 62*, 633–652.

Hu, C. W., & Ahmad, S. (1994). Total site heat integration using the utility system. *Computers & Chemical Engineering, 18*, 729–742.

Jiang, D., & Chang, C.-T. (2013). A new approach to generate flexible multiperiod heat exchanger network designs with timesharing mechanisms. *Industrial and Engineering Chemistry Research, 52*, 3794–3804.

Kapil, A., Bulatov, I., Smith, R., & Kim, J.-K. (2012). Site-wide low-grade heat recovery with a new cogeneration targeting method. *Chemical Engineering Research and Design, 90*, 677–689.

Klemeš, J., Dhole, V. R., Raissi, K., Perry, S. J., & Puigjaner, L. (1997). Targeting and design methodology for reduction of fuel, power and CO2 on total sites. *Applied Thermal Engineering, 17*, 993–1003.

Kovač Kralj, A., GLAVIČ, P., & KRAVANJA, Z. (2005). Heat integration between processes: Integrated structure and MINLP model. *Computers & Chemical Engineering, 29*, 1699–1711.

Linnhoff, B., & Hindmarsh, E. (1983). The pinch design method for heat exchanger networks. *Chemical Engineering Science, 38*, 745–763.

Matsuda, K., Hirochi, Y., Tatsumi, H., & Shire, T. (2009). Applying heat integration total site based pinch technology to a large industrial area in Japan to further improve performance of highly efficient process plants. *Energy, 34*, 1687–1692.

Rodera, H., & Bagajewicz, M. J. (1999). Targeting procedures for energy savings by heat integration across plants. *AIChE Journal, 45*, 1721–1742.

Rodera, H., & Bagajewicz, M. J. (2001). Multipurpose heat-exchanger networks for heat integration across plants. *Industrial and Engineering Chemistry Research, 40*, 5585–5603.

Stijepovic, M. Z., & Linke, P. (2011). Optimal waste heat recovery and reuse in industrial zones. *Energy, 36*, 4019–4031.

Suaysompol, K., & Wood, R. M. (1993). Estimation of the installed cost of heat exchanger networks. *International Journal of Production Economics, 29*, 303–312.

Wang, Y., Chang, C., & Feng, X. (2015). A systematic framework for multi-plants Heat Integration combining direct and indirect heat integration methods. *Energy, 90*(Part 1), 56–67.

Wang, Y., Feng, X., & Chu, K. H. (2014). Trade-off between energy and distance related costs for different connection patterns in heat integration across plants. *Applied Thermal Engineering, 70*, 857–866.

Yee, T. F., & Grossmann, I. E. (1990). Simultaneous optimization models for heat integration–II. Heat exchanger network synthesis. *Computers & Chemical Engineering, 14*, 1165–1184.

Part IV
Low-Carbon Energy Systems

Chapter 22
Synthesis of Sustainable Biofuel Production Processes: A Generic Methodology for Superstructure Optimization and Data Management

Maria-Ona Bertran, Alberto Orsi, Flavio Manenti, John M. Woodley and Rafiqul Gani

22.1 Introduction

Energy demand and emissions are rising steadily, and are forecast to double by 2050 (IEA 2012). In fact, energy-related carbon dioxide CO_2 emissions have reached historic highs (IEA 2012). Moreover, the reliance on a narrow set of technologies and fossil fuels is a threat to energy security, which raises concerns (IEA 2012). It is therefore clear that current trends in energy supply are unsustainable—economically, environmentally and socially (IEA 2011).

With a global energy-related CO2 target in 2050 of 50 % below current levels, the International Energy Agency IEA, is developing a roadmap, which forecasts 32 exajoules of biofuels used globally by 2050 that is 27 % of world transport fuel (IEA 2011). Biofuels contribute in shifting to low-carbon, non-petroleum fuels, with minimal changes to the current distribution infrastructure and vehicles (IEA 2011). Moreover, the CO_2 streams from biofuel production (fermentation, gasification) are relatively pure, making its capture less laborious than CCS of flue gasses from fossil-fuel power plants.

M.-O. Bertran · J.M. Woodley · R. Gani (✉)
Department of Chemical & Biochemical Engineering,
Technical University of Denmark, Kgs. Lyngby, Denmark
e-mail: rag@kt.dtu.dk

M.-O. Bertran
e-mail: marber@kt.dtu.dk

J.M. Woodley
e-mail: jw@kt.dtu.dk

A. Orsi · F. Manenti
Politecnico di Milano, Dipartimento di Chimica, Materiali E Ingegneria Chimica Giulio Natta, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
e-mail: orsi.alberto8@gmail.com

F. Manenti
e-mail: flavio.manenti@polimi.it

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_22

This shift from crude oil to biomass feedstocks for the production of chemicals and fuels, the latter being the focus of this chapter, requires suitable methods and tools for synthesizing, evaluating and selecting process alternatives, and for designing promising processes. Systematic process synthesis and design methods have been developed for traditional chemical processes. However, these are not directly applicable to biorefinery synthesis and design, due to their unique characteristics. For example, the economic feasibility of biofuels production is highly dependent on the location due to variations in availability, characteristics and costs of biomass in different geographic locations. Moreover, unlike crude oil, each different biomass source provides a different feedstock in terms of chemicals, their composition and properties, and even the same feedstock varies in its characteristics based on the region and climate. Therefore, unlike the optimal petrochemical refinery, the optimal biorefinery network problem needs to have location-dependent solutions.

22.2 State-of-the-art

Numerous methods have been proposed to solve the process synthesis problem, which belong to three main classes: heuristics- or knowledge-based methods, mathematical programming methods, and hybrid methods.

The basis for solving process synthesis problems via mathematical programming using a systematic modeling framework has been described in various studies (Grossmann 1985, 1990; Kravanja and Grossmann 1997; Yeomans and Grossmann 1999) as well as the solution methods and algorithms for the optimization problem (Floudas 1995; Grossmann 1989). This approach has been applied to various problems, Drobez et al. (2009) solved the synthesis problem for biogas production from animal waste formulated as a mixed-integer nonlinear programming (MINLP) problem. A superstructure for the optimal design of a multi-product biorefinery was proposed by Zondervan et al. (2011), which contained options for the production of five products (ethanol, butanol, succinic acid, gasoline and gasohol) and was formulated as a mixed-integer programming (MIP) problem using a generic process model based on processing intervals. The idea was taken further by developing an integrated business and engineering framework that includes a step-by-step procedure and integrates different methods and tools for the problem definition, formulation and solution (Quaglia et al. 2012). This framework uses the concept of processing interval with an improved generic model. The framework was applied to industrial optimal synthesis problems including utilization of soybean oil and treatment and reuse of water effluent produced by an oil refinery (Quaglia et al. 2014). Rizwan et al. (2013) proposed a superstructure and a processing network model, with the processing interval concept, for the production of biodiesel from microalgal biomass. A large biomass processing network superstructure with biofuel products has been proposed by Garcia and You (2015), the bioconversion network optimization problem is formulated as a multiobjective problem (maximization of profit and minimization of GWP) taking into account supply chain considerations but using simple

input-output models for the processing technologies. González-Delgado and Kafarov (2015) show a combined method for the synthesis of a microalgae-based biorefinery, it starts with hierarchical approach to identify promising pathways and continues with superstructure optimization.

There have also been efforts towards building generic models and model libraries that are applicable to biorefinery processes, and organizing and collecting biorefinery-related data for process synthesis. Martín and Grossmann (2012) developed a library of equation-based models for design of biofuel processes using superstructure optimization. The models are based on a combination of design of experiments, empirical correlations based on physical insights, rules of thumb, among others.

22.3 Problem Statement: Mathematical Description

The synthesis problem can be defined mathematically as follows (Floudas 1995).

$$\begin{cases} \min\limits_{x,y} f(x,y) \\ \text{s.t. } h(x,y) \\ \quad\quad g(x,y) \\ \quad\quad x \in X \leq R^n \\ \quad\quad y \in Y = \{0,1\}^l \end{cases} \tag{22.1}$$

where x is a vector of continuous variables (such as temperature, pressure, flowrates) and y is a vector of binary variables.

The problem of optimal synthesis of biorefinery networks is defined as: *given* a geographic location or a set of locations, *find* a set of available biomass-based feedstocks, a set of products that can be produced from them, a set of processing routes (represented by a sequence of relevant processing steps) connecting feedstocks and products, and a set of technological alternatives for each processing step and *determine* the optimal processing network topology in terms of feedstocks, processing routes-technologies and products, as well as the associated optimal process variables, based on a pre-defined objective function. Note that in the problem definition "*find*" has been used for database or literature search and "*determine*" for mathematical programming.

Based on the above definition, different types of network optimization problems can be formulated, depending on the number of alternatives considered of each type (that is feedstock alternatives, route-technology alternatives, and product alternatives). Examples of these problem types are shown in Fig. 22.1: for example, problem type (a) is a route selection problem, feedstock and product materials are specified, and there is a number of alternative routes; and problem type (b) is a product selection problem, feedstock is fixed, but product is not, and there is only one route per product.

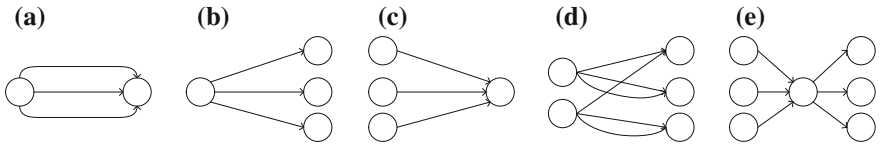

Fig. 22.1 Different problem types in network optimization problems: **a** route selection, **b** product selection, **c** raw material selection, **d** simultaneous raw material, route and product selection, and **e** raw material and product selection via intermediate

22.4 Methodology and Tools

A model-based methodology for synthesis of processing networks has been proposed by Bertran et al. (2016). The aforementioned methodology is described in this section, along with related concepts and tools. The section is structured as follows: first the three-stage approach is briefly introduced and the scope of the methodology is defined with respect to it, next definitions and rules for the superstructure representation are given in Sect. 22.4.1, then the optimization model is described in Sect. 22.4.2, and Sect. 22.4.3 contains the details about the management of the large amount of data needed to solve the problem with the given model. All of the aforementioned are integrated in a step-by-step methodology described in Sect. 22.4.4, which has been made available as a software tool under the name of Super-O, as can be read in Sect. 22.4.5.

A three-stage approach for sustainable process synthesis, design and innovation (Fig. 22.2) has been described by Babi et al. (2015). The approach consists in decomposing the design problem into three stages. In the first stage, the synthesis stage, after defining the problem, the search space needs to be determined and represented and the top or top ranked processing network(s) are selected. The selected networks are further designed and analysed in the second stage, the design stage. Finally, the third stage is concerned with using the results from the analyses performed in stage two in order to target process bottlenecks and generate innovative designs. The synthesis method presented here represents the first stage (the synthesis stage) of the aforementioned three-stage approach.

22.4.1 Superstructure Representation

A superstructure is a representation of the alternatives comprised within the search space that is being considered in a decision-making problem. In the case considered here, namely synthesis of biorefinery processing networks, the superstructure represents all the alternatives with respect to the topology of the network: raw material options are inputs, products are outputs, processing elements are nodes and connections between elements are arcs (Quaglia 2013).

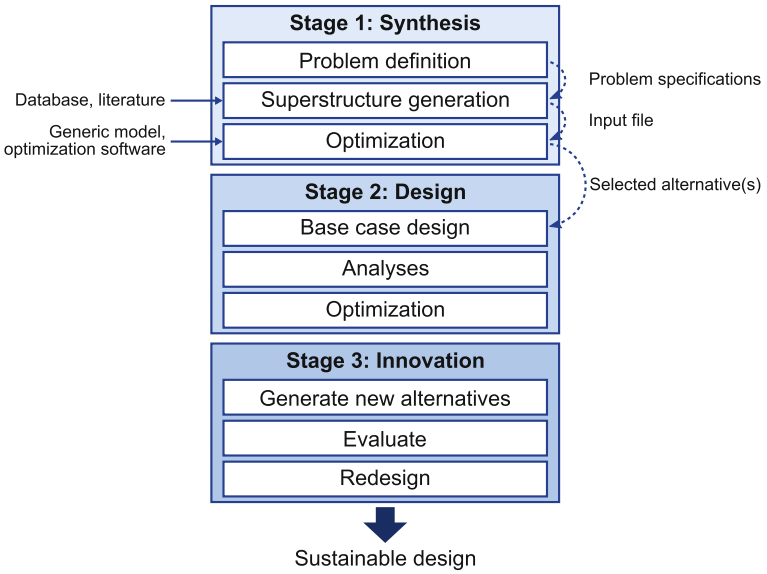

Fig. 22.2 The 3-stage approach with related tools and inputs-outputs for the first stage (tools on the *left*, inputs-outputs on the *right*)

A stage-wise superstructure representation is used here, which was presented by Quaglia (2013), in which a series of columns are used to organize the alternatives. In the first column, the raw material alternatives are collected, the last column contains the product alternatives, and the remaining columns correspond to the processing steps required for the transformation of the raw materials into the products. Each processing step contains one or more processing intervals, the technological alternatives to execute the processing step. A generic superstructure is displayed in Fig. 22.3, which serves to illustrate the concept. In terms of graphical representation, vertical lines are used to define boundaries between processing steps and boxes represent processing intervals.

The superstructure in Fig. 22.3 contains the following: 3 processing steps, 2 raw material alternatives, 9 processing intervals, including 1 bypass (interval 3-3), and 3 products. In terms of connections, the superstructure has a total of 18 connections, 16 of which are primary connections whereas 2 are secondary, 1 out of the 2 secondary connections is a recycle stream. The secondary connections are identified by their starting point not being placed on the right-hand side of the boxes representing processing intervals, but being placed elsewhere, normally on the top or bottom of the box. When representing in a superstructure the alternative of bypassing a processing step, two options are allowed: either using a bypass interval, that is placing "empty" interval, or directly connecting intervals from the adjacent steps, in Fig. 22.3 the former is used.

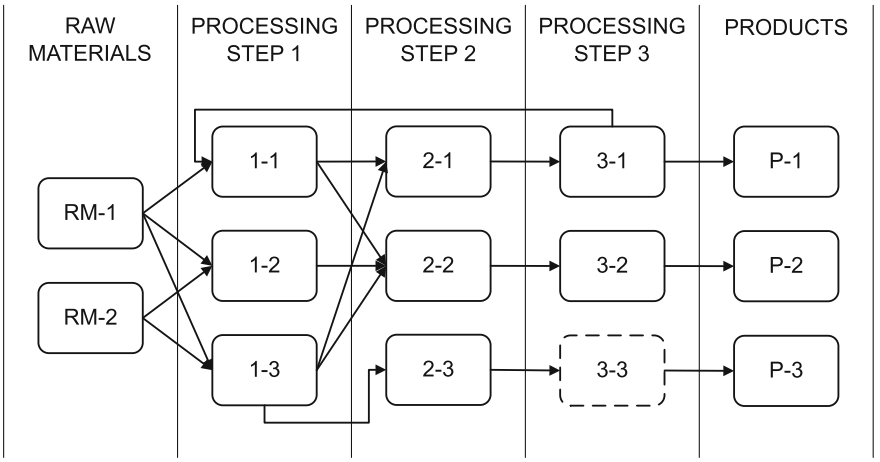

Fig. 22.3 Generic superstructure for a synthesis problem of processing networks illustrating the concepts of processing step (*column*), processing interval (*box*), by-pass (*box with dotted outline*), primary connection (*arrow* starting from the right-hand side of the interval), secondary connection (*arrow* starting from the *top* or *bottom* of the interval)

The above presented superstructure representation has the ability of representing synthesis problems at different levels, that is, different types of superstructure can be posed with varying level of aggregation of the alternatives as well as in their organization. Consider the example in Fig. 22.4, which shows two alternative separation sequences for the downstream processing of biodiesel (Zhang et al. 2003).

The representation of the two alternative sequences in Fig. 22.4 in a superstructure is not trivial. The level of detail and organization in the superstructure depends on the problem that is being solved. First of all, the processing steps need to be determined and placed in order. A first approach to determining the sequence of processing steps could be based on the type of unit operation, hence obtaining four processing steps, namely washing, neutralization, gravity separation and distillation. The disadvantage of this approach is the difficulty in finding an order of the processing steps that fits all the alternative sequences. As illustrated in Fig. 22.5, in the case considered here, the two alternative sequences differ in the order of the units, hence if the type of unit is taken as processing step, the superstructure representation will either have one or more repetitions of processing steps (Fig. 22.5a) or a connection going from left to right without being a recycle (Fig. 22.5b), which is not desirable, and even if some units are condensed in the same interval (Fig. 22.5c) the problem might still appear.

Another possible criterion to organize the superstructure is using the order of the operations itself, hence the processing steps would be: first separation step, second separation step, third separation step, and so on (see Fig. 22.6). Therefore, the first processing step of the sequence would be *separation 1*, and the alternatives would be performing separation between oil-phase and glycerol-phase first, or separating the catalyst first (via neutralization and gravity separation). Using this approach, various

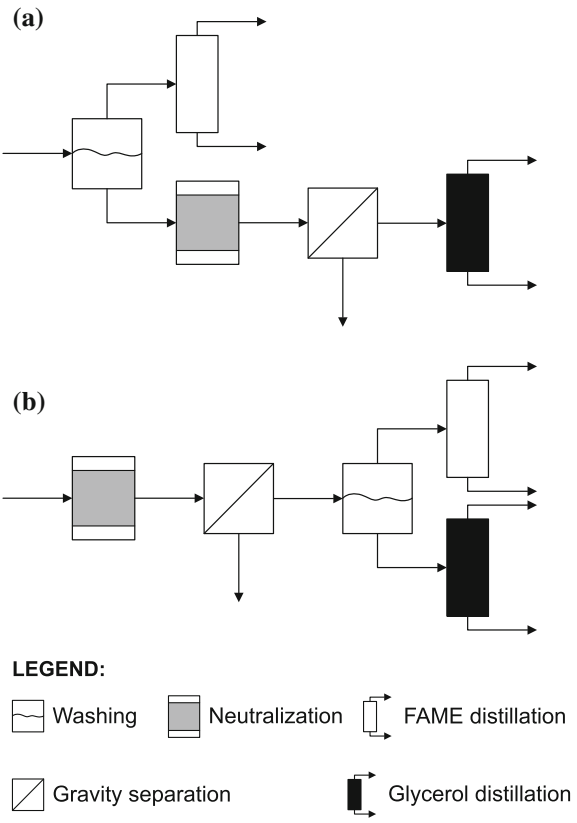

Fig. 22.4 Two alternative downstream processing sequences for the biodiesel production process via transesterification of triglycerides: in **a** water washing is performed first, and neutralization of the catalyst in the aqueous phase is done afterwards, whereas in **b** neutralization is performed first and water washing is done afterwards

representations can be obtained: Fig. 22.6a shows a superstructure based on separation steps where the final separation of the primary and secondary product streams are considered in the same processing step, whereas in Fig. 22.6b they are placed in subsequent steps, in Fig. 22.6c the two units for catalyst removal (neutralization and gravity separation) are considered together in the same processing interval, finally Fig. 22.6d is the most condensed version of the superstructure with only three steps since it the catalyst removal units are merged in one interval and the final separation of primary and secondary products are placed in one step. The choice of superstructure representation should be made based on the problem that is being solved, for example for a problem where a unique neutralization reactor with multiple choice of gravity separation is considered, the versions in Fig. 22.6c and d are not convenient since by merging the two units in one interval, separate alternatives are not allowed.

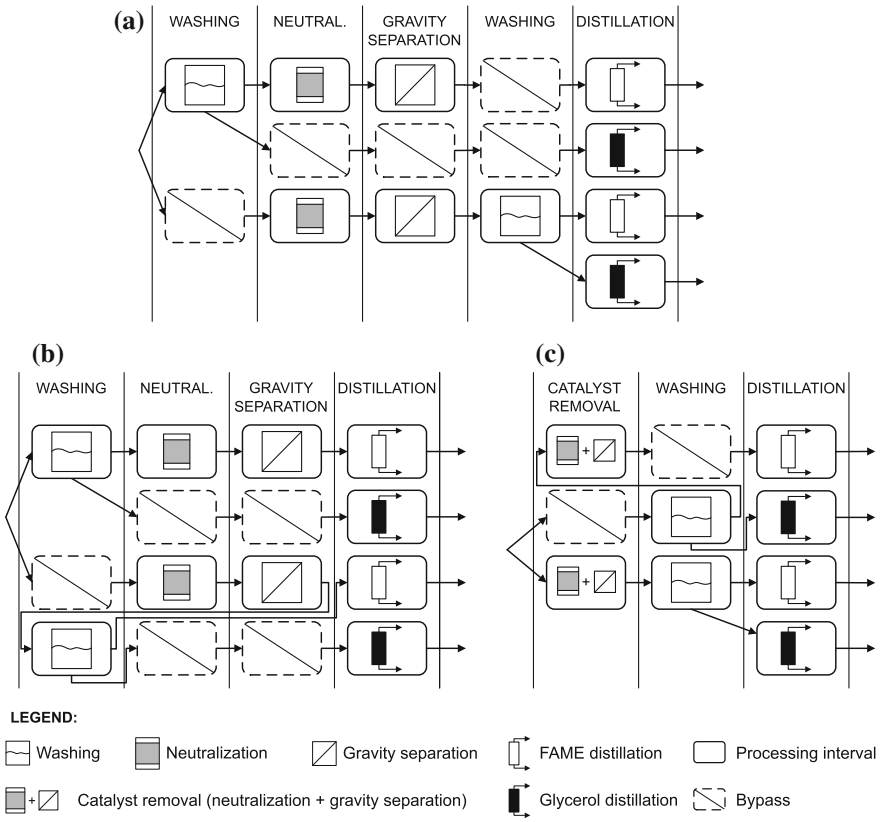

Fig. 22.5 Alternative superstructure representations of two downstream processing sequences with processing steps based on type of equipment: in **a** washing appears twice as step to avoid connections from *right* to *left*, in **b** only one washing interval is used, but a connection from *right* to *left* appears, and in **c** two steps are merged, yet still obtaining a connection from *right* to *left*

22.4.2 Mathematical Model

The synthesis problem is formulated as a mixed-integer (non)linear programming problem, that is an MI(N)LP problem, consisting of an objective function and a series of equality and inequality constraints. The equality constraints include the process model representing the processing technologies and the possible connections between them.

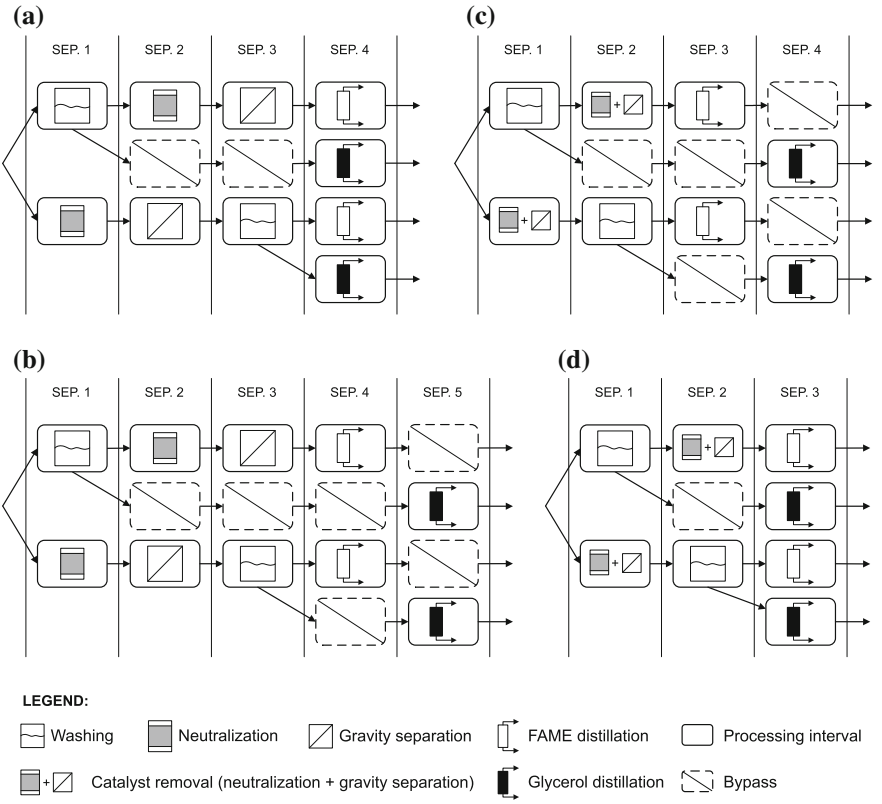

Fig. 22.6 Alternative superstructure representations of two downstream processing sequences with processing steps based on sequence of separation: in **a** four processing steps are considered, whereas in **b** five are needed since the primary and secondary product purification are separated, in **c** this is maintained but two units are merged in represented together in one processing interval, neutralization and gravity separation, and in **d** the same is used as well as representing the final separation for biodiesel and glycerol in the same step. Note that with any of the variations of this approach all streams go from *left* to *right*

The generic modeling approach proposed by Quaglia (2013) has been adopted. The basis of this approach is a generic process model which is able to represent every alternative in the superstructure. The model represents each alternative as a series of processing tasks, namely mixing of chemicals, reaction, waste separation, product separation and utility consumption. A schematic representation of the generic process interval can be seen in Fig. 22.7.

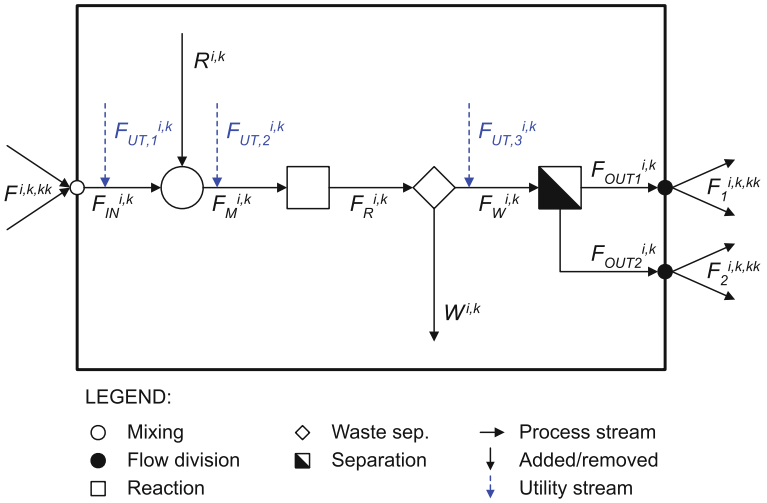

Fig. 22.7 Scheme of the generic process interval with internal variables

The advantage of the use of the generic process interval concept is the ability of the model to represent all different kinds of processing alternatives as well as providing the values of the internal flowrates.

22.4.2.1 Model Equations

The model equations are given in this section, they are divided into: (i) objective function, (ii) logical constraints, (iii) process interval model equations, and (iv) connection equations.

Objective Function

Equation 22.2 is the objective function,

$$z = S - C_{RM} - C_C - C_U - C_W - C_T - CAPEX/\tau \tag{22.2}$$

where the product sales S are calculated from the price and flowrate of products (Eq. 22.3),

$$S = \sum_i \sum_k P_P^k F_W^{i,k} \tag{22.3}$$

the raw materials cost C_{RM} is obtained from the price and flowrate of raw materials (Eq. 22.4),

$$C_{RM} = \sum_i \sum_{kk} P_F^{kk} F_W^{i,kk} \tag{22.4}$$

C_C corresponds to the cost of chemicals and solvents added to the process (Eq. 22.5),

$$C_C = \sum_{kk} \sum_{i} P_C^i R^{i,kk} \tag{22.5}$$

C_U is the cost of utilities (Eq. 22.6),

$$C_U = \sum_{kk} \sum_{ut} P_U^{ut} F_{UT}^{ut,kk} \tag{22.6}$$

C_W are the waste handling costs (Eq. 22.7),

$$C_W = P_W \sum_{i} \sum_{kk} W^{i,kk} \tag{22.7}$$

C_T is the transportation cost (Eq. 22.8),

$$C_T = \sum_{k} \sum_{kk} ctr^{k,kk} \tag{22.8}$$

where $ctr^{k,kk}$ is obtained from (Eq. 22.9),

$$ctr^{k,kk} = DIST^{k,kk} \sum_{i} F^{i,k,kk} \tag{22.9}$$

and the capital cost is calculated from (Eq. 22.10).

$$CAPEX = \sum_{kk} invl^{kk} \tag{22.10}$$

Logical Constraints

Equations 22.11–22.15 are the logical constraints of the model. Equation 22.11 is a constraint to restrict the amount of intervals selected in a step to one, except if they are followed by a blender, in that case the maximum number of intervals that can be selected is the number of intervals connected to a blender plus one. Therefore, any interval where more than one connection is to be allowed simultaneously from the previous step needs to be declared as blender.

$$\sum_{kk} y^{kk} v^{kk,step} \leq 1 + \sum_{k} \sum_{kk} \left(v^{k,step} S^{k,kk} \omega^{kk} \right) \tag{22.11}$$

Equations 22.12–22.15 are activation constraints. Equation 22.12 is an activation constraint that allows $F_W^{i,kk}$ to take any value between zero and an upper bound M only if the interval kk is selected, otherwise it forces $F_W^{i,kk}$ to be zero.

$$F_W^{i,kk} \leq y^{kk} M \tag{22.12}$$

Equation 22.13 is an activation constraint that allows $R^{i,kk}$ to take any value between zero and an upper bound M only if the interval kk is selected, otherwise it forces $R^{i,kk}$ to be zero.

$$R^{i,kk} \leq y^{kk} M \tag{22.13}$$

Equation 22.14 is an activation constraint that allows y^{kk} to be 1 only if the flowrate $F_W^{i,kk}$ is not zero.

$$y^{kk} \leq \sum_i F_W^{i,kk} \frac{M}{1000} \tag{22.14}$$

Equation 22.15 is an activation constraint that lets $F_{IN}^{i,kk}$ take any value between zero and a large number M only if the interval kk is selected, otherwise it forces $R^{i,kk}$ to be zero.

$$\sum_i F_{IN}^{i,kk} \leq y^{kk} M \tag{22.15}$$

Process Interval Model Equations

Equations 22.16–22.25 are the mass balance equations of the processing intervals. Equation 22.16 is the overall component mass balance of each interval.

$$F^{i,k,kk} = F_1^{i,k,kk} + F_2^{i,k,kk} \tag{22.16}$$

Equation 22.17 equates the flowrate of the inlet stream of the mixing task of an interval kk to the sum of the streams going from all the intervals k from the previous step to the present interval.

$$F_{IN}^{i,kk} = \sum_k F^{i,k,kk} \tag{22.17}$$

The flowrate used for transportation calculations in interval kk, F_{tr}^{kk}, is the total flowrate of the outlet stream of the mixing task of the interval, i.e. the summation of component flowrates $F_M^{i,kk}$ over the component set i.

$$F_{tr}^{kk} = \sum_i F_M^{i,kk} \tag{22.18}$$

Mixing

The quantity $R_{i,kk}$ of chemical i added to interval kk is specified through the parameter $\mu^{i,ii,kk}$, which corresponds to the ratio of added compound i in interval kk with respect to a reference compound ii.

$$R^{i,kk} = \sum_{ii} \mu^{i,ii,kk} F_{IN}^{ii,kk} \tag{22.19}$$

The flow of component i in interval kk after the mixing task is calculated by adding the flowrate before mixing $F_{IN}^{i,kk}$ and the flowrate $R_{i,kk}$ of chemical i added to the interval.

$$F_{M}^{i,kk} = F_{IN}^{i,kk} + R^{i,kk} \tag{22.20}$$

Reaction

The flowrate of component i in interval kk after the reaction task $F_{R}^{i,kk}$ is calculated in Eq. 22.21 from the flowrate of the same component i before the reaction task $F_{M}^{i,kk}$ and the amount of component generated or consumed by the reaction.

$$F_{R}^{i,kk} = F_{M}^{i,kk} + \sum_{rr,react} F_{M}^{react,kk} \theta^{react,kk,rr} \gamma^{i,kk,rr} \frac{MW^{i}}{MW^{react}} \tag{22.21}$$

Waste Separation

The modelling of the waste separation task requires two equations. Equation 22.22 is used to obtain the flowrate of component i that continues to the next task in the process interval, $F_{W}^{i,kk}$.

$$F_{W}^{i,kk} = F_{R}^{i,kk} \left(1 - \delta^{i,kk}\right), \quad \forall (kk > NRAW) \tag{22.22}$$

The flowrate of the waste stream $W^{i,kk}$ is calculated in Eq. 22.23.

$$W^{i,kk} = F_{R}^{i,kk} - F_{W}^{i,kk}, \quad \forall (kk > NRAW) \tag{22.23}$$

Product Separation

The product separation task is modelled in Eq. 22.24 with the user-defined parameter $\sigma^{i,kk}$, which corresponds to the ratio of flowrate of component i leaving the task via the primary product stream to the flowrate of component i in the inlet stream.

$$F_{OUT,1}^{i,kk} = F_{W}^{i,kk} \sigma^{i,kk} \tag{22.24}$$

The secondary product stream flowrate is calculated by simple mass balance, as shown in Eq. 22.25.

$$F_{OUT,2}^{i,kk} = F_{W}^{i,kk} - F_{OUT,1}^{i,kk} \tag{22.25}$$

Consumption of Utilities

Equations 22.26–22.29 model the utility consumption in each processing interval. The amount of utility ut consumed in each task of interval kk is defined as a function of the total flowrate of process stream in the interval. Since, along the interval, the flowrate changes due to changes in composition and addition or separation of streams, different points have been selected in which addition of utilities can be specified. Afterwards, the total consumption of utilities is calculated as the sum of the

consumption of utilities in each utility consumption point. Three different utility consumption points (1, 2, and 3) have been defined based on the fact that different values of the total mass flowrate can be obtained in different points of the same processing interval, and three are considered representative. Equation 22.26 obtains the amount of utility *ut* mixed in point 1, which is defined with respect to the inlet flowrate $F_{IN}^{ii,kk}$ in interval *kk*. The parameter $\beta_1^{ut,kk}$ is the ratio between amount of utility added in the mixing task and inlet flowrate to this task.

$$F_{UT,1}^{ut,kk} = \beta_1^{ut,kk} \sum_{ii} F_{IN}^{ii,kk} \tag{22.26}$$

Equation 22.27 calculates the flowrate of utility *ut* added in the second utility consumption point of interval *kk*, that is calculated with respect to flowrate $F_M^{ii,kk}$. Since, by mass balance, $\sum_{ii} F_M^{ii,kk} = \sum_{ii} F_R^{ii,kk}$, only one utility point has been defined among this two streams.

$$F_{UT,2}^{ut,kk} = \beta_2^{ut,kk} \sum_{ii} F_M^{ii,kk} \tag{22.27}$$

The third utility consumption point is the outlet of the waste separation task, or the inlet of the product separation task. The amount of utility *ut* consumed in this point is calculated through Eq. 22.28.

$$F_{UT,3}^{ut,kk} = \beta_3^{ut,kk} \sum_{ii} F_W^{ii,kk} \tag{22.28}$$

Finally, the total amount of utility *ut* consumed in the different tasks of interval *kk*, $F_{UT}^{ut,kk}$ is calculated in Eq. 22.29.

$$F_{UT}^{ut,kk} = F_{UT,1}^{ut,kk} + F_{UT,2}^{ut,kk} + F_{UT,2}^{ut,kk} \tag{22.29}$$

Connection Equations

Equation 22.30 is a constraint that allows the (primary) flowrate between two intervals to take a value only if the mentioned intervals are connected between them (through a primary connection). If they are indeed connected, then the value that the connecting flowrate can take has an upper bound that corresponds to the outlet flowrate of the first interval. This is the upper bound because what goes from one interval to another through a primary connecting flowrate cannot be larger than what leaves from the first interval via primary outlet, it can though be smaller, since the outlet can be divided and sent to different intervals.

$$F_1^{i,k,kk} \leq F_{OUT,1}^{i,k} SP^{k,kk} \tag{22.30}$$

Equation 22.31 constraint is similar to Eq. 22.30 but for secondary connections between intervals, therefore it uses the secondary outlet flowrate as upper bound for the secondary connecting flowrate.

$$F_2^{i,k,kk} \leq F_{OUT,2}^{i,k} \left(S^{k,kk} - SP^{k,kk}\right) \tag{22.31}$$

Equation 22.32 is the mass balance around the stream divider of the primary stream at the outlet of each interval.

$$F_{OUT,1}^{i,k} = \sum_{kk} F_1^{i,k,kk} \qquad \forall\ (k < (NRAW + NPROC + 1)) \tag{22.32}$$

Equation 22.33 is the mass balance around the stream divider of the primary stream at the outlet of each interval

$$F_{OUT,2}^{i,k} = \sum_{kk} F_2^{i,k,kk} \qquad \forall\ (k < (NRAW + NPROC + 1)) \tag{22.33}$$

If intervals k and kk are connected and kk is a blender, Eq. 22.34 allows the selection of interval kk only if interval k is selected, that is the a blender must have at least one inlet.

$$y^{kk}\omega^{kk} \leq y^{k} \qquad \forall\ \left(S^{k,kk} = 1\right) \tag{22.34}$$

Eq. 22.35 assigns the flowrate and composition of raw material to the selected interval of the raw material step.

$$F_W^{i,kk} = \phi^{i,kk} y^{kk} \qquad \forall\, (kk < (NRAW + 1)) \tag{22.35}$$

22.4.3 Data Management

The relevant information for this problem includes the alternatives for each processing step and connections between alternatives in different processing steps, as well as detailed data for each alternative, which corresponds to the parameters of the model described in Sect. 22.4.2. This represents a large amount of data, which calls for the development of a data architecture for its efficient storage, search and retrieval. The data structure needs to be flexible enough so as to be applicable to a diverse set of problems, yet robust enough so as to provide a strong common format for all data. This means that data obtained from various sources and disciplines, including open literature, academic and industrial collaborators, and reaction path synthesis, should be easily combined, allowing the solution of assorted problems together. To this end, an ontology-based data architecture has been developed, upon which specific databases are built, in this case the ProBioRefine database. The developed ontology, shown in Fig. 22.8, consists of two main sections, namely a materials section and a units section. The materials section contains the list of materials, and each of them is assigned an attribute (feedstock, intermediate, or product), which creates a link between the material and one of the material sub-sections. The material

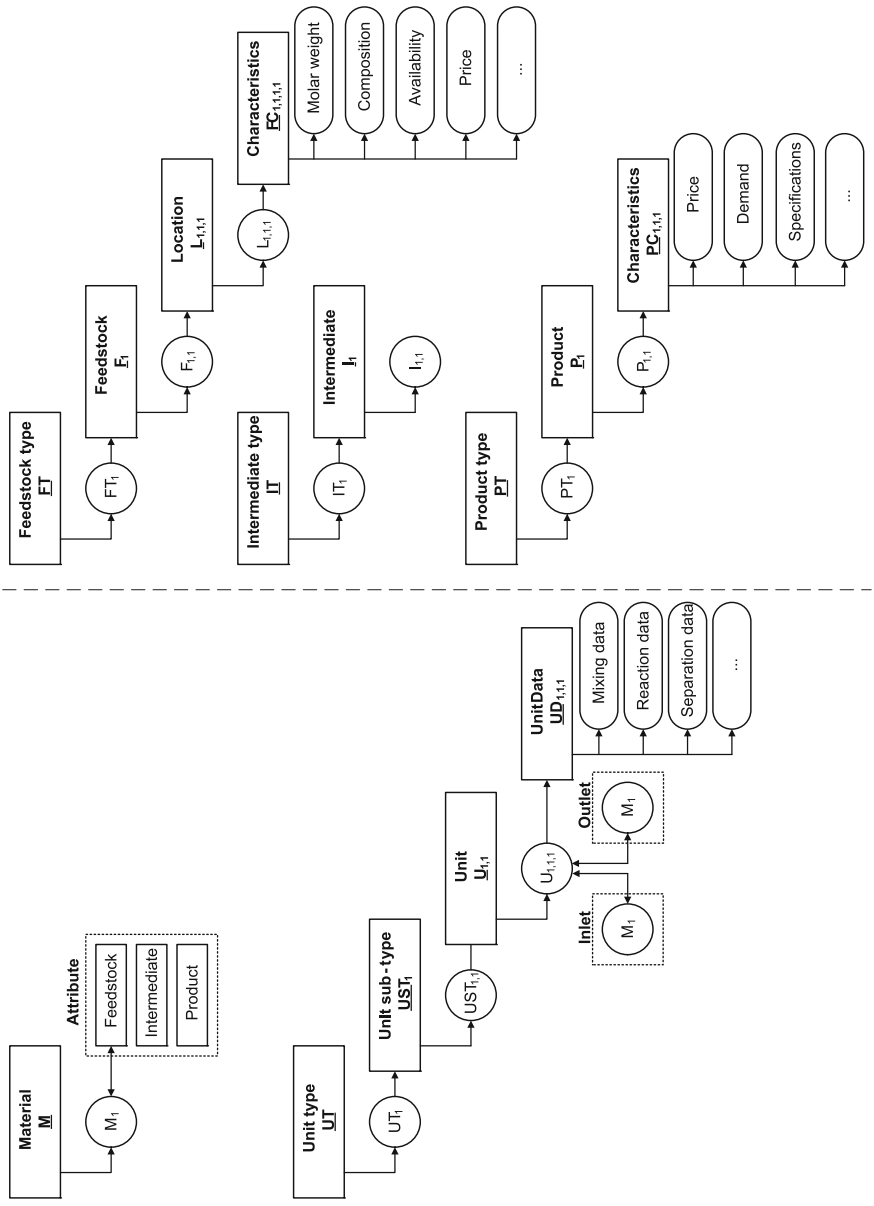

Fig. 22.8 Ontology-based data architecture (*left* main sections, *right* material sub-sections)

sub-sections are: a feedstocks sub-section, an intermediates sub-section, and a products sub-section. This means that the materials with attribute feedstock, appear in the feedstock list, where they are classified (into types) along with all the necessary data. The units section classifies the units in types and sub-types, as well as containing all information about each unit. The link between the materials and the units section is done through assigning one material of the materials list as inlet and another one as outlet of each unit. Therefore, connections between units are inferred by comparison of inlet and outlet material names, hence obtaining a superstructure of alternatives.

Transformations might be needed after retrieving data from the database for its use in superstructure optimization. Data in the database is stored in terms of units, while a superstructure is generated in terms of processing intervals. A processing interval can include a part of a unit, a whole unit, or various units (even a whole plant). Therefore, the definition of processing intervals and what they include depends on the problem scale. Each processing interval is modelled using a generic model, thus the same data is required for each of them. Simple algorithms and calculations that allow for the conversion of units data (from the database) to the processing intervals data (for the superstructure) have been developed.

22.4.4 Synthesis Method

A methodology with a step-by-step workflow has been developed for the synthesis of biorefinery networks integrating the superstructure representation in Sect. 22.4.1, modelled as explained in Sect. 22.4.2, with data organized with the ontology described in Sect. 22.4.3.

The synthesis method consists of three steps: (1) problem formulation; (2) superstructure generation; and (3) solution of the optimization problem.

22.4.4.1 Step 1: Problem Formulation

The problem formulation step consists of defining the problem characteristics, such as the objective function (where different scenarios need to be accommodated), the feedstocks that are to be converted, the products that are to be made, the geographical location and supply chain issues; this leads to a wide range of problems being formulated. Some of these characteristics need to be specified, while the specification of others is optional. For example, if the geographical location is specified, then location-dependent data (available feedstocks, known process characteristics, products, etc.) are retrieved from the database. If a specific feedstock is selected for a problem then the data corresponding to different locations where it is available, is retrieved. Alternatively, all feedstocks stored in the database are considered if one is interested in the optimal feedstock for a specific set of products. This way, different types of network optimization problems can be formulated, as mentioned in Sect. 22.3.

22.4.4.2 Step 2: Superstructure Generation

To generate the superstructure, one or various databases are used to retrieve all the necessary information, that is the alternatives, connections between them, and data for each alternative. The alternatives in the superstructure are represented as processing intervals (Quaglia et al. 2012). Connections between processing intervals are inferred by comparison of inlet and outlet materials of each unit in the database (see Sect. 22.4.3). If new information that is not in any of the databases is to be used, it is first input to one of the databases, so that the data has the same structure as the rest and it can be reused in future problems.

22.4.4.3 Step 3: Solution of the Optimization Problem

First, the mathematical representation of the superstructure is developed. Next, the optimization (MINLP or MILP) problem is solved. This problem includes the user-defined objective function, subject to a series of constraints, namely mass and energy balance constraints, flow model constraints, logic constraints, and other constraints (Quaglia et al. 2012). The developed model is generic and data-independent. This means that it does not have to be further edited by the user, the necessary data needs to be provided as an input file and the problem is solved using external software. Through the developed interface, Super-O, the input file is automatically generated by retrieving the necessary data from the databases.

22.4.5 Software Tool

In order to automate the application of the synthesis methodology described in Sect. 22.4.4, a software tool named Super-O has been developed in C# environment (Bertran et al. 2016). Super-O is a user interface for the synthesis of processing networks, it allows for the reduction of the time needed for the formulation and solution of network optimization problems. A schematic representation of the methodology as implemented in Super-O including the data flow and tools integration is shown in Fig. 22.9.

Super-O integrates especially structured databases, based on the ontology described in Sect. 22.4.3, enabling the systematic storage and retrieval of the necessary information for different problem types. In Super-O, the user specifies the problem and all necessary data are retrieved from databases. Based on the retrieved information, Super-O generates a GAMS-readable binary file containing all problem data for the solution of the optimization problem through the GAMS model file containing the generic model equations. Then, Super-O directly calls an optimization software, in this case GAMS, to solve the optimization problem and generate an output file containing the results. The results file contains the solution of the formulated

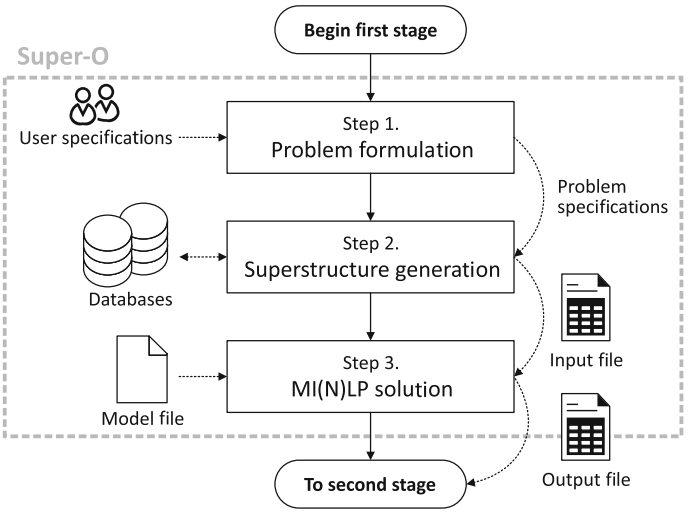

Fig. 22.9 Schematic representation of the work flow, data flow and tools integration implemented in the software interface Super-O

synthesis problem (such as the selected routes, or materials) as well as the values for all model variables, so they can in addition be used in subsequent stages of the three-stage approach (detailed design and analysis, and innovative and more sustainable design). As a result of the systematic structure and the templates generated, standard processing network problems can be formulated and solved through the user interface of Super-O, without requiring any additional programming. This reduces the time needed for the formulation of this class of problems and, at the same time, it broadens the range of potential users of this optimization-based synthesis methodology, by making it accessible to professionals who are not experts in formulating and solving process synthesis problems using superstructure optimization.

22.5 Case Study: Production of Biodiesel

Biodiesel is an alternative fuel which consists of monoalkyl esters of long chain fatty acids derived from renewable lipid feedstock (Zhang et al. 2003). The most common way to produce it is by transesterification of lipids, which can be catalyzed (by alkali, acid or enzymatic catalyst) or non-catalyzed. The aforementioned chemical reaction involves a molecule of triacylglycerol (triglyceride from the oil source) reacting with three molecules of an alcohol (for example methanol) to yield three a fatty acid alkyl ester (FAME) and one molecule of glycerol (Zhang et al. 2003).

The production of biodiesel and co-products from different feedstocks in various locations is considered in this section to illustrate the use of the methodology. The

following assumption is used throughout the case study: only triglycerides and free fatty acids are considered, hence no mono- and di-glycerides are taken into account, due to the lack of reaction data for the step-wise reaction.

The complete superstructure of alternatives for biodiesel production currently available in the ProBioRefine database is shown in Fig. 22.10, which illustrates the complete search space for the problems that are solved here. Each interval in the superstructure is described in Table 22.1.

Various problems can be formulated and solved based on the superstructure of Fig. 22.10, we show two example problems in Sects. 22.5.1 and 22.5.2.

22.5.1 Problem 1

The first example problem is described in this section.

22.5.1.1 Step 1: Problem Definition

The objective of this problem is to select the raw material and determine the optimal topology for biodiesel production in Malaysia by maximizing the Gross Operating Income (GOI) calculated from sales of product, purchase of raw materials and consumption of chemicals and solvents.

22.5.1.2 Step 2: Superstructure Generation

A database search in the ProBioRefine database reveals palm oil and waste cooking palm oil as available raw materials in the selected geographical location.

The complete superstructure of alternatives for biodiesel production from data currently available in the ProBioRefine database is shown in Fig. 22.10. Based on the definition of the problem, the superstructure is reduced by disregarding microalgae as raw material and therefore all intervals which are only connected to it.

22.5.1.3 Step 3: Solution of the Optimization Problem

The data related to the superstructure containing alternatives in terms of raw materials and technologies is available in the database in such a form that can directly be used by the optimization software along with the generic model file to solve the optimization problem. The problem statistics are given in Table 22.2. The optimal topology obtained for this problem is given in Table 22.3.

Fig. 22.11 provides a representation of the complete superstructure for biodiesel production where the disregarded alternatives due to unavailability of the raw material are shaded and the optimal solution in terms of process topology is highlighted in

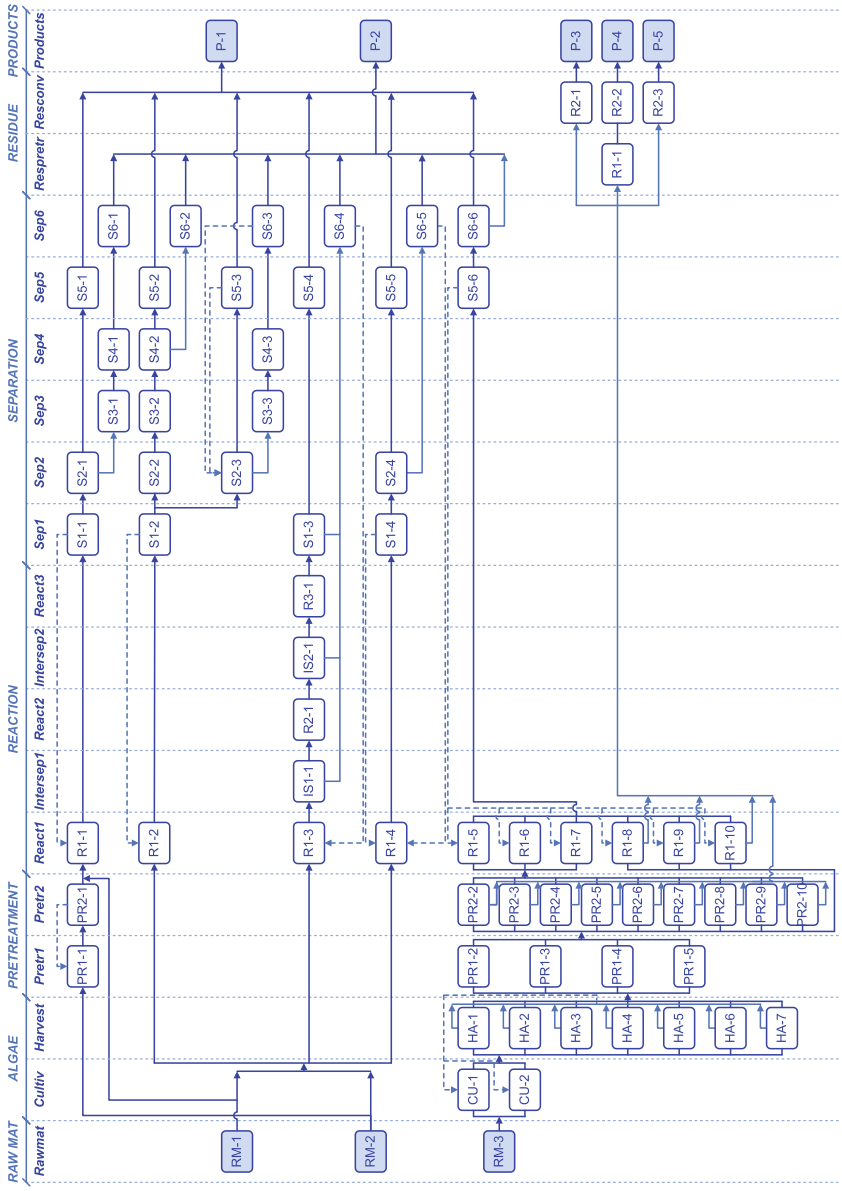

Fig. 22.10 Superstructure for biodiesel and co-products

Table 22.1 Interval list

Raw materials		
RM-1	Virgin palm oil	(Zhang et al. 2003)
RM-2	Waste cooking palm oil	(Zhang et al. 2003)
RM-3	Microalgae	(Rizwan et al. 2015)
Cultivation		
CU-1	Open pond system	(Rizwan et al. 2015)
CU-2	Photobioreactor	(Rizwan et al. 2015)
Harvesting		
HA-1	Flocculation with poly electrolyte	(Rizwan et al. 2015)
HA-2	Flocculation with NaOH	(Rizwan et al. 2015)
HA-3	Flocculation with PGA	(Rizwan et al. 2015)
HA-4	Flocculation with chitosan acid solution	(Rizwan et al. 2015)
HA-5	Bioflocculation + centrifugation	(Rizwan et al. 2015)
HA-6	Centrifugation	(Rizwan et al. 2015)
HA-7	Auto flocculation (induced by high pH)	(Rizwan et al. 2015)
HA-8	Microfiltration + centrifugation	(Rizwan et al. 2015)
Pretreatment step 1		
PR1-1	Acid-catalyzed esterification	(Zhang et al. 2003)
PR1-2	Grinding in liquid nitrogen	(Rizwan et al. 2015)
PR1-3	Drying + ultrasound	(Rizwan et al. 2015)
PR1-4	Drying + grinding + microwave + ultrasound	(Rizwan et al. 2015)
PR1-5	Drying	(Rizwan et al. 2015)
Pretreatment step 2		
PR2-1	Glycerine washing and methanol recovery	(Zhang et al. 2003)
PR2-2	Grinding-assisted lipid extraction	(Rizwan et al. 2015)
PR2-3	Ultrasound-assisted extraction by [Bmim][$MeSO_4$]	(Rizwan et al. 2015)
PR2-4	Ultrasound- and microwave-assisted lipid extraction	(Rizwan et al. 2015)
PR2-5	Wet lipid extraction	(Rizwan et al. 2015)
PR2-6	Solvent extraction (Bligh and Dyer's method)	(Rizwan et al. 2015)
PR2-7	Solvent extraction (Modified Bligh and Dyer's Method)	(Rizwan et al. 2015)
PR2-8	Supercritical fluid extraction	(Rizwan et al. 2015)
PR2-9	Extraction by ionic liquids mixture	(Rizwan et al. 2015)
PR2-10	Extraction by [Bmim][$MeSO_5$]	(Rizwan et al. 2015)
Reaction step 1		
R1-1	Alkali-catalyzed transesterification	(Zhang et al. 2003)
R1-2	Acid-catalyzed transesterification	(Zhang et al. 2003)
R1-3	Solvent-free enzyme-catalyzed transesterification (step 1)	(Sotoft et al. 2010)
R1-4	Co-solvent enzyme-catalyzed transesterification	(Sotoft et al. 2010)
R1-5	Base catalyzed transesterification	(Rizwan et al. 2015)
R1-6	Acid catalyzed transesterification	(Rizwan et al. 2015)

(continued)

Table 22.1 (continued)

Reaction step 1		
R1-7	Enzymatic transesterification	(Rizwan et al. 2015)
R1-8	Alkaline in-situ transesterification	(Rizwan et al. 2015)
R1-9	Acidic in-situ transesterification	(Rizwan et al. 2015)
R1-10	Enzymatic in-situ transesterification	(Rizwan et al. 2015)
Intermediate separation 1		
IS1-1	Removal of aqueous phase	(Sotoft et al. 2010)
Reaction step 2		
R2-1	Solvent-free enzyme-catalyzed transesterification (step 2)	(Sotoft et al. 2010)
Intermediate separation 2		
IS2-1	Removal of aqueous phase	(Sotoft et al. 2010)
Reaction step 3		
R3-1	Solvent-free enzyme-catalyzed transesterification (step 2)	(Sotoft et al. 2010)
Downstream separation step 1		
S1-1	Methanol recovery via distillation + recycle	(Zhang et al. 2003)
S1-2	Methanol recovery via distillation + recycle	(Zhang et al. 2003)
S1-3	Removal of aqueous phase in decanter	(Sotoft et al. 2010)
S1-4	Methanol recovery via distillation + recycle	(Sotoft et al. 2010)
Downstream separation step 2		
S2-1	Aqueous phase separation from oil-phase	(Zhang et al. 2003)
S2-2	Neutralization	(Zhang et al. 2003)
S2-3	Phase separation	(Zhang et al. 2003)
S2-4	Phase separation	(Sotoft et al. 2010)
Downstream separation step 3		
S3-1	Neutralization	(Zhang et al. 2003)
S3-2	Solids removal	(Zhang et al. 2003)
S3-3	Neutralization	(Zhang et al. 2003)
Downstream separation step 4		
S4-1	Solids removal	(Zhang et al. 2003)
S4-2	Phase separation	(Zhang et al. 2003)
S4-3	Solids removal	(Zhang et al. 2003)
Downstream separation step 5		
S5-1	FAME purification	(Zhang et al. 2003)
S5-2	FAME purification	(Zhang et al. 2003)
S5-3	FAME purification	(Zhang et al. 2003)
S5-4	FAME purification	(Sotoft et al. 2010)
S5-5	FAME purification	(Sotoft et al. 2010)
S5-6	FAME purification	(Rizwan et al. 2015)

(continued)

Table 22.1 (continued)

Downstream separation step 6		
S6-1	Glycerol purification	(Zhang et al. 2003)
S6-2	Glycerol purification	(Zhang et al. 2003)
S6-3	Glycerol purification	(Zhang et al. 2003)
S6-4	Glycerol purification	(Sotoft et al. 2010)
S6-5	Glycerol purification	(Sotoft et al. 2010)
S6-6	Glycerol purification	(Rizwan et al. 2015)
Pretreatment of residue		
PR-1	Enzymatic hydrolysis	(Rizwan et al. 2015)
Conversion of residue		
CR-1	Fast pyrolysis	(Rizwan et al. 2015)
CR-2	Fermentation	(Rizwan et al. 2015)
CR-3	Anaerobic digestion	(Rizwan et al. 2015)
Products		
P-1	Biodiesel	
P-2	Glycerol	
P-3	Bio-oil	
P-4	Bioethanol	
P-5	Biogas	

Table 22.2 Problem statistics (superstructure and model)

No. feedstock	2
No. products	2
No. processing steps (excl. RM & P)	13
No. processing intervals (excl. RM & P)	34
No. equations	79118
No. variables	76678
Discrete variables	78
Problem type	MIP
Solver	CPLEX

Table 22.3 Selected intervals

RM-2, PR1-1, PR2-1, R1-1, S1-1, S2-1, S3-1, S4-1, S5-1, S6-1, P-1, P-2

green and with thicker lines. It has been found that the optimal raw material is waste cooking palm oil and the optimal processing includes a pretreatment for conversion of free fatty acids.

As highlighted in the Introduction (Sect. 22.1) of this chapter, the presented methodology corresponds to the synthesis stage (the first stage in the three-stage

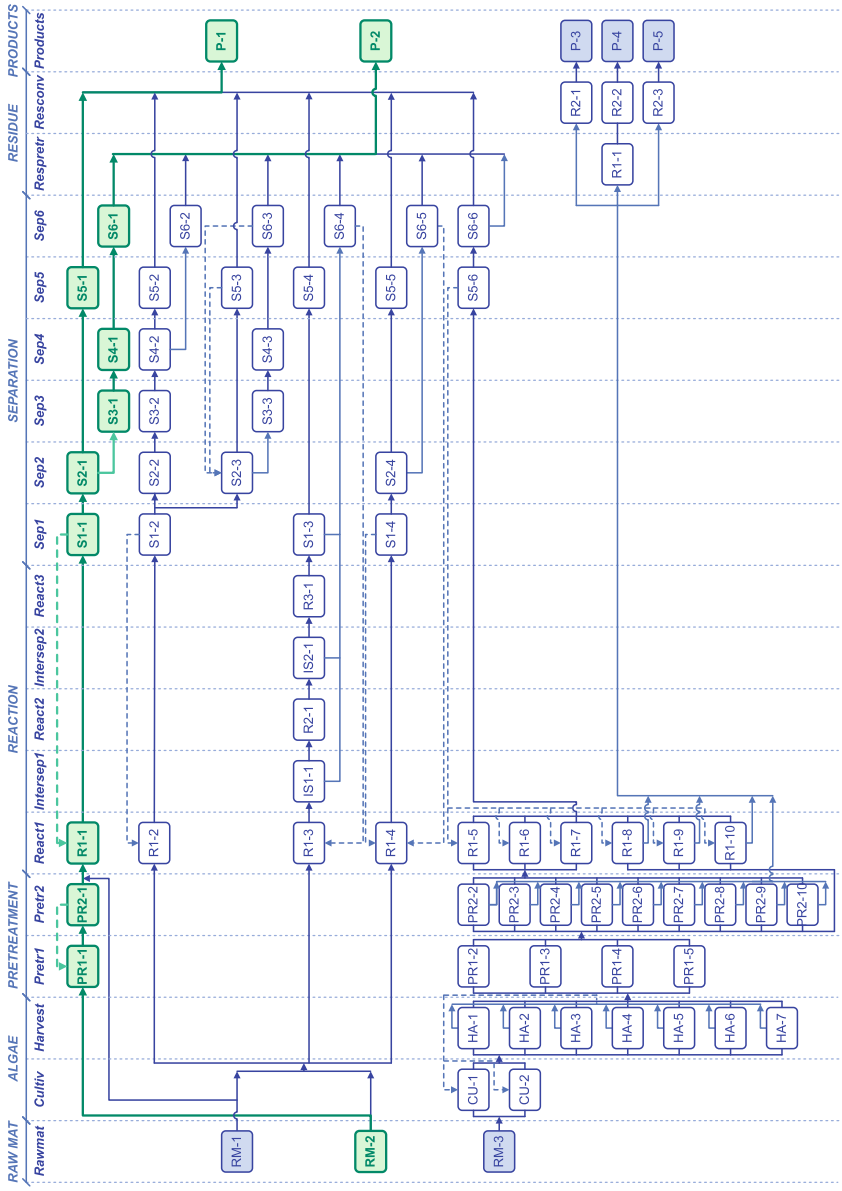

Fig. 22.11 Superstructure for problem 1 and solution. The processing intervals that are not considered due to unavailability of the raw material are *shaded*, the remaining are considered, the optimal topology is highlighted in *green* color and *thicker lines*

Table 22.4 Stream summary

	Process interval										
	PR1-1	PR2-1	R1-1	S1-1	S2-1	S3-1	S4-1	S5-1	S6-1	P-1	P-2
Triolein	94.000	94.000	93.060	4.653			4.653	4.653			
Oleic	6.000										
Methanol	16.795	17.871	12.370	13.159	0.434	0.434	0.790	0.356	0.434	0.071	
Moleate		6.298	6.298	95.108	0.113	0.113	95.108	94.995	0.113	94.621	0.113
Glycerol				9.195	9.195	9.195	9.195		9.195		9.195
Water		0.383	0.046	0.046	0.041	0.881	0.046	0.004	0.881	4.9E-04	0.527
NaOH		0.934	0.934	1.864	1.864		1.864				
H3PO4						0.013					
Na3PO4						2.547					

approach). The output of this stage not only includes the selected raw materials, technologies and products, but also calculated values of variables related to the technologies which allow for further analysis even prior to further simulation or manipulation of the results. For example, the simple mass balance data is obtained, which is given in Table 22.4.

This represents the end of stage 1 in the three-stage approach mentioned in Sect. 22.4, in the next stage, the design stage, detailed process simulation and analyses tools are used to simulate and analyse the selected process network(s) with respect to sustainability. The results of the analyses are used to identify process bottlenecks and find design alternatives that match the improvement targets.

22.5.2 Problem 2

This section contains the second example problem.

22.5.2.1 Step 1: Problem Definition

The objective is to determine the optimal topology for the production of biodiesel in South Korea by maximizing the Gross Operating Income (GOI) calculated from sales of product, purchase of raw materials, cost of chemicals and solvents added to the process, and consumption of utilities. This problem has been formulated and solved by Rizwan et al. (2015). We have used the problem formulation and data and applied the methodology of Sect. 22.4 in Super-O.

22.5.2.2 Step 2: Superstructure Generation

Based on the data available in ProBioRefine database, one raw material is considered, namely microalgae, giving as main product biodiesel and four co-products.

22.5.2.3 Step 3: Solution of the Optimization Problem

The optimization problem is solved using an external optimization software, the generic model and the data from the database regarding the superstructure of alternatives. The problem statistics are given in Table 22.5.

The complete superstructure for the problem is given in Fig. 22.12, where the disregarded alternatives due to unavailability of raw materials are shaded and the optimal solution in terms of process topology is highlighted.

The optimal solution has a negative objective function value of −46.493, corresponding to the highlighted route in Fig. 22.12. The negative objective function indicates that the production of biodiesel from microalgae in South Korea, according to the current data and available technologies, is not economically feasible. The aforementioned conclusion provides little motivation towards moving on to the second stage (within the three-stage approach) and performing detailed design and analysis of the selected alternative(s). However, in order to direct further process developments, a systematic analysis of the results obtained from the optimization problem can help target improvements. This can be done with the obtained results and no or very simple manipulations of them.

The objective function for this problem consists of sales of product and operating costs, which include the consumption raw materials, chemicals and utilities. Therefore, the distribution of operating costs should be analyzed to target improvements, which is shown in Fig. 22.13. It is seen from Fig. 22.13 that the largest contribution to the operating cost is due to the first pretreatment step, followed by the transesterification reaction. The former has the larger contribution in terms of utility cost, whereas the latter contributes in terms of chemical cost, values are given in Table 22.6.

Table 22.5 Problem statistics (superstructure and model)

No. feedstock	1
No. products	5
No. processing steps (excl. RM & P)	8
No. processing intervals (excl. RM & P)	34
No. equations	79118
No. variables	76680
Discrete variables	80
Problem type	MIP
Solver	CPLEX

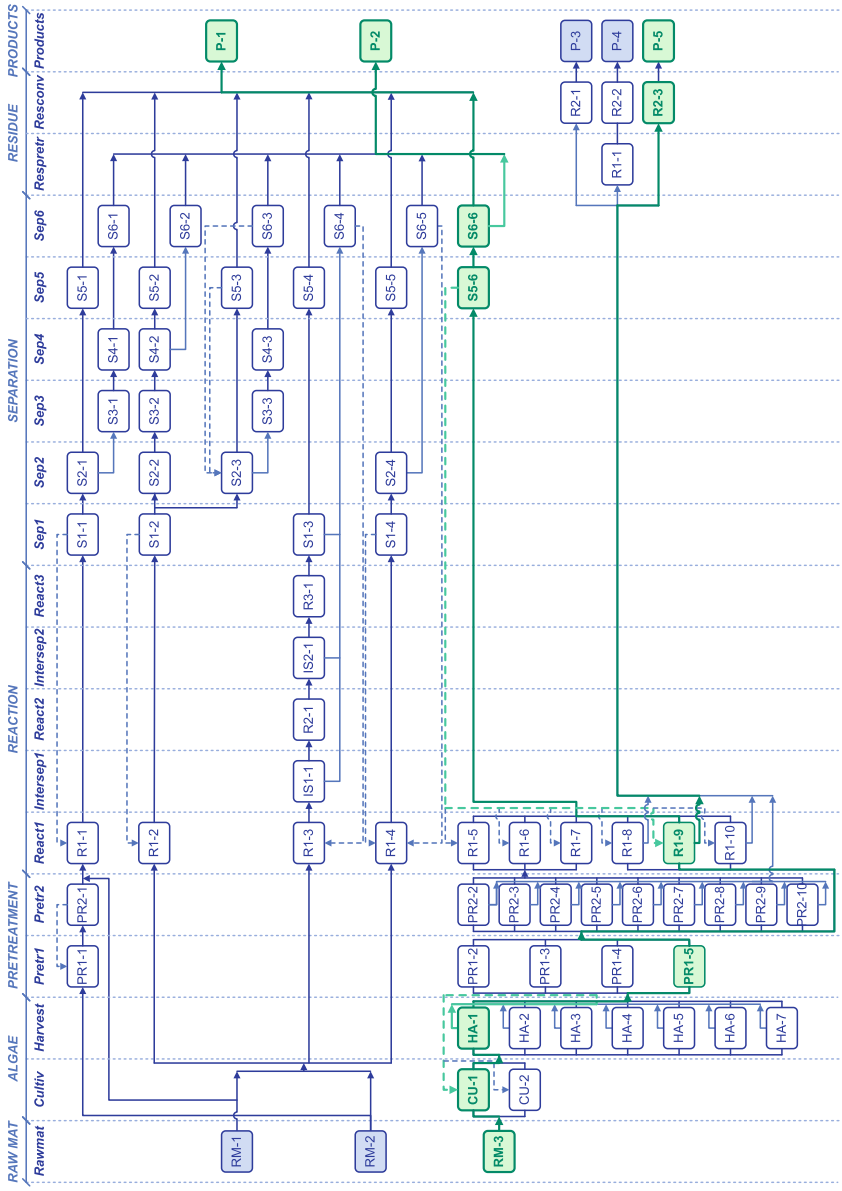

Fig. 22.12 Superstructure for problem 2 and solution. *Red* indicates infeasible alternatives, *blue* are feasible alternatives and *green* is for selected alternatives

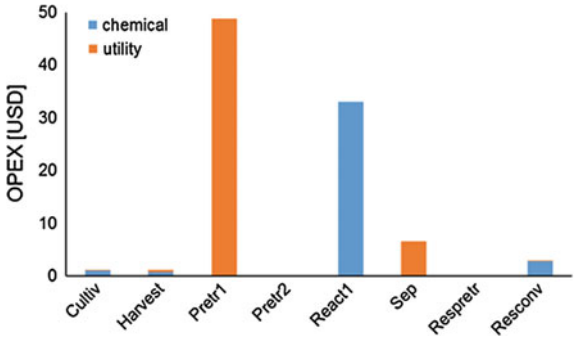

Fig. 22.13 Distribution of operating costs for problem 2

Table 22.6 Distribution of operating costs for problem 2

	Processing step							
	Cult	Harvest	Pretr1	Pretr2	React1	Sep	Respretr	Resconv
Chemical	1.124	0.750			32.977	1.6E-05		2.740
Utility	0.026	0.500	48.737	1.4E-03		6.555		0.258

From this simple analysis based on results obtained in the optimization, one can conclude that improvements in steps of the processing other than the two previously mentioned would not be significant towards obtaining an economically feasible process under the current prices. In order to obtain a positive objective function, a minimum reduction of the total operating cost of both steps of 43.1 % is required. Changes in costs of raw materials and products could also bring improve the economics of the selected process. A sensitivity analysis can be performed based on data from the optimization, shown in Table 22.7.

The sensitivity analysis reveals that an increase of biodiesel price between 110 and 120 % would be required in order to obtain a profitable process, if no other changes are considered.

Table 22.7 Distribution of operating costs for problem 2

Increase in biodiesel price (%)	Objective function
0	–46.493
1	–46.052
10	–42.35
100	–5.243
120	3.042

22.6 Outlook

A methodology for sustainable synthesis of biorefinery networks has been presented, described and applied to the synthesis of biofuel production processes. The developed methodology integrates the steps with the necessary models and tools, and has been implemented in Super-O, a user-friendly software interface, which has been tested for various problems of varying type, size and complexity. The case studies solved so far have shown that the use of the methodology through Super-O enables a fast formulation, efficient data management, and allows large problems to become manageable.

Future developments are directed towards populating the ProBioRefine database and solving problems of increased size and complexity, that is with more alternatives and including more data, for example the capital cost of the processing intervals.

Acknowledgments The authors would like to thank the members of the ProBioRefine initiative.

References

Babi, D. K., Holtbruegge, J., Lutze, P., Gorak, A., Woodley, J. M., & Gani, R. (2015). Sustainable process synthesis-intensification. *Computers and Chemical Engineering*, *81*, 218–244.

Bertran, M. O., Frauzem, R., Zhang, L., & Gani, R. (2016). A generic methodology for superstructure optimization of different processing networks. In: Z. Kravanja (Ed.), *Proceedings of the 26th European Symposium on Computer Aided Process Engineering—ESCAPE 26*. Portoroz, Slovenia: Elsevier B.V.

Drobez, R., Pintaric, N. Z., & Pahor, B. (2009). MINLP synthesis of processes for the production of biogas from organic and animal waste. *Chemical and Biochemical Engineering Quarterly*, *23*(4), 445–459.

Floudas, C. (1995). *Nonlinear and mixed-integer optimization: fundamentals and applications*. New York: Oxford University Press.

Garcia, D. J., & You, F. (2015). Multiobjective optimization of product and process networks: General modeling framework, efficient global optimization algorithm, and case studies on bioconversion. *AIChE Journal*, *61*(2), 530–554.

González-Delgado, A., & Kafarov, V. (2015). Development of a topology of microalgae-based biorefinery: Process synthesis and optimization using a combined forwardbackward screening and superstructure approach. *Clean Technologies and Environmental Policy*, *17*, 2213–2228.

Grossmann, I. (1985). Mixed-integer programming approach for the synthesis of integrated process flowsheets. *Computers and Chemical Engineering*, *9*(5), 463–482.

Grossmann, I. (1989). *MINLP optimization strategies and algorithms for process synthesis*. Tech. rep.: Carnegie Mellon University.

Grossmann, I. (1990). Mixed-integer nonlinear programming techniques for the synthesis of engineering systems. *Research in Engineering Design*, *1*, 205–228.

IEA. (2011). *Technology roadmap: Biofuels for transport*. OECD/IEA: Tech. rep.

IEA. (2012). *Energy technology perspectives 2012: Scenarios and strategies to 2050*. OECD/IEA: Tech. rep.

Kravanja, Z., & Grossmann, I. (1997). Multilevel-hierarchical MINLP synthesis of process flowsheets. *Computers and Chemical Engineering*, *21*, S421–S426.

Martín, M., & Grossmann, I. (2012). BIOpt: A library of models for optimization of biofuel production processes. *Computer Aided Chemical Engineering*, *30*, 16–20.

Quaglia, A. (2013). An integrated business and engineering framework for synthesis and design of processing networks. PhD thesis, Technical University of Denmark.

Quaglia, A., Sarup, B., Sin, G., & Gani, R. (2012). Integrated business and engineering framework for synthesis and design of enterprise-wide processing networks. *Computers and Chemical Engineering*, *38*, 213–223.

Quaglia, A., Pennati, A., Bogataj, M., Kravanja, Z., Sin, G., & Gani, R. (2014). Industrial process water treatment and reuse: A framework for synthesis and design. *Industrial & Engineering Chemistry Research*, *53*(13).

Rizwan, M., Lee, J. H., & Gani, R. (2013). Optimal processing pathway for the production of biodiesel from microalgal biomass: A superstructure based approach. *Computers and Chemical Engineering*, *58*, 305–314.

Rizwan, M., Lee, J., & Gani, R. (2015). Optimal design of microalgae-based biorefinery: Economics, opportunities and challenges. *Applied Energy*, *150*, 69–79.

Sotoft, L., Rong, B., Christensen, K., & Norddahl, B. (2010). Process simulation and economical evaluation of enzymatic biodiesel production plant. *Bioresource Technology*, *101*, 5266–5274.

Yeomans, H., & Grossmann, I. E. (1999). A systematic modeling framework of superstructure optimization in process synthesis. *Computers and Chemical Engineering*, *23*(6), 709–731.

Zhang, Y., Dube, M., McLean, D., & Kates, M. (2003). Biodiesel production from waste cooking oil: 1. process design and technological assessment. *Bioresource Technology*, *89*, 1–16.

Zondervan, E., Nawaz, M., de Haan, A. B., Woodley, J. M., & Gani, R. (2011). Optimal design of a multi-product biorefinery system. *Computers and Chemical Engineering*, *35*(9), 1752–1766.

Chapter 23
Systematic Design of Biorefinery Downstream Processes

Michele Corbetta, Ignacio E. Grossmann, Carlo Pirola and Flavio Manenti

Abstract Downstream processing of biofuels and bio-based chemicals often represents the bottleneck for the economic sustainable development of new processes. It is also a challenging problem for process synthesis and optimization, due to the intrinsic nonideal thermodynamics of the liquid mixtures derived from the (bio)chemical conversion of biomass. In this Chapter, a recent mathematical framework is outlined for the structural and parameter optimization of process flowsheets with rigorous and detailed models. The optimization problem is formulated within the Generalized Disjunctive Programming (GDP) framework and the solution of the reformulated MINLP problem is approached with a decomposition strategy based on the Outer-Approximation algorithm. At first, the mathematical formulation and the numerical implementation are outlined. In the second portion of the Chapter, several validation examples in the field of biorefineries are proposed spanning from the economic optimization of single distillation columns, the dewatering task of diluted bio-mixtures, up to the distillation sequencing with simultaneous mixed-integer design of each distillation column for a quaternary mixture in the presence of azeotropes.

23.1 Introduction and State of the Art

The renewed interest in the field of distillation has been recently promoted by the consistent research on biomass conversion technologies to biofuels and bio-based chemicals. These technologies are based on (bio)chemical reactors that produce

M. Corbetta (✉) · F. Manenti
Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano, P.zza L. Da Vinci 32, 20133 Milan, Italy
e-mail: michele.corbetta@polimi.it

I.E. Grossmann
Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, 15213 Pittsburgh, PA, USA

C. Pirola
Dipartimento di Chimica, Università degli Studi di Milano, Via Golgi 19, 20133 Milan, Italy

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_23

highly diluted aqueous solutions. The downstream processing of those mixtures usually involves distillation, leading to high operating costs due to the high heat of vaporization of water (Chen 2009; Xiu and Zeng 2008). For this reason, attempts to optimize and thermally integrate the purification step (Ahmetovic et al. 2010; Dias et al. 2009; Karuppiah et al. 2008) result in a relevant lowering of the production costs that reduces the economic gap with respect to cheaper fossil-based products (Hermann and Patel 2007; Sauer et al. 2008). In addition, the optimization of this type of downstream processes, as opposed to hydrocarbon distillation, involves highly nonideal liquid mixtures that demand rigorous thermodynamic models.

In this context, process simulators offer a reliable and rigorous modeling environment that rely on extensive thermodynamic properties databanks and tailored distillation algorithms, in contrast with equation-oriented GDP/MINLP optimization tools that are usually based on shortcut models for the unit operations and for the estimation of physical and thermodynamic properties (Navarro-Amoros et al. 2013). Unfortunately, it has been demonstrated that the optimization tools available within commercial simulation packages are not as effective and flexible as it would be required (Biegler 1985) due to the high nonlinearity of the equation systems, and to the impossibility to optimize structural (integer) decision variables. This was the motivation for several authors to develop ad hoc interfaces for the process simulator-based optimization with MINLP optimization algorithms. Two main strategies have been proposed; the one based on the augmented penalty/equality relaxation outer-approximation (AP/ER/OA) deterministic algorithm (Viswanathan and Grossmann 1990), and the ones based on metaheuristic methods, such as the evolutionary algorithms (Gross and Roosen 1998).

Starting from the deterministic approach, Harsh and co-workers developed an MINLP algorithm for the retrofit of chemical plants with fixed topology based on the FLOWTRAN process simulator, and they applied it to the ammonia synthesis process (Harsh et al. 1989). Other authors (Diaz and Bandoni 1996) derived an MINLP approach to optimize the structure and the operating parameters of a real ethylene plant in operation, interfacing a specific simulation code. Caballero et al. (Caballero et al. 2005) proposed an optimization algorithm for the rigorous design of single distillation columns using Aspen HYSYS and Brunet et al. (2012) applied the same methodology to assist decision makers in the design of environmentally conscious ammonia–water absorption machines for cooling and refrigeration. Very recently, Navarro-Amoros et al. (2014) proposed a new algorithm for the structural optimization of process superstructures within the Generalized Disjunctive Programming (GDP) framework. Finally, Garcia et al. (2014) proposed a hybrid simulation-multiobjective optimization approach that optimizes the production cost and minimizes the associated environmental impacts of isobutane alkylation. The simultaneous process optimization and heat integration approach has been also addressed by coupling process simulators with external equation systems (Chen et al. 2015; Navarro-Amoros et al. 2013).

On the other hand, several authors have proposed optimization algorithms based on evolutionary methods in order to overcome some difficulties that arise from the use of deterministic nonlinear programming solvers with real-world complex

problems. For instance, Gross and Roosen (1998) addressed the simultaneous structural and parameter optimization in process synthesis coupling Aspen Plus with evolutionary methods. Similarly, an optimization framework is proposed (Leboreiro and Acevedo 2004) for the synthesis and design of complex distillation sequences, based on a modified genetic algorithm coupled with a sequential process simulator, succeeding in problems where deterministic mathematical algorithms had failed. Vazquez-Castillo et al. (2009) addressed the optimization of intensified distillation systems for quaternary distillations with a multiobjective genetic algorithm coupled to the Aspen Plus process simulator. Subsequently, Gutierrez-Antonio and Briones-Ramirez (2009) implemented a multiobjective genetic algorithm coupled with Aspen Plus to obtain the Pareto front of Petlyuk sequences. Bravo-Bravo and co-workers (Bravo-Bravo et al. 2010) proposed a novel extractive dividing wall distillation column, which has been designed using a constrained stochastic multi-objective optimization technique, based on the use of GA algorithms. Finally, Eslick and Miller (2011) developed a modular framework for multi-objective analysis aimed at minimizing freshwater consumption and levelized cost of electricity for the retrofit of a hypothetical 550 MW subcritical pulverized coal power plant with an MEA-based carbon capture and compression system.

In the following, the novel interface between the process simulator PRO/II (SimSci, Schneider-Electric) and GAMS based on the recent work of Corbetta et al. (2016), Corbetta (2015) is proposed and applied to the structural and parameter optimization of biorefinery downstream processes. The optimization tool is applied to several case studies, including the distillation sequencing with simultaneous mixed-integer optimal design of each distillation column for a quaternary mixture in presence of azeotropes.

23.2 Mathematical Methods

The mathematical framework is based on three fundamental pillars: the modeling, the simulation and the optimization of biorefinery concepts, as shown in Fig. 23.1.

Looking at the block diagram, the methodology requires two inputs that are a process superstructure and experimental data, and it returns an optimized and validated flowsheet as output.

This framework moves from the acquisition of experimental data on kinetics and phase equilibria of the bio-based mixtures under analysis. This step is of paramount importance for the modeling of biorefinery conversion processes because of the lack of reliable predictive submodels. For the sake of clarity, submodels are here intended as systems of closure equations for first-principles models. As an example, the rigorous modeling of a distillation column includes the UNIQUAC model as a closure equation that allows to evaluate the liquid activity coefficient, along with constitutive MESH equations (Mass balances, Equilibrium equations, Stoichiometric constraints, Heat balances).

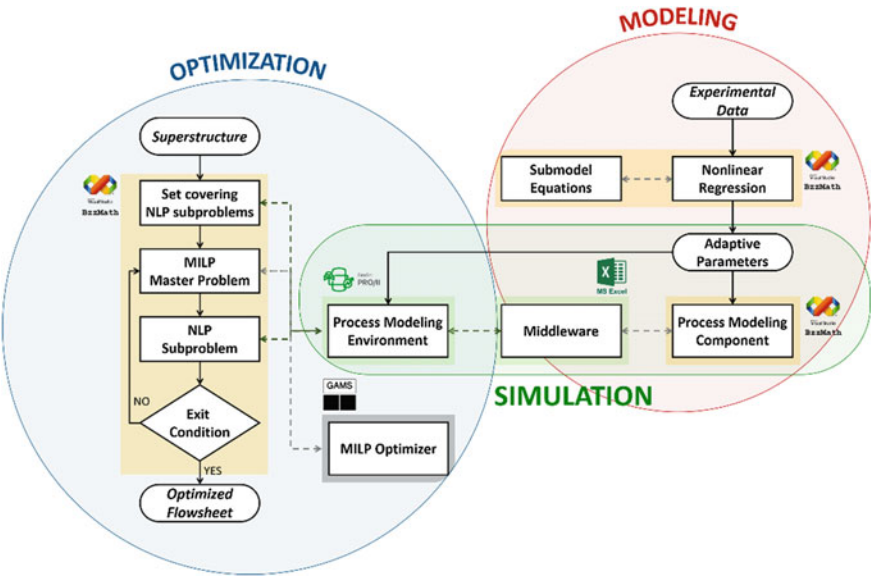

Fig. 23.1 Methodology for the modeling, simulation, and optimization of biorefinery concepts

Experimental data are the main input of the modeling activity to derive adaptive parameters of thermodynamic and kinetic submodels by means of nonlinear regression tools. After the tuning of the models, a critical comparison between model predictions and experimental data allows to validate the simulation tools or to identify possible improvements.

The collection of constitutive equations (e.g. material and energy balances) and closure submodels (e.g. UNIQUAC liquid activity coefficient model, heat transfer coefficient models) results in NLS, ODE and PDE systems of equations that require robust numerical solvers. With respect to surrogate models, the main feature of mechanistic models depends on their reliability and validity on a wider space of operating variables.

Validated models, also called Process Modeling Components (PMC), can be embedded inside the Process Simulation Environment (PSE) of commercial software as outlined in the work of Corbetta et al. (2014). This step opens the possibility to link the experiments-based modeling step with the model-based simulation and optimization of conversion and purification processes.

Finally, new rigorous MINLP process optimization algorithms are applied to find the best flowsheet design starting from a comprehensive superstructure that includes all the possible technical solutions.

23.2.1 Optimization Theory

Discrete-continuous optimization problems are commonly modeled in algebraic form as Mixed-Integer Linear (MILP) or Nonlinear (MINLP) Programming models.

Over the last few years, there has been an evident increase in the development of these kind of models, particularly in process systems engineering (Kallrath 2000; Karuppiah and Grossmann 2006). Mixed-integer linear programming (MILP) methods and codes have been available and applied to many practical problems for more than 20 years (Nemhauser and Wolsey 1988). On the other hand, only recently several new methods and codes are becoming available for mixed-integer nonlinear problems (Grossmann and Kravanja 1995; Grossmann 2002).

The general formulation of a MINLP problem is reported in Eq. (23.1), where x and y are bounded continuous and integer decision variables, f is the objective (non)linear function to minimize, and g is a (non)linear constraint.

$$\begin{aligned} \min \quad & Z = f(\mathbf{x}, \mathbf{y}) \\ \text{s.t.} \quad & g(\mathbf{x}, \mathbf{y}) \leq 0 \\ & \mathbf{x}^{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}^{\mathrm{U}} \\ & \mathbf{y}^{\mathrm{L}} \leq \mathbf{y} \leq \mathbf{y}^{\mathrm{U}} \\ & \mathbf{x} \in R^n,\ \mathbf{y} \in N^m \end{aligned} \tag{23.1}$$

Figure 23.2 provides a graphical representation of an application example dealing with the optimization of the reflux ratio and number of trays of a distillation column. The two decision variables are bounded and define a feasible region along with a nonlinear product purity specification (g). Since the number of stages is an integer variable, the feasible region is only the union of the red segments. Moreover, the contour plot of a hypothetic objective function are plotted.

Among the MINLP solution methods, the Outer Approximation (Duran and Grossmann 1986; Viswanathan and Grossmann 1990) is a decomposition strategy that splits the MINLP problem in a series of NLP subproblems and MINLP master problems. The NLP subproblem S(yk) derives from the original MINLP problem fixing the value of the integer variables (yk), while the MILP master problem (MOA, k) is an accumulation of k linearizations of the original MINLP problem, where the linearization points are given by the NLP subproblems.

Figure 23.3 shows an example of the accumulated linearizations of an objective function and a feasible region with two continuous decision variables. It should be

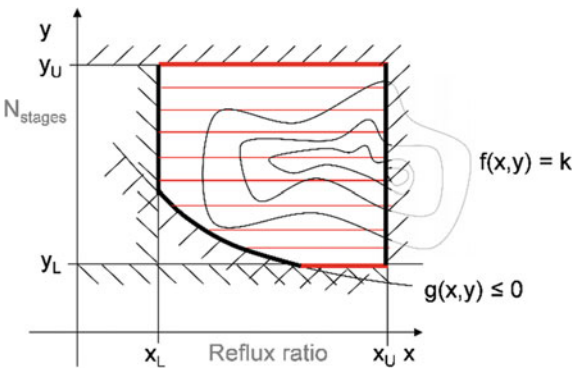

Fig. 23.2 Graphical interpretation of a MINLP problem for a distillation column

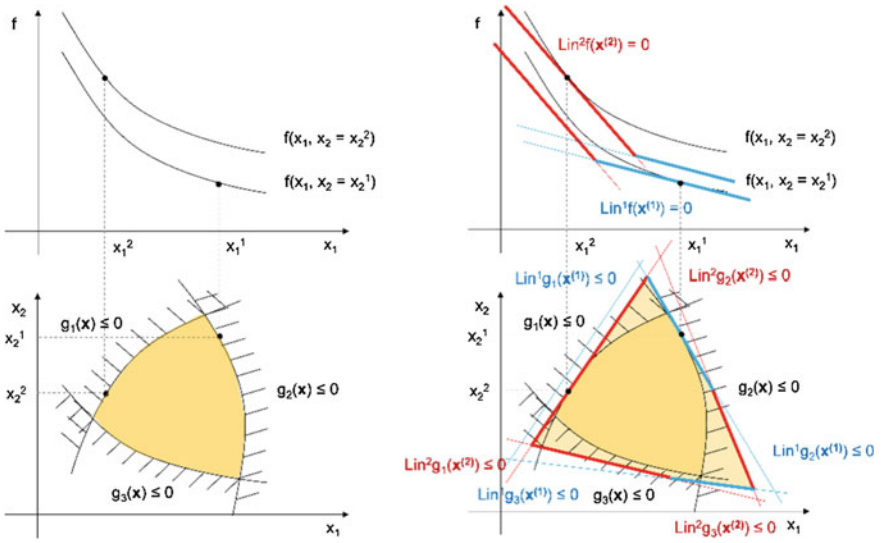

Fig. 23.3 Accumulated linearizations of a convex objective function and feasible region

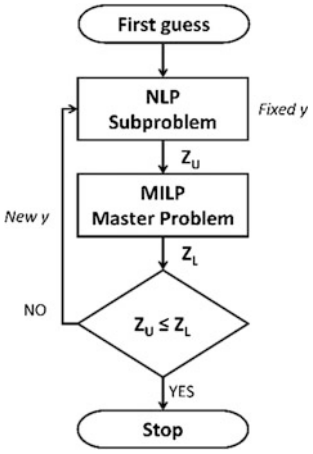

Fig. 23.4 Block diagram of the OA algorithm for the solution of MINLP problems

noted that for convex functions the linearizations underestimate the objective function and overestimate the feasible region.

The algorithm (Fig. 23.4) moves from the definition of a first guess for the decision variables that is given to the NLP subproblem. The output of this step is an optimized value of the continuous variables and the upper bound for the objective function. The corresponding linearizations are added to the MILP master problem, which, in turn, gives the updated value of the integer decision variables and the lower bound of the objective function (it is proved for convex problems for the geometric implications discussed above). The algorithm iterates between NLP and

MILP problems until the exit condition is satisfied (ZU $\leq$ ZL) assuming integer cuts are added at each iteration.

Since MINLP models can be formulated in different ways, leading either to solvable or non-solvable problems, there is a need for a systematic modeling framework that provides a fundamental understanding on the nature of these models (Grossmann and Trespalacios 2013).

Generalized Disjunctive Programming (GDP) is a modeling framework that represents problems in terms of Boolean and continuous variables, allowing the representation of constraints as algebraic equations, disjunctions and logic propositions. Unlike direct MINLP formulations, this higher-level modeling framework makes the formulation process more intuitive and systematic, while retaining in the model the underlying logic structure of the problem. It should be noted that since (MINLP) is expressed through algebraic equations, the modeler has to directly express the logic in the format of f(x, y) and g(x, y), while in (GDP) the disjunctions capture the logic in continuous form and the logic propositions capture the logic in the Boolean space. Particularly in process synthesis, the decisions are normally associated as to whether certain equipment should be included or not in a process flowsheet. If the equipment is selected, then the mass and energy balance, physical and chemical equilibrium (if any), and cost constraints need to be satisfied. If it is not selected, then all the equations can be ignored. Therefore, in PSE GDP problems will normally take the following form (Eq. 23.2):

$$
\begin{aligned}
\min \quad & Z = \sum_{k \in K} c_k + f(\mathbf{x}) \\
\text{s.t.} \quad & r(\mathbf{x}) \leq 0 \\
& \vee_{j \in J_k} \begin{bmatrix} Y_{jk} \\ g_k(\mathbf{x}) \leq 0 \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K \\
& \Omega(Y) = true \\
& \mathbf{x}^{\mathbf{L}} \leq \mathbf{x} \leq \mathbf{x}^{\mathbf{U}} \\
& \mathbf{x} \in R^n,\ \mathbf{c} \in R^m,\ \mathbf{Y} \in \{true; false\}^m
\end{aligned}
\tag{23.2}
$$

The GDP formulation is divided in three main sections:

1. Objective function and global constraints. The objective is a function of the continuous variables x and of the conditional equipment costs (c_k). It contains the global constraints $r(\mathbf{x}) \leq 0$ that must hold true regardless of the discrete decisions.
2. Disjunctions. The logic in the continuous space is represented by a set of disjunctions, k $\in$ K, each of which contains j $\in$ J_k terms, linked by an OR operator. Each term of the disjunction has a Boolean variable Y_{jk}, an associated set of inequalities and the cost associated with the conditional equipment. Exactly one of the Boolean variables can be selected in each disjunction. For an active term in a disjunction, the corresponding inequalities and cost are enforced. When the term is not active, the corresponding constraints are ignored.

3. Logic propositions. The symbolic equation $\Omega(Y) = \text{true}$ represents the set of logic propositions that relates the Boolean variables. It is possible to transform the logic propositions in algebraic form. This step is realized replacing the Boolean variables with binary variables and applying the Boolean algebra rules (removing implications, applying De Morgan's laws and distributivity of OR over AND).

To solve GDP problems it is possible to use some special techniques such as disjunctive branch and bound (Lee and Grossmann 2001) and Logic-Based Outer Approximation (Turkay and Grossmann 1996), or they can be reformulated as MINLP to exploit the developments in these solvers.

This theory can be applied to the topology optimization of complex process superstructures based on rigorous thermodynamic models, with special emphasis on distillation downstream processes in the biorefining area. Specifically, the distillation sequencing problem with simultaneous design of number of trays and feed tray location is addressed. Both continuous (e.g. split ratio, reflux ratio, pressure) and integer (e.g. number of trays, feed trays, equipment existence) decision variables are optimized under the Generalized Disjunctive Programming (GDP) framework (Grossmann and Trespalacios 2013), using the process modeling environment of SimSci PRO/II, and the optimization environment of GAMS.

The optimization algorithm basically requires a superstructure, the propositional logic to define the topology of the superstructure, selected flowsheets implemented in PRO/II, a set of bounded continuous and integer decision variables, nonlinear (potentially implicit) constraints (e.g. purity and safety constraints), and an economic objective function.

23.2.2 Process Superstructure

The optimization procedure starts from the definition of the process superstructure. The most general superstructure that can be handled by the algorithm developed by Corbetta et al. (2016) is based on the interconnection of permanent units with elementary conditional unit and trays modules. While permanent units are present in each possible optimal flowsheet originated from the superstructure, the elementary conditional unit and trays modules are introduced in the superstructure to describe the conditional units (or conditional sections with more than one unit) that are not necessarily present in the final optimized flowsheet. Conditional trays are introduced within the conditional unit and trays module for the rectification and stripping sections of the distillation columns potentially present (Fig. 23.5). The GDP conditional tray representation is adopted to define feed stage and number of stages of the distillation column (Barttfield et al. 2004).

It is worth to note that the superstructure is never completely solved as a unique process simulator flowsheet. Rather the model could be depicted as a collection of different possible black-box simulations, which are defined by disjunctions, in contrast with fully equation-oriented models. For this reason, only permanent units and selected

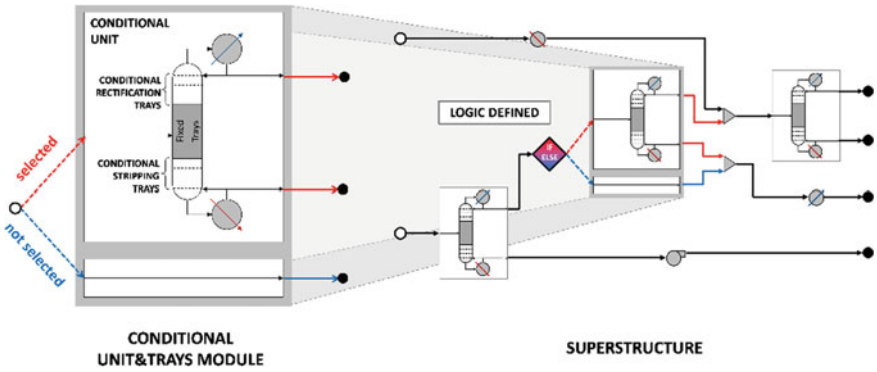

Fig. 23.5 Representations of the conditional unit and trays elementary module and of a typical superstructure

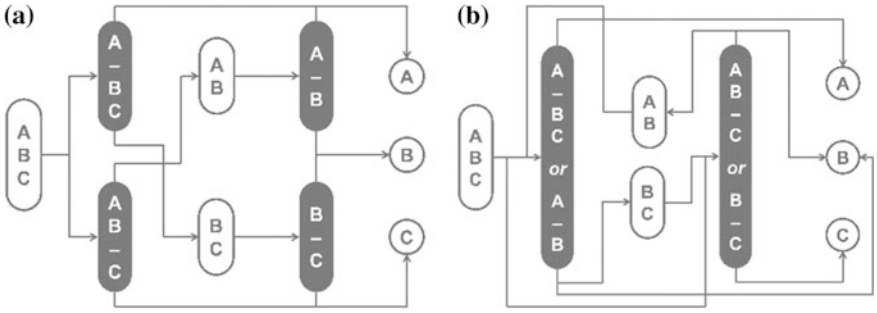

Fig. 23.6 **a** State task network and **b** state equipment network representations for the sharp distillation of a ternary mixture

conditional units are solved at each call of the process simulator. In this way, no splitters are required for conditional units, and zero-flow units are avoided.

A relevant case of this kind of superstructure arise from the solution of distillation sequencing problems. At this level, it is possible to represent the sequencing with either a State Task Network (STN) or a State Equipment Network (SEN) (Yeomans and Grossmann 1999). Figure 23.6 reports the two different superstructures for the distillation of a ternary mixture. It is possible to highlight that the SEN requires a smaller number of columns but it introduces recycles. Nevertheless, since the modeling is accomplished by the process simulator with a logic-based definition of the input file that considers only selected units, it does not matter if either STN or SEN superstructure is adopted.

23.2.3 Process Modeling

The detailed modeling can be achieved with a process simulator (SimSci PRO/II) taking advantage of thermodynamic databanks for the estimation of physical

properties and ad hoc algorithms for the solution of nonlinear systems derived from distillation columns and other unit operations. Within this process modeling environment there is also the flexibility to introduce custom modeling components, which can be required in case of nonconventional unit operations, as discussed elsewhere (Corbetta et al. 2014).

23.2.3.1 Thermodynamic Modeling

When the target is to address the synthesis of biorefinery downstream processes, it is worth mentioning that predictive thermodynamic models, such as UNIFAC, frequently fail. This is due to the highly complex nature of the interactions between oxygenated chemicals in the aqueous phase. These complex liquid mixtures can be obtained, for instance, in the form of fermentation broth withdrawn from a bioreactor or from the outlet of catalytic deoxygenation converters, and they present some common characteristics. Usually they are diluted organic aqueous solutions, in which water can represent up to 80–90 wt%. Moreover, they are made up of a large amount of components that belong to the same chemical class (e.g. ketones, alcohols, polyols), with a frequent occurrence of homogeneous azeotropes and pinch points in the corresponding VLE equilibrium diagrams. Finally, a heavy cut composed by soluble solids is usually present due to incomplete biomass conversion, presence of an inert lignin fraction, and/or production of high molecular weight components by side reactions. Consequently, developing reliable VL(L)E thermodynamic models based on nonlinear regression of (UNIQUAC or NRTL) binary interaction parameters on experimental data (Pirola et al. 2014) plays a major role in correctly predicting the distillation behavior, as will be highlighted in Sect. 23.3.

23.2.3.2 Cost Functions

Once the process simulator is set up with a proper thermodynamic model, the convergence of a flowsheet provides the value of the implicit variables for the evaluation of the economic objective function and for measuring the violation of nonlinear constraints. Specifically, the techno-economic assessment is approached with nonlinear cost functions (Douglas 1988) and rigorous sizing models embedded within the process simulator (e.g. tray and packing hydraulics). For distillation columns, the minimization of the cost objective function is performed by considering both annualized capital costs (CAPEX) and operating cost (OPEX), which added together determine the Total Annualized Costs (TAC) as reported in Eq. (23.3).

$$TAC = \frac{C_{inv}}{payback\ time} + C_{op} = \left(\frac{C_{col} + C_{internals} + C_{reb} + C_{cond}}{payback\ time} \right) + C_{steam} + C_{cw} + C_{entrainer} \tag{23.3}$$

Operating costs include utility costs (steam, cooling water and, potentially, entrainers), while column investment costs include trays or packing, column vessel, condenser and reboiler installation and purchase costs, which depend on the value of the decision optimization variables and on the size of the equipment. For CAPEX evaluation, the following function, Eq. (23.4), is adopted, where constants c_1, c_2, e_1 and e_2 depend on the equipment type, L_1 and L_2 are relevant size, FC is a parameter depending on the fabrication material and operating pressure, while M and S is the Marshall and Swift economic index.

$$C_{equipment}[\$] = c_1 \left(\frac{M \text{ and } S}{280} \right) L_1^{e_1} L_2^{e_2} (c_2 + F_{C,equiment}) \tag{23.4}$$

23.2.3.3 Generalized Disjunctive Programming Formulation

The optimization problem is formulated within the Generalized Disjunctive Programming framework (Grossmann and Trespalacios 2013), in which also implicit variables (x_I) are assumed, along with continuous decision variables (x) and Boolean decision variables (Y), as outlined in Eq. (23.5).

$$\begin{aligned}
\min \quad & Z = \sum_{k \in K} TAC_k + f(\mathbf{x}, \mathbf{x_I}) \\
\text{s.t.} \quad & \mathbf{x_I} = \left[\mathbf{x'_I}\, \mathbf{x''_I} \right] \\
& \mathbf{x'_I} = f'_I(\mathbf{x}) \\
& g_n(\mathbf{x}, \mathbf{x_I}) \le 0 \quad n \in G \\
& \left[\begin{array}{r} Y_k \\ \left[\begin{array}{r} Y_j \\ TAC_j = \gamma_j \end{array} \right] \vee \left[\begin{array}{r} \neg Y_j \\ TAC_j = 0 \end{array} \right] \\ x''_{I,k} = f''_I(\mathbf{x}) \\ g_n(\mathbf{x}, \mathbf{x_I}) \le 0 \\ TAC_k = \sum_{j \in CT_k} \gamma_j(x, y, x_I) \end{array} \right] \vee \left[\begin{array}{r} \neg Y_k \\ x''_{I,k} = 0 \\ TAC_k = 0 \end{array} \right] \quad j \in CT_k,\ n \in G_k,\ k \in K \\
& \Omega_{topology}(Y_k) = true \\
& \neg Y_k \Rightarrow \neg Y_j \quad k \in K,\ j \in CT_k \\
& \vee Y_i \quad i \in RCT_k,\ k \in K \\
& \vee Y_i \quad i \in SCT_k,\ k \in K \\
& \mathbf{x^L} \le \mathbf{x} \le \mathbf{x^U} \\
& \mathbf{x} \in R^n,\ \mathbf{x_I} \in R^{n_I},\ \mathbf{c} \in R^m,\ \mathbf{Y} \in \{true; false\}^m
\end{aligned} \tag{23.5}$$

These implicit variables are evaluated by the process simulator, which is treated as a black-box and is represented by the implicit function (f_I) that links implicit variables with decision variables. Nonlinear constraints (g) can be both global (set G) and

conditional (set G_k). The mathematical programming formulation involves the definition of a set K of disjunctions, corresponding to conditional unit and trays modules, for which additional constraints, equations, and cost contributions are defined. Each disjunction involves only two terms, corresponding to the selection or not of a conditional unit (Y_k), with embedded disjunctions for the conditional trays (Y_j) belonging to each elementary module (set of CT_k conditional rectification and stripping trays of column k). The economic objective function includes these contributions along with a general function (f) of the decision variables. Finally, the propositional logic defines the units' interconnectivity. Topology logic ($\Omega(Y) =$ true) defines the interconnection between conditional units, while the three subsequent sets of logic constraints are required to ensure that no conditional trays are selected if the conditional unit is not selected, at most only one Boolean variable should be true for the rectification (RCT_k) and for the stripping (SCT_k) conditional trays of each distillation column (either permanent or conditional).

23.2.4 Optimization Algorithm

The optimization problem is solved with a decomposition strategy based on the Logic-Based Outer-Approximation (LBOA) algorithm (Turkay and Grossmann 1996). The algorithm involves NLP subproblems that arise from fixed values at the Boolean variables Y in Eq. (23.5), and MILP master problems that correspond to a linear approximation at the GDP in Eq. (23.5). NLP subproblems are solved with a Derivative Free Optimizer (C++/BzzMath), and MILP master problems are solved with Branch and Cut methods (GAMS/CPLEX). The corresponding block diagram of this algorithm is shown in Fig. 23.7.

The LBOA algorithm requires a first guess for both the continuous and integer optimization variables. An initial NLP subproblem gives the initial values of the continuous decision variables (x_0). In case of conditional units, i.e. when the feasible flowsheet derived from the superstructure does not include necessarily all the units, a set covering problem is considered from which several NLP subproblems are solved to provide linearization points for all the process units (Turkay and Grossmann 1996). Subsequently, a main loop iterates between the solution of the MILP master problem and the NLP subproblem until the exit condition is satisfied. The MILP master problem updates the value of the integer decision variables (yMILP) and provides a first guess (xMILP) to the NLP subproblem, which in turn, updates the value of the continuous optimization variables (xNLP). The loop is terminated as soon as both the NLP worsening condition and the crossing of lower and upper bounds occur.

The C++ programming environment provides the communication platform between PRO/II and GAMS, by writing ASCII input files and retrieving results from ASCII output files of the simulator and of the optimizer, respectively. Auxiliary C++ functions allow to translate the GDP with fixed Boolean variables to a specific topology that corresponds to an MINLP with feed tray and number of trays

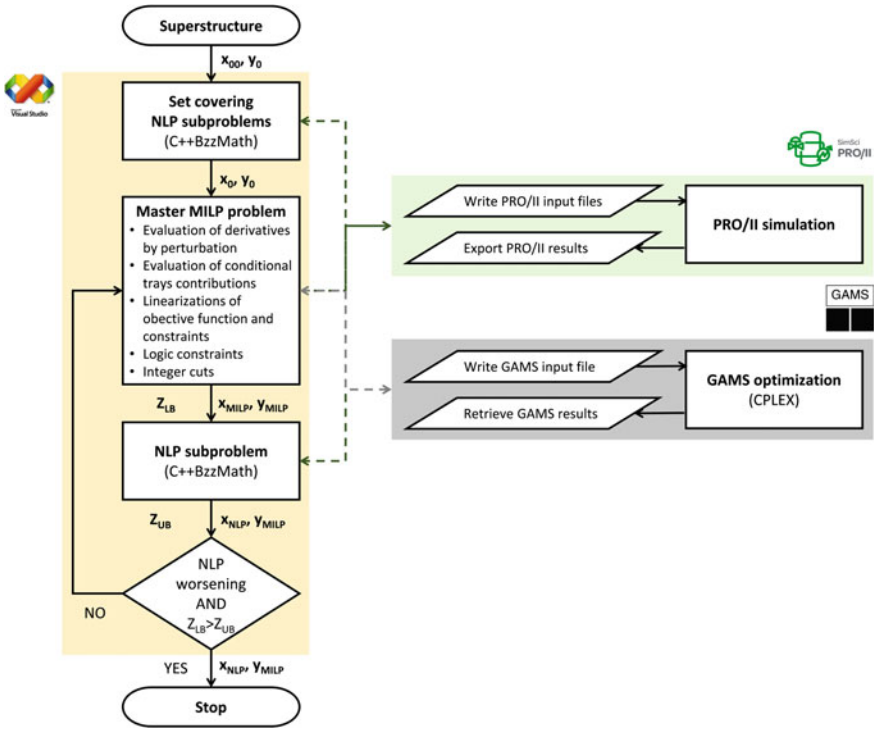

Fig. 23.7 Optimization algorithm block diagram

as variables to be determined, and to evaluate the economic objective function and the violation of nonlinear constraints. It is important to emphasize that NLP subproblems involve only selected units, thus avoiding convergence difficulties due to zero flows.

23.2.4.1 NLP Subproblems

NLP subproblems are solved with the Derivative Free Optimizer belonging to the BzzMath numerical library (Buzzi-Ferraris and Manenti 2012), which is available online at http://super.chem.polimi.it/. The BzzMinimizationRobust is based on a modified Nelder-Mead Simplex direct search method (OPTNOV variant), which has proved to handle problems with highly nonlinear, nondifferentiable and discontinuous functions, and problems with narrow valleys (Buzzi-Ferraris 1967; Buzzi-Ferraris and Manenti 2010a, b).

Nonlinear constraints of the NLP subproblems are handled with a penalty function proportional to the constraint violation that is added to the objective function, allowing to start with an infeasible first guess. A Sequential Unconstrained

Minimization Technique (Correia et al. 2010) is adopted by progressively increasing the penalty weight (w) in order to limit difficulties with narrow valleys.

The *n*th subproblem, corresponding to the yn integer variables, is reported in Eq. (23.6), where the first set of constraints derives from the convex-hull reformulation of the disjunctions (Grossmann and Trespalacios 2013). These constraints force the decision variables of the non-selected units to be zero.

$$\begin{aligned} \min_{\mathbf{x}} \quad & Z_n^{UB}(\mathbf{y}_n) = \sum_{k \in K} TAC_k(\mathbf{y}_n) + f(\mathbf{x}, \mathbf{x_I}) + \sum_{i \in G \cup G_k} w_{g_i} \max(0; g_i(\mathbf{x}, \mathbf{x_I})) \\ \text{s.t.} \quad & y_k \cdot x_k^L \le x_k \le y_k \cdot x_k^U \quad k \in K \\ & \mathbf{x} \in R^n \end{aligned} \tag{23.6}$$

23.2.4.2 MILP Master Problem

The MILP master problem, reported in Eq. (23.7), is a linearization of the original nonlinear problem, which includes accumulated linearizations of the objective function and of the constraints, logic constraints representing the interconnectivity among process units within the superstructure, integer cuts, and convex-hull constraints derived from the reformulation of the disjunctions.

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \quad & Z_{OA}^{LB} = \alpha + w \cdot \left[\sum_{l \in L} \left(slack_f^l + slack_g^l \right) \right] \\ s.t. \quad & f_{obj}(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l) + \sum_{i \in iX} \tfrac{\partial f_{obj}}{\partial x_i}(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l)(x_i - x_i^l) + \sum_{t \in CT} \gamma_t y_t \le slack_f^l + \alpha \quad l \in L \\ & g_n(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l) + \sum_{i \in iX} \tfrac{\partial g_n}{\partial x_i}(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l)(x_i - x_i^l) + \sum_{t \in CT} \eta_t y_t \le slack_{g_i}^l \quad n \in G,\ l \in L \\ & \mathrm{y}_k \cdot g_n(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l) + \sum_{i \in iX} \tfrac{\partial g_n}{\partial x_i}(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l)(x_i - x_i^l) + \sum_{t \in CT} \eta_t y_t \le slack_{g_i}^l \ n \in G_k,\ k \in K,\ l \in L \\ & \mathbf{A}\mathbf{y} \le \mathbf{b} \\ & \sum_{i \in B} y_i - \sum_{i \in N} y_i \le |B| - 1 \\ & y_k \cdot x_k^L \le x_k \le y_k \cdot x_k^U \quad k \in K \\ & \mathbf{x}^\mathbf{L} \le \mathbf{x} \le \mathbf{x}^\mathbf{U} \\ & \mathbf{x} \in R^n, \mathbf{y} \in \{0;1\}^m,\ \alpha \in R,\ slack_f^l \in R^+,\ slack_g^l \in R^+ \quad l \in L \end{aligned} \tag{23.7}$$

Accumulated linearizations are obtained using as linearization points $(\mathbf{x}^l, \mathbf{y}^l, \mathbf{x}_\mathbf{I}^l)$ those that are obtained from all the previous NLP subproblems, derivatives with respect to the continuous decision variables, and delta contributions (γ, η) associated to each conditional tray. It should be noted that the set L defines all the previous main iterations, including the ones of the set covering problem. Moreover, slack variables are added to the linearization cuts to handle feasible region and objective function nonconvexities. They allow to find possible better solutions in the surrounding of the linearized nonconvex feasible region (Viswanathan and Grossmann 1990). It should be noted that the integer constraints in the form Ay ≤ b, are the translation of the propositional logic in Eq. (23.5) that defines the superstructure topology in the GDP representation (Grossmann and Trespalacios 2013).

Derivatives of objective function and constraints are found to be very important for the effectiveness of the master problem. In fact, an inaccurate linearization of the objective function can cut-off optimal solutions, while inaccurate linearizations of constraints can cut-off feasible region areas (potentially excluding the optimal solution). Using a process simulator does not allow one to directly access analytical derivatives. For this reason, the finite difference based on perturbations has been adopted. From an analytical point of view, perturbations should be as small as possible to ensure good derivative estimates. However, it is crucial to mention that decreasing the value of the perturbation results in increasing the error due to the noise of the NLP solvers embedded within process simulators. As a result, there is a trade-off between accuracy and noise for small values of the perturbation, and depending on the decision variable and on the linearization point, there is an optimal range of perturbation. For this reason, a perturbation test is performed for each continuous decision variable at each MILP main iteration. The derivative is estimated starting from a larger perturbation; then the perturbation is halved and this procedure is iterated until the relative change of the derivative is below a certain tolerance. Typical values of the perturbations are in the range of 10.3–10.4.

Delta contributions of a conditional tray for the objective function (γ) and for nonlinear constraints (η) are computed by converging a flowsheet that differs from the *i*th linearization point only for the number of trays of the distillation section to which the conditional tray belongs, as outlined in Fig. 23.8.

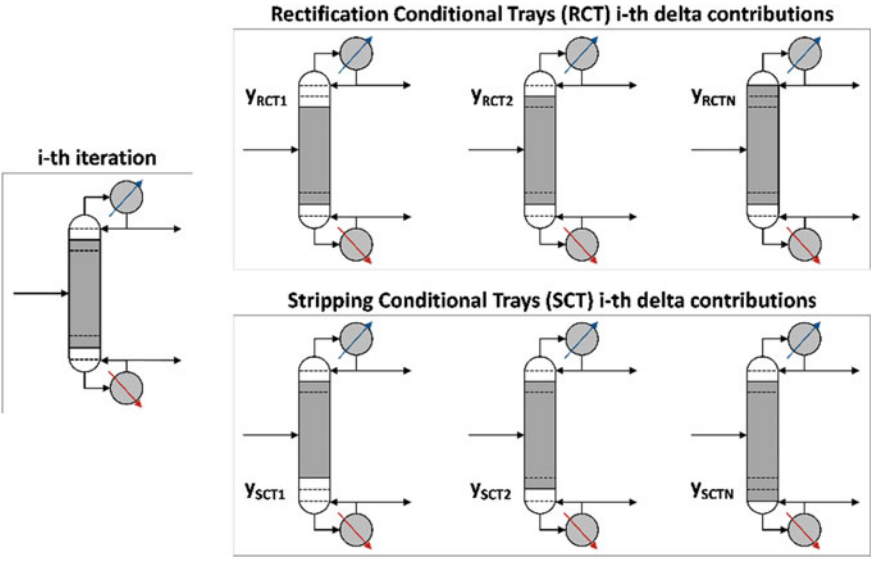

Fig. 23.8 Procedure to evaluate delta contributions of rectification and stripping conditional trays (Caballero et al. 2005)

23.2.5 Solution Strategy and Remarks

The success of this optimization algorithm is strongly influenced by a careful formulation of the problem and by the selection of proper parameters and settings.

The problem should be formulated selecting continuous and discrete decision variables that lead to a fast and easy convergence of the process simulator flowsheet with all the values in the range between lower and upper bounds. This is true especially for distillation columns, for which two specifications should be provided, along with the number of trays and feed tray. The two continuous decision variables are usually imposed as reflux ratio and bottom to feed ratio, or, in case of distillation sequencing, they can be the light key component recovery at the overhead and the heavy key component recovery at the bottom. It is sometimes useful to insert a few fixed trays in the rectification and in the stripping sections. This means that the permanent stages are not only the feed tray, the condenser and the reboiler, but also a suitable small number of trays above and below the feed, based on the separation requirements that can be checked at the end of the optimization procedure. This strategy helps the convergence of the column and avoids the solution of columns with a number of trays that is too small (such as only 3 trays), thus reducing the number of evaluation for the conditional trays delta contributions and speeding up the solution of the MILP master problem.

On the other hand, there are two main sets of parameters to consider: the ones related to the optimization algorithm, and the ones related to the SimSci PRO/II process simulator settings. Optimization parameters include perturbation step size and penalty weights for the NLP and MILP problems. Perturbation step size is selected based on the aforementioned adaptive step size strategy. Penalty weights for the violation of nonlinear constraints can be selected with the same order of magnitude of the objective function and can be successively increased with a SUMT procedure (Correia et al. 2010), while penalty weights for slack variables are found not to have a significant impact on the solution of the optimization problem. SimSci PRO/II simulation settings, in turn, should be carefully tuned, depending on the distillation column that is to be solved. At first, a suitable distillation algorithm should be selected. In this work, the CHEMDIST algorithm (Bondy 1991) has been selected for the VLLE heteroextractive distillation column case study, while the INPUT-OUTPUT algorithm (Russell 1983) for the other conventional columns. To ensure the convergence of the columns, three key settings are mandatory, i.e. initial estimates model, damping factor, and homotopy. The initialization model provides temperature estimates and either vapor or liquid molar flow rate estimates to initiate the iterative calculations. The selection of a suitable method equally depends on the distillation system. The CHEMICAL initialization model is adopted for complex thermodynamic systems and it is based on a multi-flash technique to bring the profiles closer to the final solution before the column algorithm takes over. Another option is to use the CONVENTIONAL initialization model that it is based on the Fenske shortcut model; it is less CPU intensive and it is recommended when nonidealities are weaker. On the other side, the damping factor should be reduced

from 1 to a value below 0.5 for oscillating systems, while homotopy could be adopted if the previous strategies fail.

Typical problem size should not exceed 10 continuous decision variables in the NLP subproblems, due to the intrinsic limit of DFO solvers (Rios and Sahinidis 2013), but can reach several hundreds of discrete decision variables in the MILP master problems. The overall number of continuous decision variables in the MILP could be much greater than 10, because in the NLP subproblems only selected conditional units are accounted for, along with their decision variables. The ratio between computing times for the NLP and MILP problems is proportional to the ratio between continuous and discrete decision variables, and the MILP computing time is largely due to the evaluation of conditional trays delta contributions for objective function and constraints.

23.3 Application Examples

The modeling and optimization framework is applied to four different case studies, dealing with the purification of aqueous solutions of oxygenated chemicals derived from biorefining. These flowsheets essentially include distillation columns with a highly nonideal liquid phase, considered with the UNIQUAC liquid activity coefficient model. The first case study introduces the methodology to the optimization of the operating parameters of a two-liquid phase heteroextractive distillation column with a heterogeneous azeotrope. This application example involves only the solution of a NLP problem with the robust derivative free optimizer (BzzMath) and demonstrates the higher performance with respect to the commercial software optimizers. From the second example, integer decision variables are introduced (MINLP) to optimize a single distillation column with one homogeneous azeotropes in terms of both operating parameters and tray structure. From the third case study, superstructures are introduced to outline how to optimize both structural (topology) and process operating parameters. The third case study includes the dewatering section of a downstream process with thermally coupled multieffect distillation columns, while the last example addresses the distillation sequencing of a quaternary complex bio-mixture.

For each application example, scope, thermodynamic modeling, problem formulation and results are provided, along with remarks and comments.

23.3.1 Heteroextractive Distillation of Water/Acetic Acid

The separation of water and acetic acid mixtures is a challenging task of industrial relevance (e.g. in the terephthalic acid and cellulose acetate production processes), due to the presence of a pinch point in the VLE x-y curve pure water end.

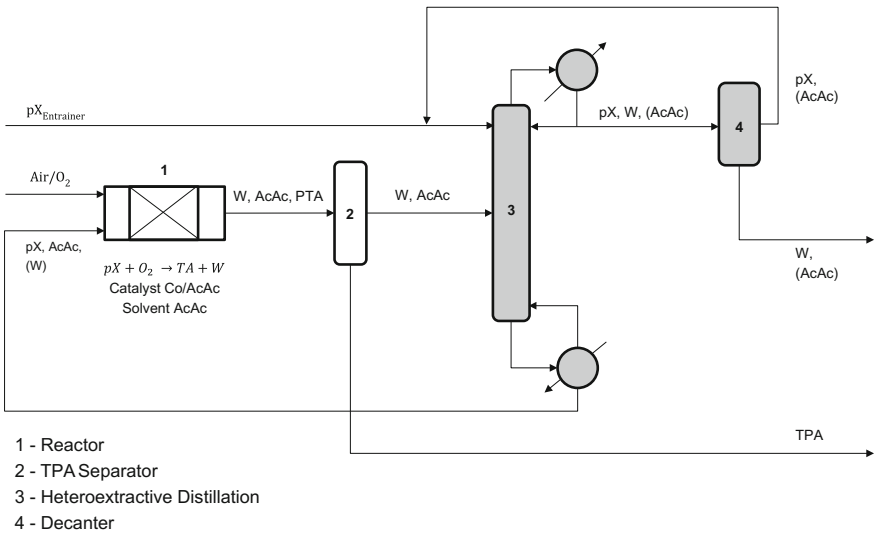

Fig. 23.9 Simplified terephthalic acid (TPA) production flowsheet

In order to overcome this technical issue, previous experimental and modeling studies (Pirola et al. 2014) showed that p-xylene (pX) is a suitable entrainer to operate an heterogeneous extractive distillation. Additionally, with respect to other entrainers suggested in the literature (Chien et al. 2004), pX is the main raw material of the terephthalic acid (TPA) production process and it is already available in this chemical plant. In addition, any contamination of AcAc with pX does not imply a problem for the recycle of the solvent (AcAc) to the terephthalic acid synthesis rector. As a result, pX can be adopted as entrainer in the heterogeneous extractive column to separate W and AcAc, as shown in Fig. 23.9. The miscibility gap between W and pX is exploited to separate the distillate stream in a decanter where an aqueous and an organic phase are obtained. The organic phase is composed of nearly pure pX that is recycled back to the top section of the column, while the aqueous stream with traces of AcAc is withdrawn and sent to wastewater treatment. From the bottom of the heteroextractive column, a stream composed by AcAc and the excess of the extractive entrainer is conveniently recycled back to the catalytic reactor where pX is converted to TPA and W with oxygen (eventually from air or enriched air), using a Co/Mn based catalyst and AcAc as solvent.

Reliable thermodynamic models, based on the UNIQUAC liquid activity model, have been developed by nonlinear regression on experimental data (Pirola et al. 2014) and are considered to simulate the purification section (Fig. 23.10) (Table 23.1).

where binary interaction parameters are in the form of Eq. (23.8).

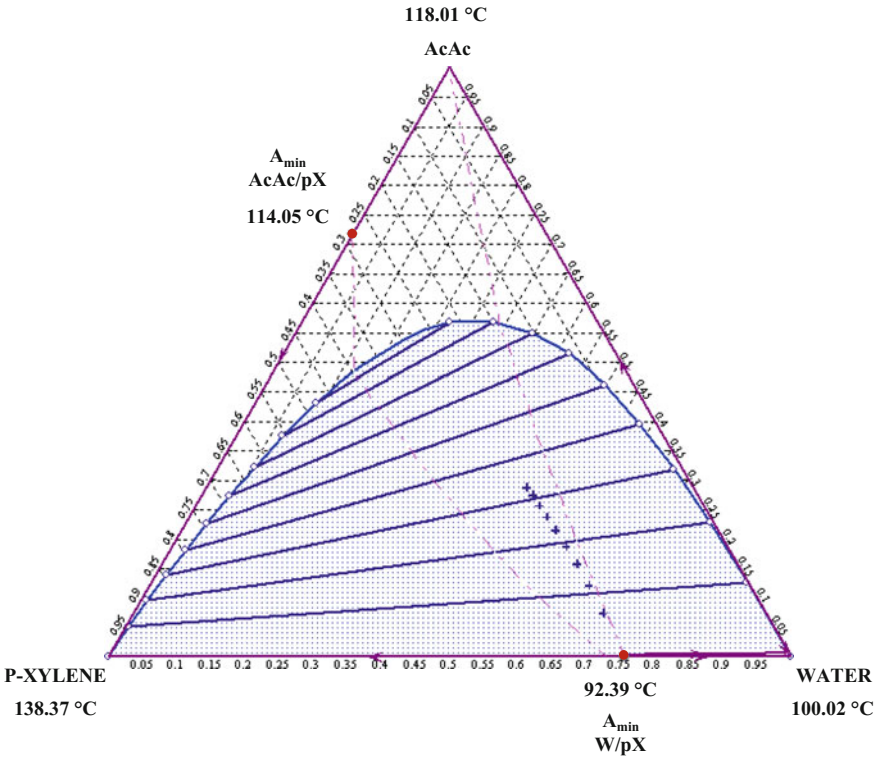

Fig. 23.10 W/AcAc/pX VLLE. Model predictions where binary parameters have been estimated by means of BzzMath nonlinear regression tools

Table 23.1 VLLE UNIQUAC binary interaction parameters for (1) W/ (2) pX/ (3) AcAc

ij	12	13	23	
a_{ij}	1.824E + 03	8.442E + 02	2.793E + 03	(J/mol)
a_{ji}	3.866E + 04	−1.059E + 03	−6.139E + 02	(J/mol)

$$\tau_{ji} = \exp\left(\frac{-(u_{ij} - u_{ji})}{RT}\right) = \exp\left(-\frac{a_{ij}}{RT}\right) \tag{23.8}$$

The heteroextractive distillation column is simulated with the commercial software SimSci PRO/II, by selecting the CHEMDIST VLLE algorithm, which is based on the full Newton-Raphson method and it is able to cope with both a single and a double liquid phase mixture, choosing in each case the appropriate set of interaction parameters.

The feed W/AcAc mixture and the entrainer (pX) are the inlet streams, while the bottom is composed by pX and AcAc and the overhead is sent to a decanter. The

organic phase (pX/AcAc) is separated from the aqueous one (W with traces of AcAc) inside the decanter. A total condenser and a kettle reboiler are assumed in the model. Feed composition is fixed to 20 % w/w of water and 80 % w/w of acetic acid, which is the typical composition of the effluent stream from the terephtalic acid catalytic reactor; its molar flowrate is set to 100 kmol/h. The plant is operated at atmospheric pressure (1 atm) and the temperature of the decanter is set to 30 °C. According to the industrial best practice, the purity specification limits the amount of AcAc in the aqueous stream to 5000 ppm, in order to minimize the loss of this compound, while in the bottom W composition must not exceed 3 % w/w.

Fixing the number of stages, six variables should be defined to solve the column simulation: the entrainer molar flowrate, the reflux ratio, the entrainer feed stage, the feed inlet stage, the temperatures of feed and entrainer. The definition of these degrees of freedom could be addressed in an optimization problem. The minimization of the objective function (TAC) is subject to MESH equations of the distillation column and the aforementioned purity specification constraints.

Two different optimizers are compared in the following: the one available within the SimSci PRO/II commercial software by Schneider Electric, and the BzzMath robust derivative free optimizer belonging to the homonymous C++ numerical library (Buzzi-Ferraris 2009).

Optimization results show that the Total Annual Cost of the column is deeply dependent from the cost of the entrainer make-up. For this reason the optimizer increases the number of stages in order to have a lower pX flowrate required to achieve the specifications.

The optimized column is reported in Fig. 23.11, where the operating and design parameters are highlighted.

A comparison with respect to the commercial simulator optimizer show that the BzzMath optimizer reaches a lower Total Annual Cost with respect to the commercial optimizer. The better performance of the BzzMath optimizer is highlighted in Fig. 23.12, where the final values of the economical objective function as a

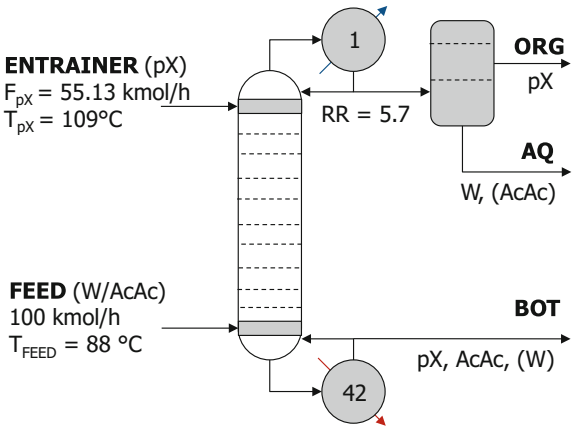

Fig. 23.11 Column optimal design and operating parameters for the separation of water and acetic acid mixtures

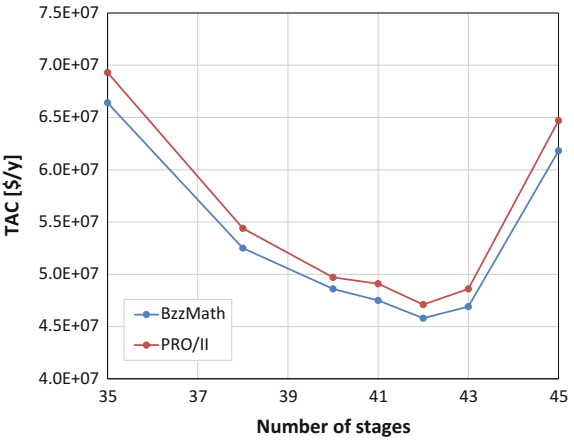

Fig. 23.12 BzzMath versus PRO/II objective function values as a function of the number of stages for the single column configuration

function of the number of stages is reported for both the BzzMath optimizer and the PRO/II one. An average economical saving of about 3.5 % is realized with the robust optimizer.

23.3.2 Purification of Bio-Based Glycols by Distillation

The second case study takes in consideration a single multicomponent distillation column that performs the cut between 1,2-propylene glycol (1, 2-PG) and ethylene glycol (EG), which are mixed with other co-products obtained from the hydroprocessing of lignocellulosic sugars. The feed is composed by 40 wt% EG, 50 wt% 1, 2-PG and 10 wt% of mixed heavier oxygenated chemicals that form also an homogeneous azeotrope with EG. The distillation column with structured packing internals is operated at atmospheric pressure with a total condenser and a kettle reboiler. The aim of the continuous-integer optimization is to determine the optimal number of trays, feed tray location, and the value of the two specifications (reflux ratio and bottom flowrate) that minimize TAC, with purity constraints at the top and at the bottom for 1, 2-PG and EG, respectively.

The first step for an effective optimization of bio-mixtures is the definition of a reliable thermodynamic model. Binary interaction parameters and predictive models included in process simulators for this kind of components are most of the times not reliable. For this reason, UNIQUAC binary interaction parameters have been estimated by nonlinear regression with the BzzNonlinearRegression class of the BzzMath library (Buzzi-Ferraris and Manenti 2010a, b; Pirola et al. 2014), using published phase equilibrium experimental data (Yang et al. 2014; Zhang et al. 2013; Zhong et al. 2014). Only two representative VLE Txy plots are reported in Fig. 23.13 for the sake of conciseness.

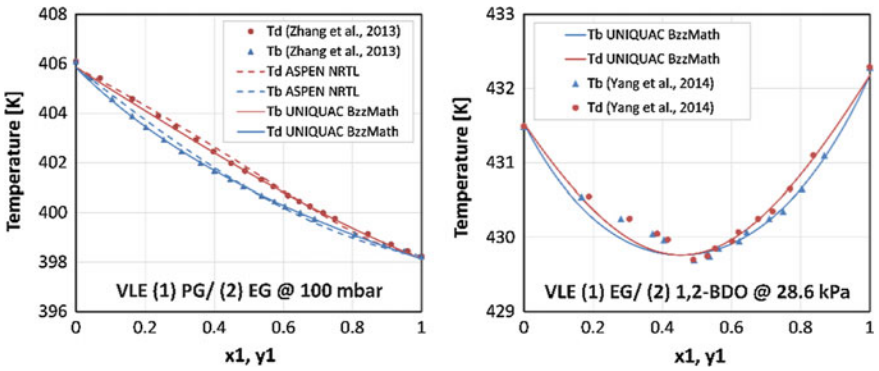

Fig. 23.13 Sample of Txy VLE binary plots of 1, 2-PG, EG and 1, 2-butanediol. Experimental data (symbols) and model predictions (lines)

The first guess for the decision variables has been derived by preliminary short-cut evaluations and involves a column with 123 stages (HETP = 0.250 m) with the feed at the 92nd stage. The evolution of the objective function over the main iterations is reported for both the MILP lower bound and the NLP upper bound in Fig. 23.14. The optimal configuration is found at the third iteration (red dashed line), with 135 stages and the feed at the 98th stage, because there is the simultaneous crossing between lower and upper bounds and the worsening of the NLP.

The problem has 2 continuous decision variables, 40 integer decision variables (due to the number of conditional trays, 20 for the rectification and 20 for the stripping sections), and 2 nonlinear constraints for product specifications. The starting value of the cost objective function is 3,986,700 $/y, while the optimal solution has a TAC of 1,803,370 $/y.

23.3.3 Dewatering of Bio-Based Polyols

As mentioned before, the dewatering processing of bio-mixtures is an energy intensive step that can significantly contribute to the total costs of the downstream operation. Thermal integration, and more specifically multieffect distillation, can substantially reduce heat duties, thus decreasing total costs.

The dewatering task of a diluted aqueous stream (85 wt% of water) of mixed bio-based glycols is addressed here. Two different process technologies are assessed (Fig. 23.15). Namely, simple distillation at atmospheric pressure (column LP1) and multieffect distillation with two heat integrated distillation columns operating at different pressures (columns LP2 and HP2). The optimization targets are to select the proper technology, and to optimize the selected columns in terms of both structural and operating parameters. Continuous decision variables include reflux ratio and B/F of the selected columns, split ratio between HP2 and LP2

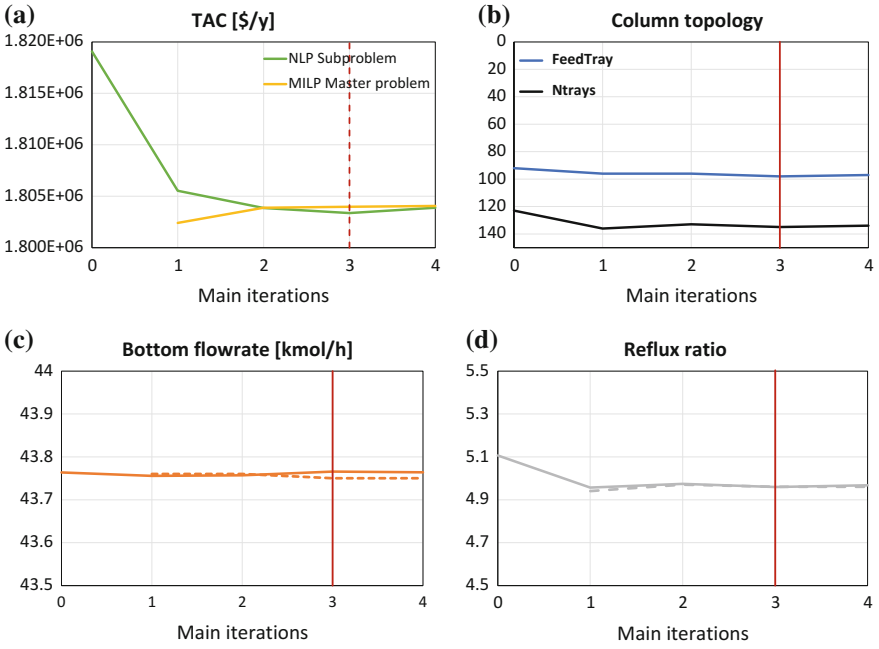

Fig. 23.14 Evolution of objective function (**a**), structural decision variables (**b**) and continuous decision variables (**c** and **d**) of case study 1 as a function of the OA main iterations. The *dashed line* reports the first guess given by the MILP master problem. The *vertical dashed line* is the optimum

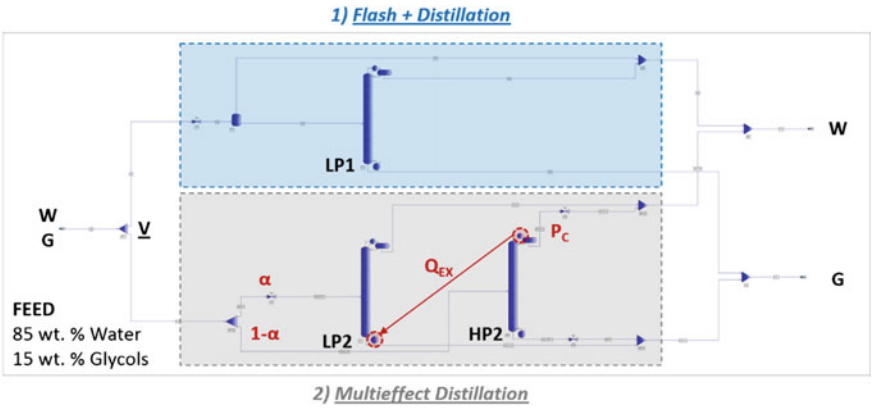

Fig. 23.15 Superstructure of the dewatering task of polyols mixtures

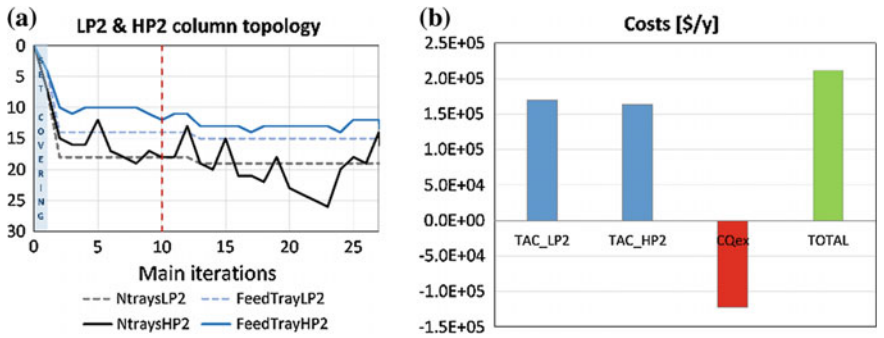

Fig. 23.16 Case study 3, dewatering section. **a** Column topology as a function of the OA main iterations. **b** Cost distribution of the optimal solution

columns (for disjunction 2), and condenser pressure of the HP2 column (for disjunction 2). Integer decision variables include the number of trays and feed tray of the selected columns, and the decision for the selection of the technology 1 (atmospheric distillation) or technology 2 (multieffect distillation). A glycols recovery specification and a water purity specification are considered as implicit global constraints, while a conditional constraint related to the feasibility for the heat integration between HP2 and LP2 columns applies only to the second disjunction. This last constraint forces the temperature of the HP2 condenser to be greater than the temperature of the LP2 reboiler plus the minimum approach temperature (EMAT = 5 °C).

The results summarized in Fig. 23.16 show that the multieffect distillation is the most favorable technology. Even in the initial (non optimized) set covering problem, the total cost associated with the single atmospheric distillation is ≈5 % higher than the one associated with the heat integrated double configuration, while the optimized solution further reduces costs to about 30 %. Interestingly, the two heat integrated distillation columns have similar topology and the same number of optimal stages. The ratio between rectification to stripping sections is about 3:1 for the LP2 column and 2:1 for HP2 column, with small reflux ratios. The split ratio between LP2 and HP2 is close to 50 %, while the HP2 condenser pressure is in the order of 10 bar. The heat integration between HP2 condenser and LP2 reboiler is greater than the 99.9 % of the condenser duty, and it allows to save steam utility costs (CQ_{EX}) of 122,289 $/y, while TAC of LP2 and HP2 are 169,813 and 163,997 $/y, respectively.

The number of continuous and integer decision variables is 8 and 62 respectively, and there are 3 nonlinear constraints. The objective function of the best first guess among the two set covering configurations is 300,492 $/y and the TAC of the optimal solution is 211,521 $/y.

23.3.4 Distillation Sequencing of Bio-Based Polyols

This subsection deals with the rigorous design of the distillation sequencing for a highly nonideal quaternary mixture of mixed bio-based glycols and butanediols (BDOs) with a homogeneous minimum boiling azeotrope between EG and 1, 2-BDO (Yang et al. 2014).

The feed mixture is composed of 50 wt% 1, 2-PG, 40 wt% EG, and 10 wt% of mixed butanediols (BDOs) at the bubble point. The design is achieved by considering a superstructure for the quaternary mixture (in which one pseudo-component, C, is the azeotrope) with 10 atmospheric distillation columns (Fig. 23.17), and optimizing both the selection of the columns and their conditional trays, as well as two continuous decision variables to specify the operation of each active column. In this case, the reflux ratio and the bottom total molar flowrate are selected as continuous decision variables to help the convergence of the columns by providing exact estimates for reflux ratio and product flowrates. Four implicit inequality constraints are considered for the purity specifications (mole fractions) and for the mass recoveries of the two main products (i.e. EG and 1, 2-PG).

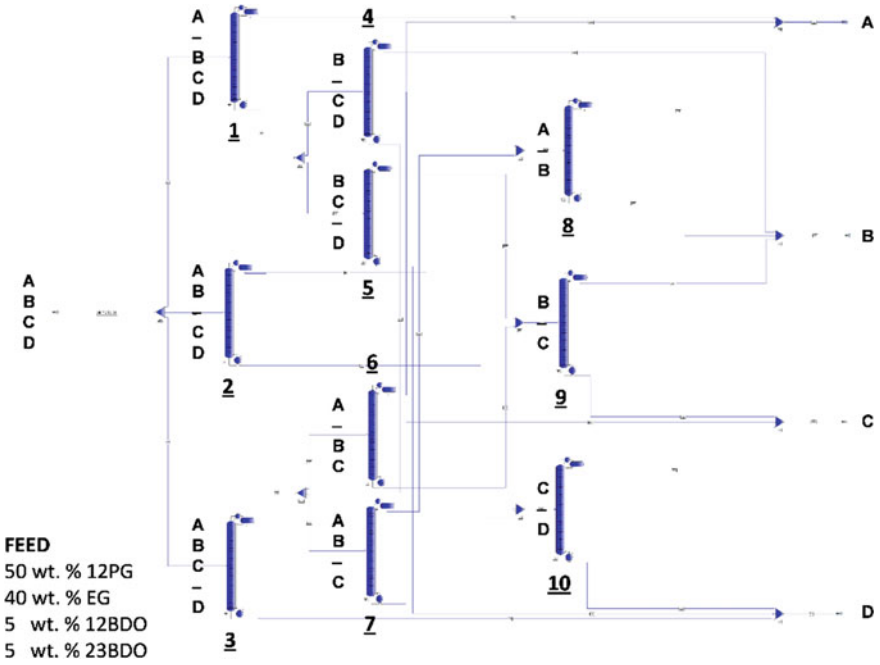

Fig. 23.17 Superstructure of the distillation sequencing for a quaternary mixture

The propositional logic defining the superstructure is reported in Eq. (23.9) for the sake of completeness, and results in the decision among 5 possible separation process layouts.

$$
\begin{aligned}
&Y_1 \dot{\vee} Y_2 \dot{\vee} Y_3 \\
&Y_4 \dot{\vee} Y_5 \\
&Y_6 \dot{\vee} Y_7 \\
&Y_8 \dot{\vee} Y_9 \dot{\vee} Y_{10} \\
&Y_1 \Rightarrow Y_4 \vee Y_5 \\
&Y_2 \Rightarrow Y_8 \wedge Y_{10} \\
&Y_3 \Rightarrow Y_6 \vee Y_7 \\
&Y_4 \Rightarrow Y_{10} \\
&Y_5 \Rightarrow Y_9 \\
&Y_6 \Rightarrow Y_9 \\
&Y_7 \Rightarrow Y_8
\end{aligned}
\tag{23.9}
$$

An additional cut, Eq. (23.10), is added to the constraints of the MILP master problems to enforce the selection of only three distillation column, which is the minimum number required for the separation of four components (NC − 1).

$$
\sum_{i \in COL} y_i = 3 \tag{23.10}
$$

The results are summarized in Fig. 23.18. The selected configuration is the sequence of columns 2/8/10 that first realizes the middle cut between 1, 2-PG and EG and then purify the two main products from butanediols (BDOs). Considering the set covering problem, keeping the same total number of trays, the configurations 1/4/10, 1/5/9, 3/6/9, 3/7/8 have 24, 96, 67 and 64 % higher extra total annualized costs. The second best configuration is the direct sequence (1/4/10), while the worst configuration is the 1/5/9 that requires higher reflux ratios to avoid violating the purity and recovery constraints. For the optimized columns topology, the cost distribution among the selected columns is 54 % for column 2, 29 % for column 8 and 18 % for column 10. The OPEX/CAPEX ratio is in the range of 0.90–1.25.

The problem has 20 and 210 continuous and integer decision variables respectively, with 4 nonlinear purity and recovery constraints.

The objective function of the best first guess among all set covering configurations is 3,249,840 $/y, wile the TAC of the optimal solution is 3,133,220 $/y.

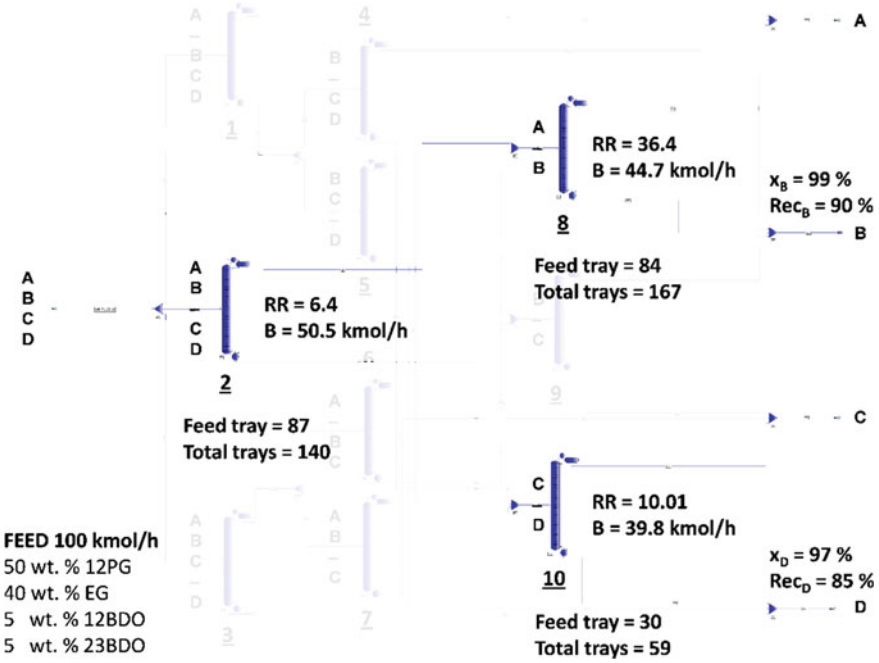

Fig. 23.18 Optimal distillation sequence and decision variables

References

Ahmetovic, E., Martin, M., & Grossmann, I. E. (2010). Optimization of energy and water consumption in corn-based ethanol plants. *Industrial and Engineering Chemistry Research, 49*, 7972–7982.

Barttfield, M., Aguirre, P. A., & Grossmann, I. E. (2004). A decomposition method for synthesizing complex column configurations using tray-by-tray GDP models. *Computers and Chemical Engineering, 28*, 2165–2188.

Biegler, L. T. (1985). Improved infeasible path optimization for sequential modular simulators. 1. The interface. *Computers & Chemical Engineering, 9*, 245–256.

Bondy, R. W. (1991). A new distillation algorithm for non-ideal system. In *AIChE Annual Meeting*.

Bravo-Bravo, C., Segovia-Hernandez, J. G., Gutierrez-Antonio, C., Duran, A. L., Bonilla-Petriciolet, A., & Briones-Ramirez, A. (2010). Extractive dividing wall column: Design and optimization. *Industrial and Engineering Chemistry Research, 49*, 3672–3688.

Brunet, R., Reyes-Labarta, J. A., Guillen-Gosalbez, G., Jimenez, L., & Boer, D. (2012). Combined simulation-optimization methodology for the design of environmental conscious absorption systems. *Computers and Chemical Engineering, 46*, 205–216.

Buzzi-Ferraris, G. (1967). Ottimizzazione di funzioni a più variabili. Nota I. Variabili non vincolate. Ing. Chim. It., 3, 101.

Buzzi-Ferraris, G., & Manenti, F. (2009). Kinetic models analysis. *Chemical Engineering Science, 64*, 1061–1074.

Buzzi-Ferraris, G., & Manenti, F. (2010a). A combination of parallel computing and object-oriented programming to improve optimizer robustness and efficiency. *Computer Aided Chemical Engineering*, *28*, 337–342.

Buzzi-Ferraris, G., & Manenti, F. (2010b). Interpolation and regression models for the chemical engineer: Solving numerical problems.

Buzzi-Ferraris, G., & Manenti, F. (2012). BzzMath: Library overview and recent advances in numerical methods. *Computer-Aided Chemical Engineering, 30*, 1312–1316.

Caballero, J. A., Milan-Yanez, D., & Grossmann, I. E. (2005). Rigorous design of distillation columns: Integration of disjunctive programming and process simulators. *Industrial and Engineering Chemistry Research, 44*, 6760–6775.

Chen, G. Q. (2009). A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. *Chemical Society Reviews, 38*, 2434–2446.

Chen, Y., Eslick, J. C., Grossmann, I. E., & Miller, D. C. (2015). Simultaneous process optimization and heat integration based on rigorous process simulations. *Computers and Chemical Engineering*.

Chien, I., Zeng, K. L., Chao, H. Y., & Hong Liu, J. (2004). Design and control of acetic acid dehydration system via heterogeneous azeotropic distillation. *Chemical Engineering Science, 59*, 4547–4567.

Corbetta, M. (2015). Green ethylene glycol. MINLP downstream process optimization and multi-scale simulation of gasification processes. Doctoral Dissertation, Politecnico di Milano.

Corbetta, M., Manenti, F., Pirola, C., Tsodikov, M. V., & Chistyakov, A. V. (2014). Aromatization of propane: Techno-economic analysis by multiscale "kinetics-to-process" simulation. *Computers and Chemical Engineering, 71*, 457–466.

Corbetta, M., Grossmann, I. E., & Manenti, F. (2016). Process simulator-based optimization of biorefinery downstream processes under the generalized disjunctive programming framework. *Computers and Chemical Engineering, 88*, 73–85.

Correia, A., Matias, J., Mestre, P., & Serôdio, C. (2010). Direct-search penalty/barrier methods. *Proceedings of The World Congress on Engineering, 2010*(3), 1729–1734.

Dias, M. O. S., Ensinas, A. V., Nebra, S. A., Maciel, R., Rossell, C. E. V., & Maciel, M. R. W. (2009). Production of bioethanol and other bio-based materials from sugarcane bagasse: Integration to conventional bioethanol production process. *Chemical Engineering Research and Design, 87*, 1206–1216.

Diaz, M. S., & Bandoni, J. A. (1996). A mixed integer optimization strategy for a large scale chemical plant in operation. *Computers and Chemical Engineering, 20*, 531–545.

Douglas, J. M. (1988). *Conceptual design of chemical processes*. New York.

Duran, M. A., & Grossmann, I. E. (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical programming, 36*(3), 307–339

Eslick, J. C., & Miller, D. C. (2011). A multi-objective analysis for the retrofit of a pulverized coal power plant with a CO_2 capture and compression process. *Computers and Chemical Engineering, 35*, 1488–1500.

Garcia, N., Fernandez-Torres, M. J., & Caballero, J. A. (2014). Simultaneous environmental and economic process synthesis of isobutane alkylation. *Journal of Cleaner Production, 81*, 270–280.

Gross, B., & Roosen, P. (1998). Total process optimization in chemical engineering with evolutionary algorithms. *Computers and Chemical Engineering, 22*, S229–S236.

Grossmann, I. E. (2002). Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and engineering, 3*(3), 227–252.

Grossmann, I. E., & Kravanja, Z. (1995). Mixed-integer nonlinear programming techniques for process systems engineering. *Computers & Chemical Engineering, 19*, S189–S204.

Grossmann, I. E., & Trespalacios, F. (2013). Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming. *AIChE Journal, 59*, 3276–3295.

Gutierrez-Antonio, C., & Briones-Ramirez, A. (2009). Pareto front of ideal Petlyuk sequences using a multiobjective genetic algorithm with constraints. *Computers and Chemical Engineering, 33*, 454–464.

Harsh, M. G., Saderne, P., & Biegler, L. T. (1989). A mixed integer flowsheet optimization strategy for process retrofits—the debottlenecking problem. *Computers and Chemical Engineering, 13*, 947–957.

Hermann, B. G., & Patel, M. (2007). Today's and tomorrow's bio-based bulk chemicals from white biotechnology—a techno-economic analysis. *Applied Biochemistry and Biotechnology, 136*, 361–388.

Kallrath, J., (2000). Mixed integer optimization in the chemical process industry experience, potential and future perspectives. *Chemical Engineering Research & Design, 78*(A6), 809–822.

Karuppiah, R., & Grossmann, I. E. (2006). Global optimization for the synthesis of integrated water systems in chemical processes. *Computers & Chemical Engineering, 30*(4), 650–673.

Karuppiah, R., Peschel, A., Grossmann, I. E., Martín, M., Martinson, W., & Zullo, L. (2008). Energy optimization for the design of corn-based ethanol plants. *AIChE Journal, 54*, 1499–1525.

Leboreiro, J., & Acevedo, J. (2004). Processes synthesis and design of distillation sequences using modular simulators: A genetic algorithm framework. *Computers and Chemical Engineering, 28*, 1223–1236.

Lee, S., & Grossmann, I. E. (2001). A global optimization algorithm for nonconvex generalized disjunctive programming and applications to process systems. *Computers & Chemical Engineering, 25*(11–12), 1675–1697.

Navarro-Amoros, M. A., Caballero, J. A., Ruiz-Femenia, R., & Grossmann, I. E. (2013). An alternative disjunctive optimization model for heat integration with variable temperatures. *Computers and Chemical Engineering, 56*, 12–26.

Navarro-Amoros, M. A., Ruiz-Femenia, R., & Caballero, J. A. (2014). Integration of modular process simulators under the generalized disjunctive programming framework for the structural flowsheet optimization. *Computers and Chemical Engineering, 67*, 13–25.

Nemhauser, G. L., & Wolsey, L. A. (1988). *Integer and combinatorial optimization*. New York: John Wiley & Sons, Inc.

Pirola, C., Galli, F., Manenti, F., Corbetta, M., & Bianchi, C. L. (2014). Simulation and related experimental validation of acetic acid/water distillation using p-xylene as entrainer. *Industrial and Engineering Chemistry Research, 53*, 18063–18070.

Rios, L. M., & Sahinidis, N. V. (2013). Derivative-free optimization: A review of algorithms and comparison of software implementations. *Journal of Global Optimization, 56*, 1247–1293.

Russell, R. A. (1983). A flexible and reliable method solves single-tower and crude-distillation-column problems. *Chemical Engineering, 90*, 53–59.

Sauer, M., Porro, D., Mattanovich, D., & Branduardi, P. (2008). Microbial production of organic acids: expanding the markets. *Trends in Biotechnology, 26*, 100–108.

Turkay, M., & Grossmann, I. E. (1996). Logic-based MINLP algorithms for the optimal synthesis of process networks. *Computers and Chemical Engineering, 20*, 959–978.

Vazquez-Castillo, J. A., Venegas-Sanchez, J. A., Segovia-Hernandez, J. G., Hernandez-Escoto, H., Hernandez, S., Gutierrez-Antonio, C., et al. (2009). Design and optimization, using genetic algorithms, of intensified distillation systems for a class of quaternary mixtures. *Computers and Chemical Engineering, 33*, 1841–1850.

Viswanathan, J., & Grossmann, I. E. (1990). A combined penalty-function and outer-approximation method for MINLP optimization. *Computers and Chemical Engineering, 14*, 769–782.

Xiu, Z. L., & Zeng, A. P. (2008). Present state and perspective of downstream processing of biologically produced 1, 3-propanediol and 2, 3-butanediol. *Applied Microbiology and Biotechnology, 78*, 917–926.

Yang, Z., Xia, S. Q., Shang, Q. Y., Yan, F. Y., & Ma, P. S. (2014). Isobaric vapor liquid equilibrium for the binary system (ethane-1, 2-diol + butane-1, 2-diol) at (20, 30, and 40) kPa. *Journal of Chemical and Engineering Data, 59*, 825–831.

Yeomans, H., & Grossmann, I. E. (1999). A systematic modeling framework of superstructure optimization in process synthesis. *Computers and Chemical Engineering, 23*, 709–731.

Zhang, L. H., Wu, W. H., Sun, Y. L., Li, L. Q., Jiang, B., Li, X. G., et al. (2013). Isobaric vapor-liquid equilibria for the binary mixtures composed of ethylene glycol, 1, 2-propylene glycol, 1, 2-butanediol, and 1, 3-butanediol at 10.00 kPa. *Journal of Chemical and Engineering Data, 58*, 1308–1315.

Zhong, Y., Wu, Y. Y., Zhu, J. W., Chen, K., Wu, B., & Ji, L. J. (2014). Thermodynamics in separation for the ternary system 1, 2-ethanediol + 1, 2-propanediol + 2, 3-butanediol. *Industrial and Engineering Chemistry Research, 53*, 12143–12148.

Chapter 24
Efficient Design of Biomass-Based Supply Chains: A Key Component of a Sustainable Energy System

J.M. Laínez Aguirre, M. Pérez-Fortes and L. Puigjaner

Abstract This chapter describes the use of mathematical programming as the tool for the design of biomass-based supply chains. This tool is helpful to devise the most appropriate manner of integrating conversion and pretreatment technologies with the channels required to convert the raw biomass, available in the collection areas, into energy in the demand points. The project analysis should be carried out adopting a holistic view. The formulation described in this chapter does so by tackling the problem from a multiple objective approach which considers financial, environmental as well as social aspects. The problem is formulated as a mixed integer linear program (MILP). The insights gained by using this approach are demonstrated through three literature case studies. The first case study comprises an illustrative hydrogen supply chain, where hydrogen is synthesised from biomass and coal gasification. The second one considers regional electrification in rural areas by using gasification combined with gas engines. In this case, a social criterion is introduced. The third case study is a biomass-based supply chain designed to partially fulfil the demand of processing coal plants existing in Spain.

Notation

Indices

a	mid point environmental impact categories
e	suppliers
f, f'	facility locations
g	end point environmental impact categories

J.M. Laínez Aguirre
University at Buffalo, 342 Bell Hall, Amherst, NY 14260, USA
e-mail: jmlainez@gmail.com

M. Pérez-Fortes
Institute for Energy and Transport Energy Technology Policy Outlook, Westerduinweg 3, 312/001, 1755 LE Petten, Netherlands
e-mail: Maria-Del-Mar.PEREZ-FORTES@ec.europa.eu

L. Puigjaner (✉)
Universitat Politècnica de Catalunya, Avda. Diagonal 647, 08028 Barcelona, Spain
e-mail: Luis.Puigjaner@upc.edu

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_24

i tasks
j equipment technology
s materials (states)
t, t' planning periods

Sets

A_g set of midpoint environmental interventions that are combined into endpoint damage factors g
FP set of materials s that are final products
$\bar{I}$ set of tasks i with variable input
I_j set of tasks i that can be performed in technology j
$\bar{J}_e$ technology j that is available at supplier e
$\tilde{J}_f$ technology j that can be installed at location f
J_i technologies that can perform task i
Mkt set of market locations
NTr set of production, or non-transport, tasks
RM set of materials s that are raw materials
Sup set of supplier locations
T_s set of tasks producing material s
$\bar{T}_s$ set of tasks consuming material s
Tr set of distribution tasks

Parameters

$FCFJ_{jft}$ [\$/h] fixed cost per unit of technology j capacity at location f in period t
i_r [adim.] discount rate
$NormF_g$ [adim.] normalising factor of damage category g
$Price_{sft}$ [\$/MJ] price of product s at market f in period t
$Water_s$ [adim.] Moisture for material s
$Water_{ij}^{max}$ [adim.] Maximum moisture for task i performed in equipment j

Greek symbols

α_{sij} [adim.] mass fraction of task i for production of material s in equipment j
$\bar{\alpha}_{sij}$ [adim.] mass fraction of task i for consumption of material s in equipment j
ζ_{ag} [adim.] g end-point damage characterisation factor for environmental intervention a
$\psi_{ijff'a}$ [points/kg] a environmental category impact CF for task i performed using technology j receiving materials from node f and delivering it at node f'
ψ_{ija}^{T} [points/(kg km)] a environmental category impact CF for the transportation of a mass unit of material over a length unit

Binary Variables

V_{jft} 1 if technology j is installed at location f in period t, 0 otherwise

Continuous Variables

$DamC_{gft}$ [point]	normalised endpoint damage g for location f in period t
$DamC_{g}^{SC}$ [point]	normalised endpoint damage g along the whole SC
$EPurch_{et}$ [$]	economic value of purchases executed in period t to supplier e
$ESales_{t}$ [$]	economic value of sales executed in period t
$FAsset_{t}$ [$]	investment on fixed assets in period t
$FCost_{t}$ [$]	fixed cost in period t
HV_{s} [MJ/kg]	lower heating value for material s
IC_{aft} [point]	midpoint a environmental impact associated to site f which rises from activities in period t
$Impact_{overall}^{2002}$ [point]	total environmental impact for the whole SC
NPV [$]	economic metric, net present value
$P_{ijff't}$ [kg]	specific activity of task i, by using technology j during period t, whose origin is location f and destination is location f'
$Profit_{t}$ [$]	profit achieved in period t
Pv_{sijft} [kg]	input/output material of material s for activity of task i with variable input/output, by using technology j during period t in location f (This must be a production activity)
$Sales_{sff't}$ [$]	amount of product s sold from location f in market f' in period t
S_{sft} [kg]	amount of stock of material s at location f in period t
SoC [unit]	surrogate social metric

24.1 Introduction

Late last century a survey on energy efficiency in industrial processes CEC (1988) warned a general trend that favours flexibility in the use of facilities dedicated to process manufacturing. This is an obvious requirement with the present fluctuations and the uncertain economic situation that characterizes today's market demand. This survey made reference essentially to the chemical process industries (CPI) as the most representative sector, although it also was of relevance to other flexible manufacturing industries which employ a network configuration. As a consequence, in the following years an intensive effort was dedicated to multi supplier operations, looking for strategies, management practices and techniques for improving the performance of supply and distribution chains. Increasing benefits made clear the importance of the subjacent energy supply chain. Additionally, economic, environmental and social factors questioned the dependence on a single source of energy based on fossil fuels which imposed energy prices subject to the increasing market uncertainty.

The panorama described above has resulted in a mounting pressure to explore alternative sources of energy that reduce environmental footprint at competitive cost. Nowadays, residual biomass is emerging as a preferred feedstock candidate Puigjaner et al. (2015). Bioenergy, or energy from biomass, from different sources (woody biomass, agricultural and land use biomass, industrial and municipal biodegradable

wastes) has an important role in the future low-carbon society to replace fossil fuels for the production of heat, electricity, transportation fuels, and to synthesize different types of chemicals. It is worth noting that while global energy demand is expected to grow by 37 % by 2040 Pérez-Fortes (2011), the European Union has established a target of 20 % share of renewable energy out of the total European energy consumption by 2020 Tchapda and Pisupati (2014). The 2030 Climate and Energy Policy Framework EU (2014) proposes the reduction of Greenhouse Gas (GHG) emissions to at least 40 % of the 1990 level by 2030 in order to meet the 2050 goal. Moreover, renewable energy source is one of the research priorities of the Strategic Energy Technologies (SET) Plan of the European Union Puigjaner et al. (2015); Commission (2015) as well as a research theme in the Integrated Roadmap of the SET Plan, whose aim is to consolidate the updated technology roadmap and to propose research and innovation actions EU (2014). In this context, bioenergy is not only relevant to the energy generation sector, but also in a number of other areas such as greenhouse gas control (as a potential zero and even negative emissions source), biofuels and waste disposal.

As a consequence, the energy sector is moving towards a new paradigm. More efficient conversion processes, renewable sources and smart grids are all encompassed by this new approach. It develops customized solutions, adapted to the particular needs and resources of each area. In this context, as an immediate and transition solution biomass can be properly co-used with fossil fuels, where technology is already mature; while 100 % biomass systems at small scale can be appropriate for residential uses and rural electrification in emerging countries. Notwithstanding, in the long term it is foreseen that there will not be a single technology or renewable source with massive implementation, but a combination of various conversion technologies to meet the energy demand. The alternatives to centralised and conventional sources of energy should be sustainable in the time, which implies a responsible resource exploitation, by balancing source availability with electricity demand, and therefore with the plant capacity Puigjaner et al. (2015).

Common challenges facing biomass to become a viable option are:

- Energy generation: Affordable biomass conversion technologies for fuel or energy production
- Sourcing: Conflict with agriculture (land and water use) and other uses of waste
- Link between technologies and site of use: Efficient supply-distribution network

Biomass can provide a larger energy share than the one that represents nowadays. For that to become a reality, technological, economic and social barriers need to be overcome. As a result, efforts are concentrated on developing integrated frameworks to support the associated decision-making process. This chapter explores the interactions between technology energy efficiency, consumption, environment, and social impact to help identify pathways toward a sustainable biomass-based energy system. In particular, we review holistic models developed to support decisions regarding the network configurations required to move towards approaches that address simultaneously the multiple dimensions of the de-carbonisation problem. This chapter is principally focused on gasification and combustion technologies.

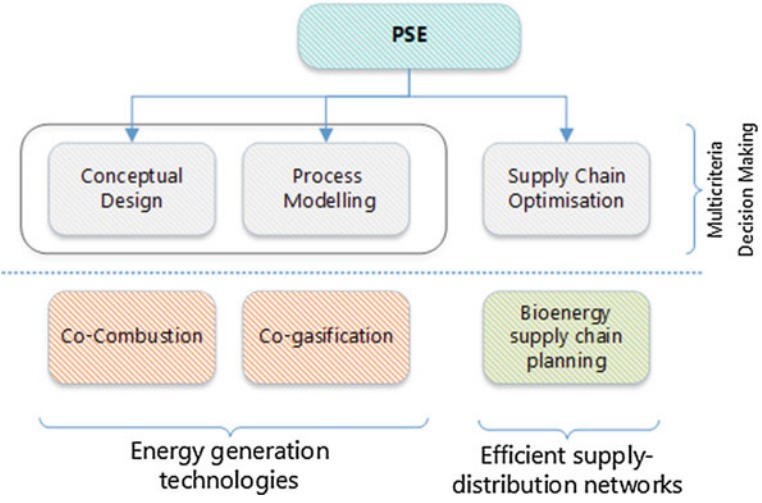

Fig. 24.1 PSE approach to bioenergy systems

As we shall see, a Process Systems Engineering (PSE) approach, depicted in Fig. 24.1, provides tools to address most of the aforementioned challenges. The first two blocks are related to developing process technologies for biomass exploitation; while the third one deals with the efficient supply-distribution networks needed to deploy such technologies. The focus of this chapter is on the latter block.

In the following, and after a brief literature review of the state of the art, this chapter focuses on the necessary integration of the entire supply/delivery chain for efficient and sustainable design of these emerging biomass energy systems, which includes the assessment of economic as well as social and environmental impacts. Moreover, emphasis is given to a rigorous approach that uses Process Systems Engineering, mathematical programming. First, a holistic framework is presented for modelling the energy and biomass, which contains a generic mathematical model. As motivating example is considered the supply chain for pure hydrogen generation from solid biomass. Then, biomass supply chains in developing economies are considered. The mathematical programming approach is applied to bio-based supply chains that use locally available biomass at or near the point of use in order to produce electricity or other bio-product. Here, a social impact metric is introduced. Next, biomass supply chains are examined in development economies. Here, the model considers the supply chain long-term strategic decisions, such as the selection of biomass sources, establishment of pre-treatment units and their location, and disposition of distribution centres, including an estimation of the potential of woody residues to supply the coal power plants that exist in the country. This model considers the possibility of collecting biomass/intermediates with different properties along the supply chain. To approach the computational expense of medium scale problems

is the focus of the following section, where the mathematical model is supporting the decision-making associated with the strategic and tactical design of biomass supply chains.

As conclusions and outlook of this chapter, further work is also underway to devise efficient strategies to decrease computational time and to add the additional constraints in the objective functions in order to tackle the mono-objective optimizations required in the multi-objective case.

24.2 State of the Art

Mathematical programming is a promising tool to assist in the quantitative evaluation of new approaches in the area of electricity generation. It is especially adequate for bio-based systems where sites of biomass generation may be far from consumption or demand points, biomass available locally may not match the biomass demand, and different generation/pre-treatment technologies may be available. Bio-based supply chains can be whether regional, i.e. concerning a community or small area with local needs to be fulfilled, or they can be global, i.e. when biomass to be supplied is required to satisfy the needs of a centralised energy system. In the medium and short term, the use of waste, which entails disposal problems, may be a continuous source of organic matter for power production Puigjaner et al. (2015).

Modelling and optimisation of SC's is becoming more popular, not only for biomass. Any type of industry or process can take benefit of this approach. Laínez and Puigjaner (2012) reviewed the application of SC optimisation in the chemical process industry. SC modelling derives from classical approaches that only consider operations, and goes a step forward by integrating business functionalities or market/operation dynamics. Supply chain decision-making tools, and tailor made approaches, will allow to appropriately exploiting the potential of biomass in power generation, heat and cooling applications, and as a transportation fuel. The optimisation of bio-based supply chains encompasses various decisions such as raw materials selection, facility location, selection and sizing of pre-treatments, products to be synthesised, and connectivity in the supply/delivery network, among others.

The biomass SC problem may be addressed using a wide range of decision-maker outlooks. As example, Caputo et al. (2005) evaluate the net present value (NPV) of 100 % biomass projects, focusing on transportation. Bowling (2011) look for an optimal SC for a biorefinery, considering overall sales and costs optimization to discern between a distributed or centralized structure with special attention on transportation costs. Ayoub et al. (2009) focus on costs and environmental impact through emissions to air, water pollutants and solid wastes. Damen and Faaij (2006) perform a life cycle inventory to compare co-combustion and combustion of only coal and Perry and Rosillo-Calle (2008) focus on CO_2 emissions along the whole SC. A more recent work from Mele et al. (2011) combines the use of mathematical programming with LCA, to perform a multi-objective optimization based on the NPV and the LCA, to produce bioethanol from sugar cane in Argentina. Environmental evaluations often

take into account a LCA Cherubini and Stromman (2011). Other attempts have been recently done to add the social criterion to the economic and environmental points of view, as the creation of places of job You et al. (2012). The bio-based works combine multi-objective optimization and mathematical programming (MILP, mixed integer non-linear program, MINLP, with and without uncertainty and risk consideration) or scenario-based optimization with geographic information systems (GIS) for spatial data analysis. The literature review from An et al. (2011) exposes that bioenergy is approaching to an important growth and needs to integrate strategic, tactical and operational decisions (i.e. the operations research point of view) to enhance and secure their viability, even if planning models have not been fully required (and therefore, developed) yet. Different works can be found that go in depth into a specific SC echelon: (i) feedstock production (growing, harvesting and collection), or waste generation, (ii) biomass pre-treatment, (iii) storage, (iv) biomass treatment, (v) electricity distribution and (vi) electricity consumption.

The work by Yue et al. (2014) reviews the major pathways for biomass to bioenergy and biofuel products. Biorefineries and carbon capture and storage are also included. The concept of superstructure is also exploited for the selection of the best technologies. The authors point out the challenges of including sustainability and uncertainties into the optimisation of the supply chain. The review by Cambero and Sowlati (2014) remarks that the use of biomass has an important potential to substitute fossil fuel, while all three aspects of sustainability (economic, environmental, and social) have to be considered in the optimisation problem. Uncertainty has been increasingly considered when modelling biomass supply chains: in Osmani and Zhang (2014) a stochastic mixed-integer linear program (MILP) model is developed which considers uncertainty in the supply of biomass-to-bioethanol, demand of biofuel, biomass and biofuel prices. The purpose of this model is to determine the location and the efficiency of the biorefineries, storage sites and selling points of bioethanol. Gebreslassie et al. (2012) also develop a stochastic MILP to address the optimal design of a biorefinery supply chain under supply and demand uncertainties. Miret et al. (2016) in their design of a bioethanol supply chain took into account the optimisation of a superstructure of first and second generation biomass conversion technologies.

High complex and computational demanding programs are resulting from the development of MILPs for the evaluation of biomass related supply chains, thus calling for decomposition methods that can attenuate the heavy computational load that is needed for the solution of stochastic programs. For example, Balaman et al. (2014) designed an anaerobic digestion supply chain, under cost and environmental criteria optimisation. Their model considers uncertainties by employing a Fuzzy multi-objective MILP. Osmani and Zhang (2014) employed a decomposition based on the Sample Average Approximation method. Gebreslassie et al. (2012) utilised the Multicut L-shaped method, while the work by Shastri et al. (2011) employed a decomposition scheme together with a distributed computing approach.

24.3 Energy/Biomass Supply Chain Modelling

This section describes the generic problem associated with the optimal design and operation of Biomass SC networks. In general, the SC strategic level determines the network through which the final product is manufactured/generated and distributed to finally reach the final consumer. The goal of a SC network design problem is to determine the optimal sourcing, manufacturing and distribution configuration for the different product lines of an enterprise. The most common approach is to formulate a large-scale Mixed Integer Linear Program (MILP) that captures the main revenue channels as well as the relevant fixed and variable operating costs for each facility and each major product Graves and Tomlin (2003). Specifically, a Biomass SC network consists of a number of potential geographical locations where either a conversion/pre-treatment site and/or distribution centre can be opened, and suppliers at fixed locations which have available biomass with different properties. The characteristics of the biomass can be changed by using the pretreatment units (e.g., drying or torrefaction) so that the treated biomass meets the characteristics required to be used in further steps. Even more, such pre-treatments increase the energy content and bulk densities of the biomass. Material flows between any facilities may appear if selecting such flow allows improving the performance of the SC. A market demand may be served by more than one site.

The mathematical model supports managers on planning decisions such as:

- The active SC nodes and links among them;
- The facilities capacity expansion in each time period;
- The product portfolio per plant, production amounts, utilization level, and transportation links to establish in the network alongside with material flows;
- The amount of final products to be sold in each market;
- The environmental impact associated to each SC node or activity.

A general schematic of the biomass energy SC is shown in Fig. 24.2. Notice that it is comprised by four blocks: (i) sourcing, (ii) pretreatment, (iii) generation, and (iv) distribution. The sourcing block consists in collecting the different biomass that may be available from different regions and suppliers. Each type of biomass has its own characterizing properties such as moisture content and heating value that determines its energy conversion efficiency. The pretreatment block considers those activities that modify the quality (primarily moisture content) and/or shape of the biomass. Examples of this kind of processes are the chipping, pelletising, drying, and torrefaction. These activities may be necessary, provided that there may be a technology in use which requires feeding material having a maximum moisture content and/or some shape requirements. The generation block converts the biomass into energy or any biofuel. Finally, the distribution block comprises those activities aiming at delivering the final product to the consumption points.

In a first approach, the SC decisions will be taken such that an economic indicator, i.e., Net Present Value (NPV), and an environmental impact metric, are optimized at the end of a predefined planning horizon.

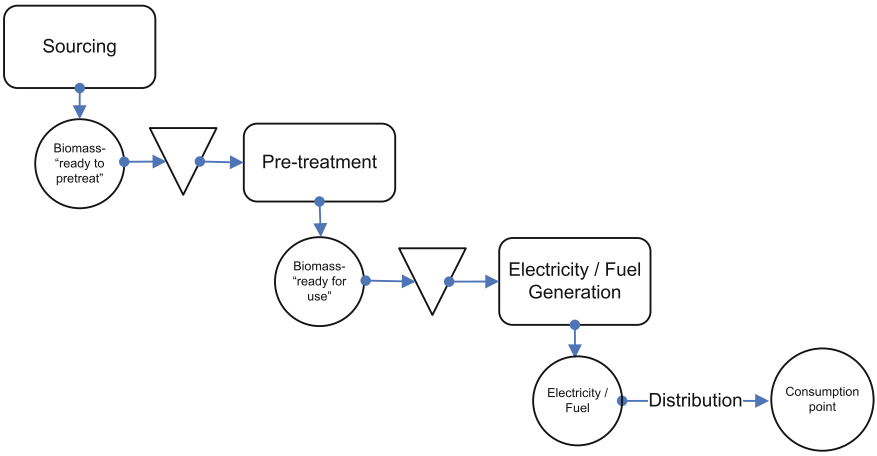

Fig. 24.2 General schematic of the biomass SC

The environmental metric selected for this model is the IMPACT 2002+ Humbert et al. (2005). This methodology proposes a feasible implementation of a combined mid-point/damage-oriented approach. It relates all types of Life Cycle Inventory results via 15 mid-point impacts (e.g., human toxicity, respiratory effects, iphoto-chemical oxidation, aquatic ecotoxicity, terrestrial ecotoxicity, global warming) to four end-point categories (human health, ecosystem quality, climate change-global warming potential and resources).

24.3.1 The Mathematical Formulation

The mathematical formulation of the biomass-based SC problem is briefly described next. This is a MILP formulation based on the works of Laínez-Aguirre et al. (2009) and Bojarski et al. (2009). In this chapter, the most relevant parts of the formulation are briefly explained. The interested reader is referred to the previous references for the complete formulation. The variables and constraints of the model can be roughly classified into three groups. The first one concerns process operation constraints. The second one deals with the environmental model, while the third refers to the economic formulation.

24.3.1.1 Operations Model

The design-planning model selected is adapted from the work of Laínez-Aguirre et al. (2009). This model translates the State-Task-Network (STN) formulation Kondili et al. (1993) to the SC context. This facilitates the consideration of

pretreatment activities and their outputs. The SC material balances can be modelled by means of a single equation set for all materials and echelons. This is possible since the formulation is centred around tasks i in contrast to the traditional product-based formulations. Thus, the most relevant variable of the model is $P_{ijff't}$, which represents the magnitude of a particular task i, performed using technology j during period t, whose origin is location f and destination is location f'. In the case of production activities, they must receive and deliver material within the same location (P_{ijfft}), while a distribution activity have different facilities f and f'. This mathematical formulation assumes that an activity consumes and produces certain materials with determined properties and can be performed in different equipments.

Mass balance is one of the main building blocks of the formulation and must be satisfied at each node of the network. The expression for the mass balance for each type of material s (raw material, pre-processed biomass, final product) processed at each potential site f in every time period t is presented in Eq. (24.1). Parameter α_{sij} is defined as the mass fraction of material s that is produced by task i using technology j. T_s is a set that refers to tasks that produce s, while $\bar{\alpha}_{sij}$ and $\bar{T}_s$ sets, are associated with tasks which consume s.

$$S_{sft} - S_{sft-1} = \sum_{f'} \sum_{i \in T_s} \sum_{j \in (J_i \cap \tilde{J}_{f'})} \alpha_{sij} P_{ijf'ft} - \sum_{f'} \sum_{i \in \bar{T}_s} \sum_{j \in (J_i \cap \tilde{J}_f)} \bar{\alpha}_{sij} P_{ijff't} \forall s, f, t \tag{24.1}$$

For biomass SC, it is particular important to include an *energy balance* equation. If we considered that biomass properties are fixed along the different process in the network, the energy balance is satisfied directly by the definition of the streams. However, we would like to relax this assumption. For that purpose, we defined the set of activities ($\bar{I}$) for which it is convenient to let the model specify the mixture of inputs required in order to achieve a given value of a specific biomass property; for instance, a specific moisture content. For such activities, the combination of feedstock and, therefore, the proportion of each feedstock is *variable*. In order to model this feature, the mass balance is modified as shown in Eq. (24.2). Note that Eq. (24.1) is a particular case of Eq. (24.2).

$$\begin{aligned} S_{sft} - S_{sft-1} = & \sum_{f'} \sum_{i \in T_s} \sum_{j \in (J_i \cap \tilde{J}_{f'})} \alpha_{sij} P_{ijf'ft} - \sum_{f'} \sum_{i \in \bar{T}_s} \sum_{j \in (J_i \cap \tilde{J}_f)} \bar{\alpha}_{sij} P_{ijff't} \\ & + \sum_{i \in (T_s \cap \bar{I})} \sum_{j \in (J_i \cap \tilde{J}_{f'})} Pv_{sijft} - \sum_{i \in (\bar{T}_s \cap \bar{I})} \sum_{j \in (J_i \cap \tilde{J}_{f'})} Pv_{sijft} \quad \forall s, f, t \end{aligned} \tag{24.2}$$

With regard to the variables $P_{ijff't}$ [kg] and Pv_{sijft}, the former is used in the mass balance (Eq. (24.2) coupled with the parameter α_{sij} or $\bar{\alpha}_{sij}$ which specify a fixed proportion of material produced or consumed for a task i. On the other hand, Pv_{sijft} is modelling flexible tasks which allow the proportion of the material produced or

consumed to vary so as to provide more degrees of freedom for biomass mixing to the model. Consequently, Pv_{sijft} is not multiplied by such parameters in Eq. (24.2).

The energy balance for "flexible" activities is represented by Eq. (24.3). Here, HV_s [MJ/kg] is the heating value of material s. Each type of biomass has a different heating value. A specific activity changes the heating value of the output stream if (i) it is a pre-treatment task that modifies explicitly the calorific value of the biomass, or (ii) it is a task whose main objective is the change of shape, but it is fed with a mixture of biomasses.

$$\sum_{s \in T_s} HV_s Pv_{sijft} = \sum_{s \in \bar{T}_s} HV_s Pv_{sijft} \qquad \forall i \in \bar{I}, j, f, t \tag{24.3}$$

In case the flexible activities must be fed by an input stream with a given moisture content (MC), constraint (24.4) is enforced. The parameters $Water_s$ and $Water_{ij}^{max}$ represent the MC for material s, and the maximum MC allowed for task i performed in equipment j, respectively.

$$\sum_{s \in S_i} Water_s Pv_{sijft} \leq Water_{ij}^{max} \sum_{s \in \bar{S}_i} Pv_{sijf't} \qquad \forall i \in \bar{I}, j, f, t \tag{24.4}$$

The previous equations are the most relevant for biomass SC models compared to a traditional supply chain. The complete operations formulation includes capacity and market related equations that are common to most strategic network configuration models which can be found in Laínez-Aguirre et al. (2009).

24.3.1.2 Environmental Formulation

The application of the life cycle assessment (LCA) methodology to the SC model allows the implementation of the environmental formulation, which uses the IMPACT 2002+ as metric. Here, environmental interventions for each activity i are translated into metrics related to impact as end-points or mid-points metrics by the usage of characterisation factors. Equation (24.5) calculates IC_{aft} which represents the mid-point environmental impact a associated with site f, as a consequence of carrying out activities in period t. In turn, $\psi_{ijff'a}$ is the a characterisation factor of the environmental category impact for task i performed using technology j, receiving materials from node f and delivering them at node f'.

$$IC_{aft} = \sum_{j \in \bar{J}_f} \sum_{i \in I_j} \sum_{f'} \psi_{ijff'a} P_{ijff't} \qquad \forall\, a, f, t \tag{24.5}$$

Equation (24.6) introduces $DamC_{gft}$ [points], which is a weighted sum of all mid-point environmental interventions. They are combined using g end-point damage factors ζ_{ag} [adim.], normalised with $NormF_g$ factors. Moreover, Eq. (24.7) calculates g normalised end-point damage along the SC ($DamC_g^{SC}$ [points]).

$$DamC_{gft} = \sum_{a \in A_g} NormF_g \zeta_{ag} IC_{aft} \qquad \forall\, g,f,t \tag{24.6}$$

$$DamC_g^{SC} = \sum_f \sum_t DamC_{gft} \qquad \forall\, g \tag{24.7}$$

Equation (24.8) aggregate the end-point environmental damages for the whole SC.

$$Impact_{overall}^{2002} = \sum_f \sum_g \sum_t DamC_{gft} \tag{24.8}$$

For further details regarding the environmental formulation the interested reader is referred to Bojarski et al. (2009).

24.3.1.3 Economic Formulation

The expressions required to compute the operating revenue, the operation costs, the total capital investment, and NPV are included in the economic formulation.

For instance, the *operating revenue* is expressed in Eq. (24.9) as the product sales during period t.

$$ESales_t = \sum_{s \in FP} \sum_{f \in Mkt} \sum_{f' \notin (Mkt \cup Sup)} Sales_{sf'ft} Price_{sft} \qquad \forall t \tag{24.9}$$

General speaking, the *operating costs* include fixed ($FCost_t$) and variable ($EPurch_{et}$) costs. The latter usually includes the cost of purchases from supplier, transport cost and production cost. Another important piece of the economic formulation is the *total capital investment* on fixed assets ($FAsset_t$), which must consider the investment made to expand the technology's capacity j in facility site f in period t.

Equation (24.10) represents the profit in period t, as operating revenues minus fixed and variable operating costs. The NPV can be calculated as in Eq. (24.11).

$$Profit_t = ESales_t - (FCost_t + \sum_e EPurch_{et}) \qquad \forall\, t \tag{24.10}$$

$$NPV = \sum_t \left(\frac{Profit_t - FAsset_t}{(1 + i_r)^t} \right) \tag{24.11}$$

The overall optimisation problem can be posed mathematically as follows:

$$\underset{\mathscr{X},\mathscr{Y}}{\text{Min}} \left\{ -NPV, Impact^{2002}_{overall}, -SoC \right\}$$
$$\text{subject to}$$
$$\textit{Operations, Environmental and Economic formulation};$$
$$\mathscr{X} \in \{0,1\}; \mathscr{Y} \in \mathbb{R}^{+}$$

where, $\mathscr{X}$ denotes the binary variables set, while $\mathscr{Y}$ corresponds to the continuous variable set.

24.3.2 A Hydrogen Supply Chain

This case study illustrates the basic concepts behind the biomass SC design-tactical model outlined in the previous section and was first presented in Laínez et al. (2011). It compares the generation of electricity and H2 from two different kinds of feedstock: (i) different biomass wastes and (ii) coal.

24.3.2.1 Case Study

A simplified potential network is proposed and restricted to Spain (see Fig. 24.3). Lugo (F1), Ciudad Real (F2) and Burgos (F3) are considered to be possible facilities location nodes. The feedstock is supposed to be available at Cordoba (LA), Lugo (LB), Cuenca (LC), Santander (LD) and Oviedo (LE). This last site is the one supplying coal. Hydrogen is supposed to be sold at three market places located at Madrid (M1), Valencia (M2) and Barcelona (M3), while electricity is fed to the Spanish electricity network at their respective generation places. Different biomasses are modelled considering that each of them possesses different energy content and humidity. Here, coal has been considered as a dry material which does not require any pretreatment.

The biomass may be pretreated before being finally processed. Figure 24.4 depicts the different pretreatment processes that may be applied to the biomass (BM) so that it achieves the adequate shape and properties (energy content and humidity) for later processing. In this case study bulk density has not been considered for the sake of simplicity. The pretreatment options considered here are: chipping, drying, torrefaction and pelletising. It is assumed that the condition for biomass to pass through torrefaction is to have a LHV lesser than 15 MJ/kg, while to be pelletised biomass must have a humidity equal to or lower than 7 %. The parameters associated with the pretreatment processes are listed in Table 24.1.

The technology that is employed to provide the final product is gasification and a gasification plant with Carbon Capture and Storage (CCS) for the H2 generation. Efficiencies of 40 % and 30 % are assumed for each plant respectively. Other

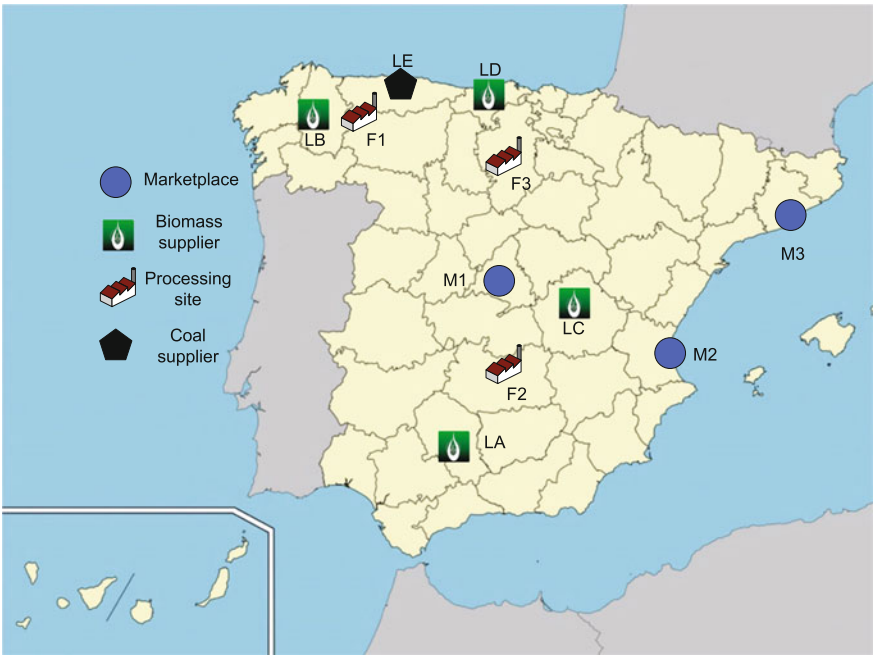

Fig. 24.3 Location map for the potential SC network

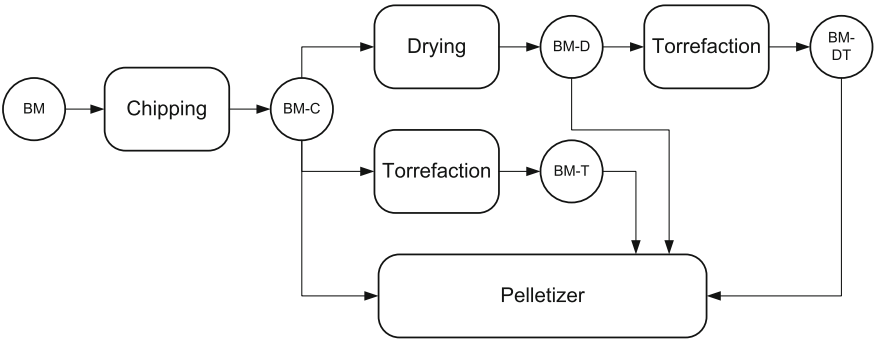

Fig. 24.4 STN representing the pretreatment activities for a generic biomass

relevant information concerning these technologies is presented in Table 24.2. In order to assess the environmental impact associated with the energy SC, the available LCI values were retrieved from the LCI database EcoinventV2.0 using SimaPro 7.1.6 SIMAPRO (2004) and converted directly to the IMPACT 2002+ mid-point indicators. For those activities which were not available, the impacts were assumed based on similar products or activities. The project is evaluated along a planning horizon

Table 24.1 Pretreatment processes and their main modelling assumptions

Activity/ equipment	Moisture losses (%)	Dry matter losses (%)	Operating cost (EUR/t)	Capacity (t/h)	Investment (1 × 106 EUR)	Electricity consumption (MWh/t)
Chipper	0	0.17	2.5	30	0.37	5
Dryer	88	0.08	55	100	5	20
Torrefactor	55	19	40	20	0.1	37
Pelletizer	0	0	3.5	6	0.485	30

Table 24.2 Parameters for the processes for electricity and H2 generation

Technology	Operating cost (EUR) (1 × 106EUR)	Capacity	Investment	Product price (EUR)	Total monthly demand
Electricity	34.2/MWh	300 MW	860	0.151/ kW-h	75000 MWh
H2	1880/t	33.6 t/h	1500	3/kg	650 t

of 25 years, considering monthly planning decisions. The model was implemented in GAMS which is algebraic modelling software.

24.3.2.2 Results

Figure 24.5 shows the obtained dominant biomass based SC that maximises NPV. It is found that the three potential locations are considered and on each one of them a facility is opened. All pretreatment technologies are installed in location F1 besides the required equipment to produce H2. From this site H2 is delivered to all markets. F1 is collecting all the forest wood residues (FWR) for which larger mass flows are required due to their low LHV. By establishing F1, which is near to the FWR collection site, significant savings in transportation are obtained. The electricity is generated in site F2. In this site; equipments to perform chipping, drying and pelletising are installed. The electricity demand of each market is satisfied from site F2. Site F3 is used just as a distribution centre for pre-treated biomass. Equipment for chipping and drying is installed in such a site. For this configuration there are some inter-site flows, clearly showing the capabilities of the model to tackle with inter-site distribution tasks. For instance, forest wood residues which have been dried and torrefied are being sent from site F1 to F2. By having materials flows of pre-treated biomass the transportation cost is reduced due to their higher bulk density.

The optimal configuration for the environmental impact has also been obtained. Figure 24.6 shows the minimum IMPACT 2002+ configuration for the biomass based SC. This supply chain fulfils with the same demand as the one obtained by optimising NPV. For this case the location F3 is not considered, and all biomass is

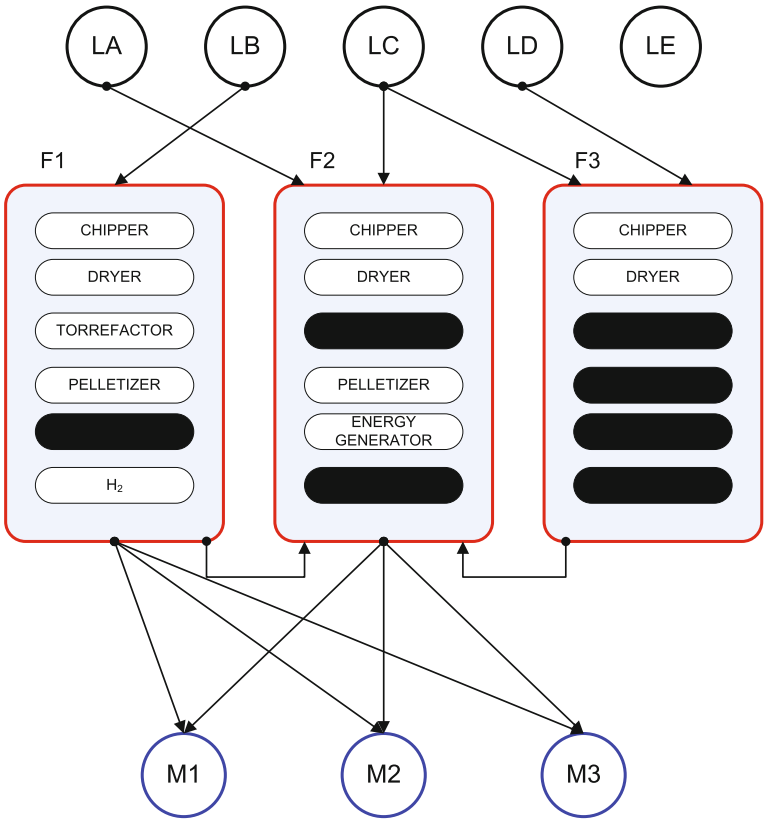

Fig. 24.5 Optimal NPV network configuration for the biomass based SC

sent from the collection sites to locations F1 and F2. This configuration is satisfying the demand of electricity from both locations F1 and F2, whereas H2 is delivered from site F2. This allows to slightly reduce the environmental impact associated with transportation. Recall that we introduce a "flexible" task to account for those tasks for which we would like the model to decide how to better mix different biomasses so as to achieve a given specified biomass property. We have assumed that the pelletizer is one of such tasks for this case study. To give an example, there are periods in which the model proposes to make the following mix: 1.4 % forest wood residues, 30.3 % dried and torrefied forest wood residues, 10.5 % dried pine waste, 14.4 % dried almond tree prunings, and 43.5 % chipped olive pomace (mass basis). This mixture is then fed to the syngas production plant. The values of humidity corresponding to these materials are 10.0 %, 6.0 %, 7.0 %, 7.0 % and 7.5 %, respectively. It can be proved that the humidity of this mix is 7.0 % which is the maximum humidity allowed for the pelletiser input.

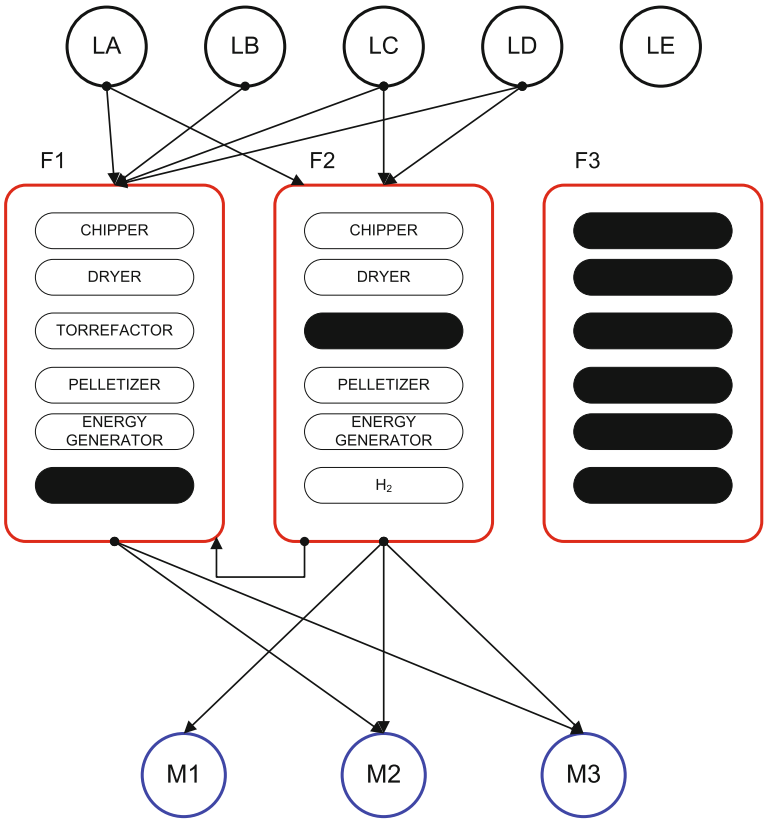

Fig. 24.6 Optimal IMPACT 2002+ network configuration for the biomass based SC

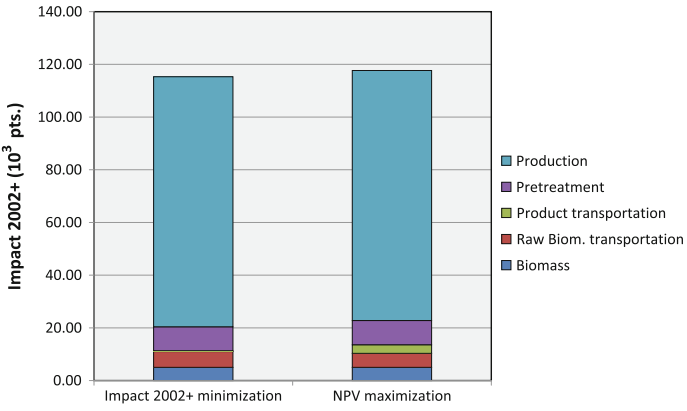

Fig. 24.7 Distribution of environmental impacts for single objective optimization solutions, according to different SC activities

Table 24.3 Environmental impacts arising from the optimisation results [Impact 2002+ pts]

End point impact category	Impact 2002+ optimisation	NPV optimisation	100 % coal-based SC
Human health	16255.29	17267.21	109640.5
Ecosystem quality	3375.79	3610.96	11077.16
Climate change	90383.37	90950.66	95334.48
Resources	5292.64	5852.73	140605.9
Impact 2002+	115307.1	117681.6	356658.1

By deploying the SC configuration corresponding to the more profitable SC configuration, a NPV equal to 228.51 M EUR is obtained. This value is reduced by 3 % when the environmental friendly configuration is established. The main difference between these two configurations is the investment required for installing the proposed capacity in the different sites. With regard to environmental interventions, electricity generation and H2 production are the most important factors contributing to the overall environmental impact in both single objective optimization cases; while biomass sourcing is the least impacting aspect (Fig. 24.7). This clearly shows that activities to reduce environmental impact should be focused on improving the technologies used to produce energy and H2.

For comparison purposes the optimal SC based on coal was also obtained. An NPV improvement of 219 % can be gained by utilizing coal as feedstock when compared to the Biomass based SC. The main difference is from the production cost which is due to the pretreatment activities that are required in the biomass based SC. This fact also makes the investment increase in the Biomass SC. However, the Impact 2002+ is increased in 203 % compared with the biomass based SC (Table 24.3). It is noteworthy that the impact associated with the climate change category is very similar for both cases. We have to bear in mind that CO2 is still emitted when using a Biomass SC, however this biomass is regenerated faster than fossil fuels. Nevertheless, the other categories are significantly increased in the coal based SC, specially for the resources and human health. This fact emphasizes the significance of having an overall impact indicator instead of a partial indicator such as CO2kg.

24.3.3 Biomass Supply Chains in Developing Economies

This second case study contemplates a specific rural area of a developing country, Ghana (Africa) which has been presented in Pérez-Fortes et al. (2012).

24.3.3.1 Case Study

Nine communities in Atebubu-Amantin district, in the Brong Ahafo Region (see Fig. 24.8) are part of the analysis. The selected communities form a region with a main characteristic: they are equipped with a multi-functional platform (MFP) that currently supplies the electricity needs; cell phone's battery charging, water refrigeration, lighting, radio, TV, computer and maize mills and cassava graters. Those communities are therefore used to pay for the electricity service. Data for the SC characterisation was provided by the Energy Center, in Kwame NKrumah University of Science and Technology (KNUST), and the NGO Kumasi Institute of Technology, Energy and Environment (KITE) (Ghana). The MFP's project in Atebubu district has been executed in collaboration with KITE and the local NGO called Women and Children Support Organisation (WACSO).

The electricity demand has been estimated on the basis of references from previous experiences on rural electrification projects, in West Africa and South American communities conducted by Arranz-Piera et al. (2011) and the company *Trama Tecnoambiental* Vallvé et al. (2007); Arranz-Piera (2008). The highest estimated gross demand is 448.65 kWh/day in Kumfia community, while the lowest is 21.17 kWh/day in Nwunwom community. Such figures take into consideration the LV microgrid losses. Figure 24.10 depicts the nine communities represented by their relative energy demands (in blue), in a square grid that represents their relative distance in km. Black points mark potential locations for pre-treatment and treatment units; all the communities and four more intermediate sites are considered as potential locations.

Cassava rhizome is the biomass waste considered in this study. Cassava is a well extented tropical crop in the country, mainly used as food (in the form of fufu or gari). Agriculture is the most important economic sector, with cassava, yam and maize as basic products. The cassava is planted once a year, in April, during the rainy season and does not need any special care. A 66.5 % of the tubercle is cassava rhizome Pattiya (2011). As the produced wastes have no current alternative use, the cost of acquisition is considered negligible. Table 24.4 summarises the cassava main

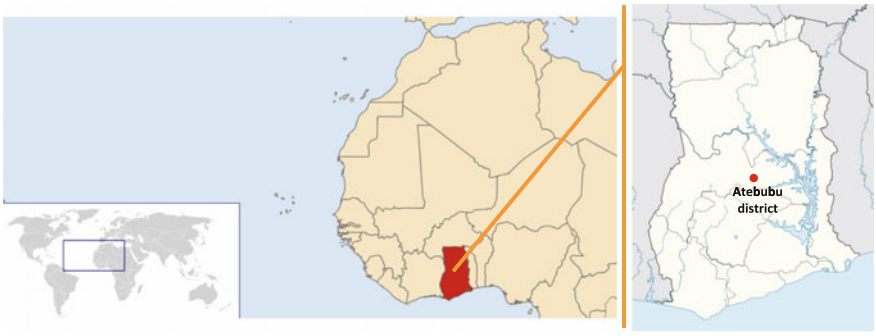

Fig. 24.8 Location of Atebubu district, in Ghana

Table 24.4 Feedstock properties

Biomass	Cost ($/t)	LHV_{ar} (MJ/kg)	MC (% wt)	Seasonality	Yearly available (t)
Cassava waste	0	10.61	42.50	June–October	1666.13

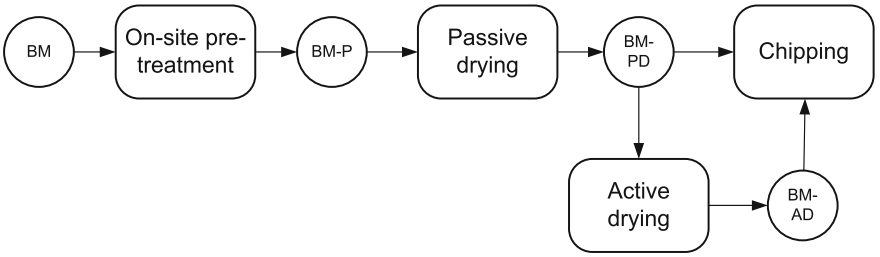

Fig. 24.9 Pre-treatment activities layout

properties. The last column shows the total amount of cassava waste produced by the 9 communities. This value has been calculated by considering that from the total amount produced in Atebubu district, only a 20 % can be taken for electricity purposes. According to Serpagli et al. (2010), 264649 t of cassava has been produced during 2009 in the whole Atebubu district, representing 17667.6 GJ/yr.

Inlet biomass can be fed to a gasifier, however it must be chipped and have 20 % or less of MC. It is supposed that the only possible biomass storage is carried out before chipping and gasification using on-field storage which is the cheapest and simplest option. Figure 24.9 shows the layout of the different pre-treatment options applied to the biomass (BM). MC, DM, shape and LHV change along the network.

The most important parameters in the biomass gasifier are the amount of inlet air (i.e., the equivalence ratio, ER) and the MC. See the main system's parameters in Table 24.5. The gasification units range between 5 and 100 kW_e. Table 24.6 lists the parameters required for cost estimation of the pre-treatment units, the G-ICE plant and transportation. The diesel price is assumed as $1133.31/t.[1]

With regard to the environmental formulation, Impact 2002+ metric is used to evaluate the environmental impact in points (pts). LCI values are retrieved from LCI database Ecoinvent-V1.3 (2006) using SIMAPRO (2004), and they are directly converted into Impact 2002+ mid-point indicators in the so-called LCIA step. The impact of the gasification has been adapted from a large scale gasification plant impact, an IGCC, considering the efficiency difference in energy terms between the two plants.

One important aspect considered in this case study is the *social impact*. The approach proposed to quantify this criterion consists in the number of demand sites that have a treatment or pre-treatment system installed. The aim is to install as many as possible to promote working places in the widest range of communities or demand

[1] On-field data.

Table 24.5 Principal output values

Parameter	Values
T_{gasif} (°C)	702.00
Producer gas composition (on a mole basis)	
CO	23.93
CO_2	10.49
N_2	37.07
H_2	20.88
CH_4	3.58
H_2O	4.03
Flowrate (kg/h)	35.33
LHV (MJ/kg)	6.32
CGE (%)	68.00
Power (kW_e)	15.80
η (%)	17.00

Table 24.6 Economic parameters for pre-treatment units, G-ICE plant, transportation and utilities consumption. Data from Hamelinck and Faaij (2002), Hamelinck et al. (2003), TRAMA (2008), *Ankur Scientific Energy Technologies Pvt. Ltd.*, KITE and WACSO

	Base scale	Base investment	O and M (% of investment)	Utility consumption	Lifetime (yr)
Drying	100 t/h	M$10.5	3	$0.06 \cdot t_{H_2Oev}$ (t diesel)	15
Chipping	80 t/h	M$1.2	20	Bond law 0.15 · t input (kW)	15
G-ICE system	20 kW_e	M$0.05[1]	4		7
Transportation biomass		Tractor full $0.32/km · t	Loading and offloading $1.32/t		
MV network[2]		$5000/km			

[1] LV network costs are included here
[2] Transformer cost is $1000

sites. Therefore, the social criterion *SoC* should be maximised (see Eq. 24.12). This criterion assigns a value of 1 to each unit installed per site f. If V is the binary variable that characterises the number of units installed per site, this metric can be expressed as follows:

$$SoC = \sum_j \sum_f \sum_t V_{jft} \qquad \forall j,f,t \tag{24.12}$$

24.3.3.2 Results

Three scenarios are obtained by optimizing the individual metrics: NPV, Impact 2002+ and *SoC*. The optimal networks are represented by four types of matter flows that connect the different sites. Those flows are: raw material from harvesting, stored raw material, dried matter and chipped matter, characterised by different colours in Fig. 24.10. Flowrates are obtained in tons/month. All networks distribute dried material. The chipped biomass is the most used alternative. Since it has been assumed that cassava waste is produced into each community, proportional to the population, there is no need to employ a MV microgrid and use intermediate sites. Moreover, the high investment costs associated with an intermediate site prevent their use. The simplest network comes out from the environmental impact minimisation, with only chipped matter being generated. Table 24.7 lists the calculated capacity for the equipments installed at each site for each scenario.

Table 24.8 summarises the three criteria evaluated for each optimal network. If the most environmental friendly option is selected, the decision maker should be willing to compromise around 60 % of the optimum NPV which implies just a 2 % improvement in the environmental metric. In order to decrease the environmental impact, there is a necessity of more decentralised units. This requires a significant greater investment in comparison with the NPV optimum alternative. Notwithstanding, the gain in the environmental impact only results from a reduction in transportation. The environmental impact for the social SC network is closer to the optimal environmental impact. The maximum value for the social criterion is 27, installing 3 units per site. In this case the capacity of the installed units is adapted to match the demand of each community. However, a certain degree of centralisation is needed in network to ensure its financial sustainability.

The needed amount of biomass to satisfy the electricity demand is 1526 t/yr. An 8.4 % of the total available cassava rhizome is not used and can be employed for other purposes. The most important differences among the three selected criteria concern transportation and investment. Due to the cassava waste disposition, which is present in all the communities, no MV microgrid is installed. The smallest communities whose demands are far from the biggest ones, such as Seneso or Nwunwom, are the communities that show more variability along the different scenarios. The largest variability comes from the chipper and dryer installations, since those are the units used to adjust the social factor during the search of optimal scenarios. Even though there is enough biomass to be self-sufficient, biomass from other communities is processed to supply the demand of small communities. By doing so, the investment in pre-treatment technologies is reduced. Such a reduction has a major contribution than the savings in transportation that would be gained if the biomass were pre-treated locally for these small communities. It is also relevant to notice that transportation of raw biomass is carried out from some small communities to other ones with installed pre-treatment capacity in order to face biomass seasonality. Also, it is important to highlight that this type of models can be used to explore the sensitivity of the solutions to important parameters. For example, the effect of the electricity price ($0.233/kWh) on the optimal NPV network can be analysed. It is found that if the

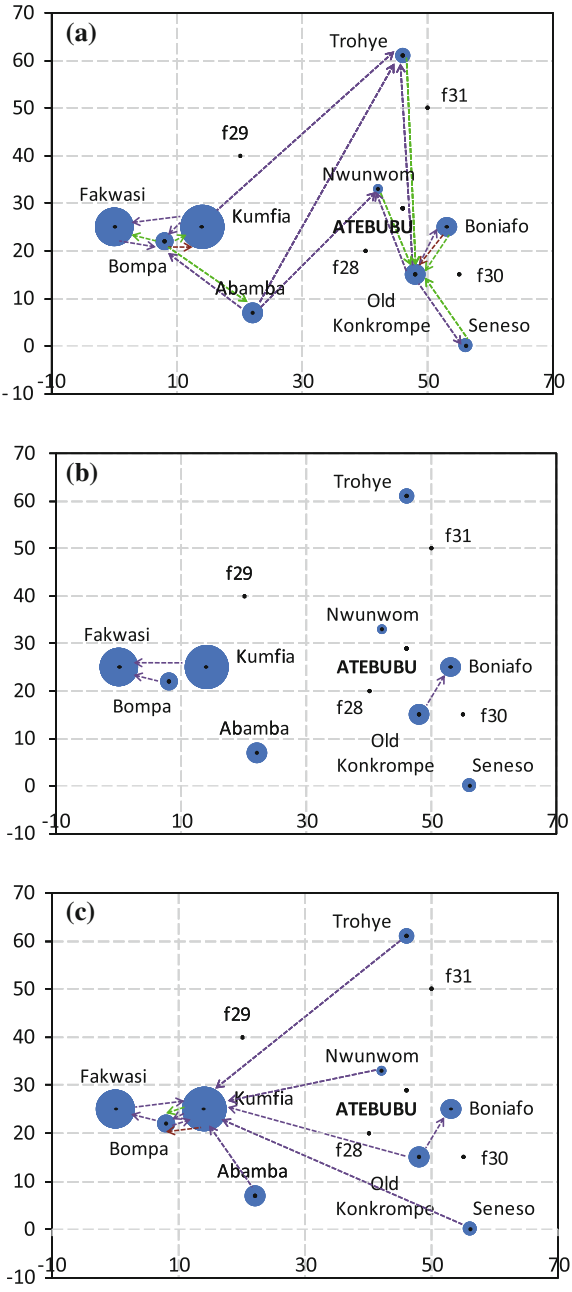

Fig. 24.10 Optimal network configurations for the three selected criteria. See in *brown* the non-stored raw matter flow, in *green*, the raw matter flow, in *orange* the dry matter flow and in *purple* the chipped matter flow. **a** Optimal NPV network configuration. **b** Optimal Impact 2002+ network configuration. **c** Optimal social network configuration

Table 24.7 Equipment capacity for the optimum networks configurations obtained for the three selected criteria

		NPV max			Impact min			*SoC* max	
	Dryer (t/h)	Chipper (t/h)	G-ICE (MJ/h)	Dryer (t/h)	Chipper (t/h)	G-ICE (MJ/h)	Dryer (t/h)	Chipper (t/h)	G-ICE (MJ/h)
Seneso			18.00	0.10	0.10	18.00	0.10	0.10	18.00
Old	0.10	0.10	21.44	0.10	0.10	21.44	0.10	0.10	21.44
Konkrompe									
Fakwasi	0.14	0.10	80.61	0.14	0.10	80.61	0.10	0.10	80.61
Kumfia	0.20	0.11	102.35	0.19	0.10	102.35	0.12	0.10	102.35
Trohye			18.00	0.10	0.10	18.00	0.10	0.10	18.00
Bompa			18.00	0.10	0.10	18.00	0.10	0.10	18.00
Nwunwom			18.00	0.10	0.10	18.00	0.10	0.10	18.00
Boniafo			20.53	0.10	0.10	20.53	0.10	0.10	20.53
Abamba	0.10	0.10	22.04	0.10	0.10	22.04	0.10	0.10	22.04

Table 24.8 Economic, environmental and social aspects for the individual objective functions optimised networks

NPV optimisation ($)	89895.95
Impact 2002+ (pts/yr)	113.46
Social criterion	17.00
NPV ($)	36867.21
Impact 2002+ optimisation (pts/yr)	110.94
Social criterion	27.00
NPV ($)	45155.60
Impact 2002+ (pts/yr)	111.43
Social optimisation	27.00

electricity price falls below $0.2/kWh and equipment investment remains the same then the viability of the network breaks down, i.e. the NPV becomes negative.

24.3.4 Biomass Supply Chains in Developed Economies

This last case study is a retrofitting proposal for coal combustion power plants in Spain that contemplates the use of biomass to replace a fraction of coal. It was first presented in Pérez-Fortes et al. (2014).

24.3.4.1 Case Study

Provided a set of biomass collection sites and the current list of power plants in Spain, the SC model assists on the decisions associated with the technology allocation problem and the flows of materials between sites, while quantifying the performance of the proposed configuration in terms of NPV and environmental impact, IMPACT 2002+.

The types of biomass waste used in this case study are forest wood residues (FWR) and agricultural woody residues (AWR) from Gómez et al. (2010a, b). The amount of biomass available in collection areas is estimated by an approach that integrates physical, geographical and technical limitations, providing an upper bound for the potential availability Gómez et al. (2010b). It is assumed that no transportation cost is charged for biomass assembly inside the collection areas. Table 24.9 sums up the main characteristics of FWR and AWR used for modelling purposes.

The technology considered for biomass usage in a combustion plant is co-firing. According to Basu et al. (2011), the investment is around 192 €/kW_{th} for this technology. Operation and maintenance (O&M) costs represent 4 % of the investment Gómez et al. (2010a). The type of coal used into each plant is specified in López-Vilariño et al. (2003). See Tables 24.10 and 24.11 for further detail about the power plants, their coal origin and power produced.

Table 24.9 Feedstock characteristics Gómez et al. (2010b)

Biomass waste	MC (%wt)	LHV_{ar} (MJ/kg)	BD (kg/m^3)	Yearly available (kton)	Adjusted availability (kton)	Seasonality	Cost (€/ton)
FWR	30	12.5	140	7,748	1,162	None summer	56
AWR	40	10.8	100	3,883	2,718	and winter	52

Table 24.10 Types of coal used in the Spanish power plants. Data for 2010 ENERCLUB (2010); López-Vilariño et al. (2003); REE (2010)

	Type	Origin	LHV_{ar} (MJ/kg)	Cost (€/t)
Coal 1	Sub-bituminous 1	Local	12.57	85
Coal 2	Sub-bituminous 2	Local	17.81	85
Coal 3	Bituminous 1	Local	22.63	85
Coal 4	Bituminous 2	Imported	27.03	80

Biomass storage is allowed after harvesting/collection and after pre-treatment sites. This case study considers open air covered storage for raw material. MC decreases due to natural drying and DM is reduced due to degradation: losses of 2 % MC and 0.25 % DM can be accounted Rentizelas et al. (2009); Maciejewska et al. (2006).

Figure 24.11 shows the general network configuration considered in this case study. After being collected, the biomass waste may be transported to different sites to be stored. Following, there are two mandatory processes before biomass pre-treatment to obtain the mandatory conditions of MC and shape: chipping and drying. Thereafter, torrefied biomass (TOR), torrefied pellets (TOP), pellets (PEL), bio-oil (OIL) or bioslurry (SLU) are produced before being stored if needed, and processed in the power plant. Trucks, adapted to carry solids or liquids are used biomass distribution.

Table 24.12 summarizes investment, O&M costs and utilities consumption for pre-treatment units, storage and transportation echelons. Utilities cost are 1,393 €/t MITYC (2010) for diesel and 0.04463 €/kWh CNEL (2010) for electricity. The electricity sold is bought at 0.03701 €/kWh OMEL (2011).

24.3.4.2 Results

We are going to focus on two scenarios to demonstrate the capabilities of the biomass SC optimization in the context of developed economies. They are proposed to evaluate the trade-offs among the state -of-the-art (SOTA) pre-treatments, i.e. torrefaction, pelletisation, pelletisation of torrefied biomass, fast pyrolysis and fast pyrolysis combined with char grinding:

Table 24.11 Spanish thermal power plants characteristics ordered by region. Installed power (MW) and used capacity (GWh) for 2010 López-Vilariño et al. (2003); REE (2010)

Model name	Name	Power (MW)	Energy (GWh)	Type of coal
m1	Puentes García Rodríguez	1,468	4,955	Coal 1
m2	Meirama	563	856	Coal 1
m3	Aboño	916	3,663	Coal 3
m4	Lada	513	698	Coal 3
m5	Soto de la Ribera	604	927	Coal 3
m6	Narcea	595	1	Coal 3
m7	Anllares	365	0	Coal 3
m8	Compostilla	1,171	209	Coal 3
m9	La Robla	655	29	Coal 3
m10	Guardo	516	63	Coal 3
m11	Pasajes de San Juan	217	487	Coal 4
m12	Cercs	162	516	Coal 2
m13	Escatrón	80	0	Coal 2
m14	Teruel	1,102	1,793	Coal 2
m15	Escucha	159	156	Coal 2
m16	Litoral de Almería	1,159	4,409	Coal 4
m17	Los Barrios	589	2,489	Coal 4
m18	Puertollano	221	255	Coal 3
m19	Puentenuevo	324	590	Coal 3
Total		11,379	22,096	

- Scenario A (SCNA). This option considers as alternatives all the proposed pre-treatments. The model is forced to select at least one SOTA pre-treatment.
- Scenario B (SCNB). This alternative optimizes the SC considering as SOTA pre-treatments only the use of fast pyrolysis and fast pyrolysis with char grinding.

Figures 24.12 and 24.13 illustrate the optimum networks obtained for each criterion and scenario. With regard to SCNA, Fig. 24.12a, b show that the difference between the NPV and IMPACT 2002+ optimisation results is around the distribution among sites. The environmental optimisation does not suggests transportation of intermediates between sites which indeed contributes to reduce the overall impact. Moreover, this is achieved by installing pre-treatment unit in each collection area. Meanwhile, the NPV takes advantage of economies of scale by suggesting equipment of higher capacity at fewer locations.

Figure 24.13a, b depict the optimum networks for fast pyrolysis combined with char grinding. Both networks are very similar. There is mostly transportation of raw biomass from suppliers to intermediate sites despite the biochar being more

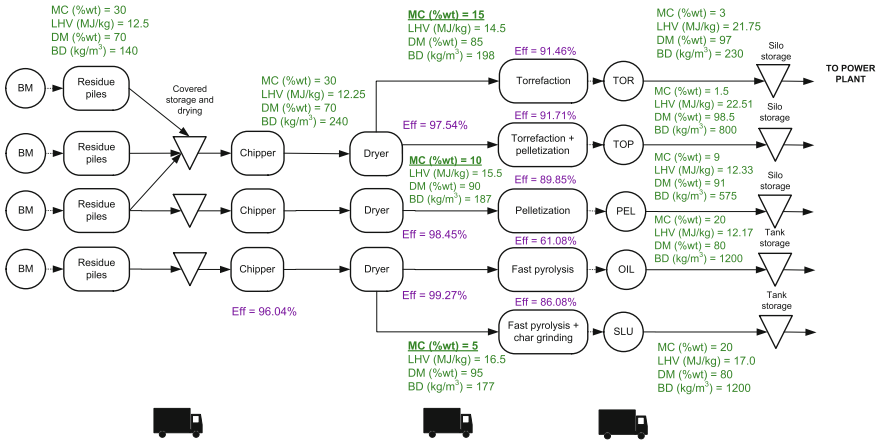

Fig. 24.11 Properties for a FWR stream along the SC. LHV is on *ar* basis. The efficiency (*Eff*) is defined in terms of LHV_{ar}

Table 24.12 Economic parameters of pre-treatment and storage units and transportation Hamelinck et al. (2003); Magalhaes et al. (2009); Uslu et al. (2008)

	Base scale	Base investment	O and M (% of investment)	Utility consumption	Lifetime (yr)
Chipping	80 t/h	0.7 M€	20	Bond law $0.15 \cdot t_{input}$ (kW)	15
Drying	100 t/h	6.9 M€	3	$0.06 \cdot t_{H_2Oev}$ (t diesel)	15
Torrefaction	40 $MW_{th_{in}}$	6.2 M€	5	92 kWh/t_{input}	10
Torrefaction +					
Pelletization	40 $MW_{th_{in}}$	7.5 M€	5	102 kWh/t_{input}	10
Pelletization	40 $MW_{th_{in}}$	5.9 M€	5	129 kWh/t_{input}	10
Fast pyrolysis	40 $MW_{th_{in}}$	10.5 M€	4	75 kWh/t_{input}	25
Fast pyrolysis +					
char grinding	5 t/h	4.9 M€	5	95 kWh/t_{input}	25
Open air covered storage			0.53 €/m^3month		
Silo storage	5000 m^3	0.45 M€	3		25
Tank storage	2272 m^3	1.11 M€	3		25
Transportation	130 m^3		0.69 €/m^3		
solid			1.16 €/km		
Transportation	33 m^3		0.69 €/m^3		
liquid			1.70 €/km		

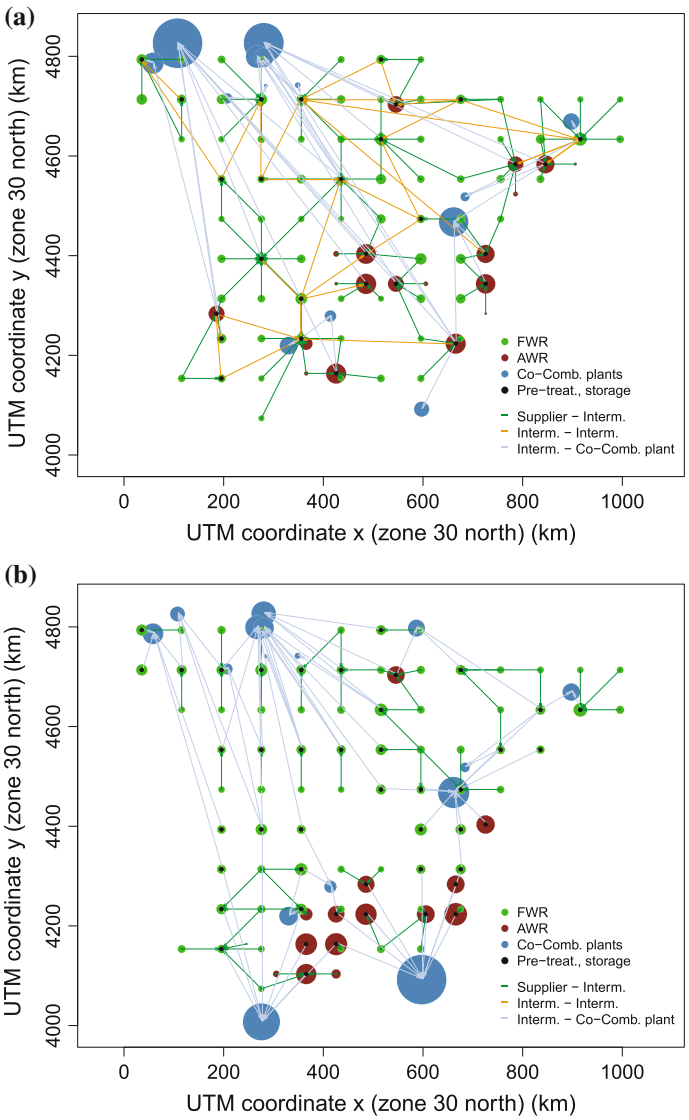

Fig. 24.12 Optimal network configurations for the two selected criteria and mandatory SOTA pretreatments (SCNA). **a** Optimal NPV network configuration for SCNA. **b** Optimal Impact 2002+ network configuration for SCNA

dense. However, observe that liquid product transportation is carried out by trucks of smaller capacity. From the size of the fast pyrolysis plants installed and also considering the efficiencies of chipping and drying (as for example calculated in Fig. 24.11), it can be observed that the minimum capacity available for chipping and drying (for

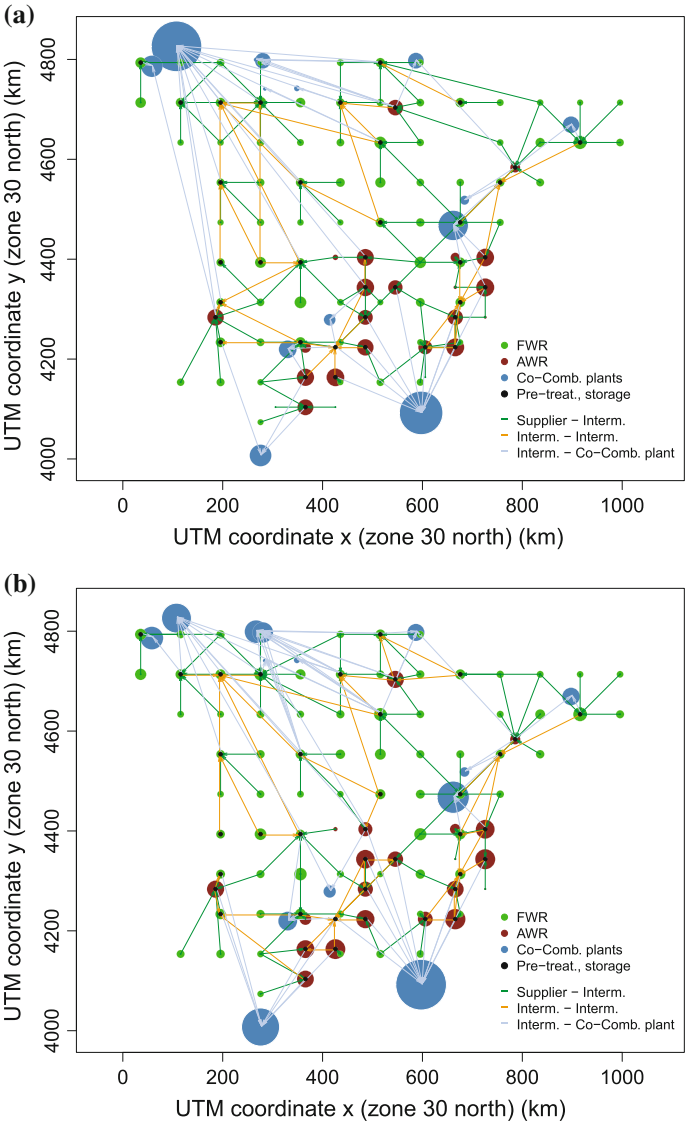

Fig. 24.13 Optimum network configurations for the two selected criteria and production of liquid fuel (SCNB). **a** Optimum NPV network configuration for SCNB. **b** Optimum Impact 2002+ network configuration for SCNB

both scenarios) is larger than input demand. On the other hand, fast pyrolysis units are installed in around 70 % of the candidate nodes at its maximum capacity.

Due the its higher bulk density, liquid biomass is adequate for long distance transportation and for storage in smaller places. Nevertheless, fast pyrolysis is still in a

Table 24.13 Revenues increase to cover the biomass SC

	EcO-SCNA	EnvO-SCNA	EcO-SCNB	EnvO-SCNB
Revenues increase	15.30 %	20.66 %	21.71 %	23.24 %

pre-commercial state and therefore, is more expensive than other alternatives. In addition to that, according to the data for this specific case study, trucks to move liquid biomass are smaller and more expensive than trucks used to transport solid biomass. Under the considered data assumptions, we can concluded that to make liquid biofuels attractive their distribution cost should drop by 30 %, or from another point of view, the distances to be covered should increase by a factor a 1.4–1.5 to justify the conversion into biochar. By comparing the (i) consumed biomass (i.e. efficiency of the process), (ii) overall cost and (iii) investment for the same networks, liquid fuel results in a needed global expense decrease of 5 % to be similarly "attractive". If this reduction is to come only from technology investment, this should decrease by 15 %.

Other type of outcome from this analysis is listed in Table 24.13. It shows the needed increment in electricity price to afford the investment and operation along 10 years for each of the optimal proposed networks. This increase oscillates between 9 and 23 %.

24.4 Conclusions

This chapter has emphasised some of the insights that an optimisation approach for biomass supply chains can provide. These may result useful to move efficiently towards the goals established in the horizon up to 2050. The approach consists of a MILP formulation that considers long-term strategic decisions such as selection of pre-treatment trains of units and their respective location, selection of biomass sources, location of processing sites, and distribution centres. The problem has been formulated with multiple objective functions: the net present value, the Impact2002+ metric to quantify the overall environmental impact, and a surrogate social metric that considers the processing units installed in each location. The latter is especially significant when addressing problems in developing economies, as the electrification problem presented for the communities in the Atebubu-Amantin district in Ghana. Another two distinctive features of the described approach are (i) the model's capability of combining feeds and determining the optimal proportions of each input for the energy generation activity and (ii) the consideration of passive drying during storage. However, further analysis is necessary for determining the benefits of providing these additional degrees of freedom in contrast to assuming a unique source with average properties. For cases where such benefits are significant, exploring the extension of this flexibility to other activities included in the biomass supply chain, such

as storage and pre-treatment, may prove worthy. Further work is devoted to envisaging decomposition strategies to reduce the computational cost of solving this type of problems. The capabilities of the mathematical modelling approach were pointed out through the different case studies. These have demonstrated that this formulation can be adapted to address different contexts (e.g., centralised and decentralised, rural or state-wide networks).

Finally, we would like to underscore that efforts to consolidate the research carried out related to pre-treatment/conversion technologies into a repository, which would be available to the PSE community, could facilitate the feasibility evaluation of biomass-based-energy projects.

Acknowledgments The authors are grateful to the CEPIMA members for the support provided. We acknowledge the financial funding received from the Generalitat de Catalunya through the ESF (FI Grants).

Disclaimer This work was performed while the author's main affiliation was the Universitat Politècnica de Catalunya. The authors confirm that this study and all related work were carried out without any conflict of interest. Responsibility for the information and views set out in this work lies entirely with the authors.

References

An, H., Wilhelm, W. E., & Searcy, S. W. (2011). Biofuel and petroleum-based fuel supply chain research: A literature review. *Biomass and Bioenergy*, *35*, 3763–3774.

Arranz-Piera, P. (2008). Development of electricity service operators for poverty alleviation in ecuador and peru. In *4th European Conference PV-hybrid and mini-grid*. Athens, Greece, May 2008.

Arranz-Piera, P., Velo, E., & Horta, F. (2011). Guidelines to introduce biomass systems in decentralised electrification programmes for mdg achievement in the sub-saharan african context. case studies of two reference countries: Ghana & mozambique. In *19th European Biomass Conference and Exhibition*. Berlin, Germany, June 2011.

Ayoub, N., Seki, H., & Naka, Y. (2009). Superstructure-based design and operation for biomass utilization networks. *Computers and Chemical Engineering*, *33*, 1770–1780.

Balaman, S. Y., & Selim, H. (2014). A fuzzy multiobjective linear programming model for design and management of anaerobic digestion based bioenergy supply chains. *Energy*, *74*, 928–940.

Basu, P., Butler, J., & Leon, M. A. (2011). Biomass co-firing options on the emission reduction and electricity generation costs in coal-fired power plants. *Renewable Energy*, *36*, 282–288.

Bojarski, A. D., Laínez-Aguirre, J. M., Espuña, A., & Puigjaner, L. (2009). Incorporating environmental impacts and regulations in a holistic supply chains modeling: An lca approach. *Computers & Chemical Engineering*, *33*(10), 1747–1759.

Bowling, I. M., Ponce-Ortega, J. M., & El-Halwagi, M. (2011). Facility location and supply chain optimization for a biorefinery. *Industrial and Engineering Chemistry Research*, *50*, 6276–6286.

Cambero, Claudia, & Sowlati, Taraneh. (2014). Assessment and optimization of forest biomass supply chains from economic, social and environmental perspectives? a review of literature. *Renewable and Sustainable Energy Reviews*, *36*, 62–73.

Caputo, A. C., Palumbo, M., Pelagagge, P. M., & Scacchia, F. (2005). Economics of biomass energy utilization in combustion and gasification plants: Effects of logistics variables. *Biomass and Bioenergy*, *28*, 35–51.

Cherubini, F., & Stromman, A. H. (2011). Life cycle assessment of bioenergy systems: State of the art and future challenges. *Bioresource Technology*, *102*, 437–451.

Club Español de la Energía; Instituto Español de la Energía (Spain). (2010). Retrieved December 2010, from http://www.enerclub.es/es/frontaction.do?action=viewcategory&categoryname=carbón&id=1086.

Comisión Nacional de Energía (Spain). (2010). Retrieved December 2011, from http://www.cne.es/cne/contenido.jsp?id_nodo=423&&&keyword=&auditoria=f.

Commission, European. (2015). *Towards an integrated strategic energy technology (set) plan: Accelerating the european energy system transformation*. Technical report, European Commission.

Damen, K., & Faaij, A. (2006). A greenhouse gas balance of two existing international biomass import chains. The case of residue co-firing in a pulverized coal-fired power plant in The Netherlands. *Mitigation and Adaptation Strategies for Global Change*, *11*, 1023–1050.

Ecoinvent-V1.3. (2006). *The ecoinvent database v1.3*. Technical report, Swiss Centre for Life Cycle Inventories.

Energy efficiency in industrial processes. (1988). Technical report, Report EUR 12046 EN, Commission of the European Communities.

Gebreslassie, Berhane H., Yao, Yuan, & You, Fengqi. (2012). Design under uncertainty of hydrocarbon biorefinery supply chains: Multiobjective stochastic programming models, decomposition algorithm, and a comparison between cvar and downside risk. *AIChE Journal*, *58*(7), 2155–2179.

General Secretariat of the Council. (2014). *Conclusions on 2030 climate and energy policy framework, brussels*. Technical report, European Commission.

Gómez, A., Zubizarreta, J., Rodrigues, M., Dopazo, C., & Fueyo, N. (2010a). An estimation of the energy potential of agro-industrial residues in Spain. *Resources, Conservation and Recycling*, *54*, 972–984.

Gómez, A., Rodrigues, M., Montañés, C., Dopazo, C., & Fueyo, N. (2010b). The potential for electricity generation from crop and forestry residues in Spain. *Biomass and Bioenergy*, *34*, 703–719.

Graves, S. C., & Tomlin, B. T. (2003). Process flexibility in supply chains. *Management Science*, *49*, 907–919.

Hamelinck, C. N., Suurs, R. A. A., & Faaij, A. P. C. (2003). International bioenergy transport costs and energy balance. Technical report, Universiteit Utrecht, Copernicus Institute, Science Technology Society, Utrecht, The Netherlands.

Hamelinck, C. N., & Faaij, A. P. C. (2002). Future propects for production of methanol and hydrogen from biomass. *Journal of Power Sources*, *111*, 1–22.

Humbert, S., Margni, M., Jolliet, O. (2005). *Impact 2002+: User guide draft for version 2.1*. Technical report, Industrial Ecology & Life Cycle Systems Group, GECOS, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland, October 2005.

Kondili, E., Pantelides, C. C., & Sargent, R. W. (1993). A general algorithm for short term scheduling of batch operations. *Computers & Chemical Engineering*, *17*, 211–227.

Laínez, J. M., & Puigjaner, L. (2012). Prospective and perspective review in integrated supply chain modelling for the chemical process industry. *Current Opinion in Chemical Engineering*, *1*(4), 430–445. Biotechnology and bioprocess engineering? Process systems engineering.

Laínez, J. M., Pérez-Fortes, M., Bojarski, A. D., & Puigjaner, L. (2011). *Syngas from waste: Emerging technologies, chapter raw materials supply* (pp. 23–54). London: Springer.

Laínez-Aguirre, J. M., Kopanos, G., Espuña, A., & Puigjaner, L. (2009). Flexible design-planning of supply chain networks. *AIChE Journal*, *55*, 1736–1753.

López-Vilariño, J. M., Fernández-Martínez, G., Turnes-Carou, I., Muniategui-Lorenzo, S., López-Mahía, P., & Prada-Rodríguez, D. (2003). Behavior of fluorine and chlorine in Spain coal fired power plants with pulverized coal boilers and fluidized bed boiler. *Environmental Technology*, *24*(6), 687–692.

Maciejewska, A., Veringa, H., Sanders, J., & Peteves, S. D. (2006). *Co-firing of biomass with coal: Constraints and role of biomass pre-treatment*. Technical report, Institute for Energy and the European Comission (EC).

Magalhaes, A. I. P., Petrovic, D., Rodriguez, A. L., Putra, Z. A., & Thielemans, G. (2009). Techno-economic assessment of biomass pre-conversion processes as a part of biomass-to-liquids line-up. *Biofuels, Bioproducts and Biorefining*, *3*, 584–600.

Mele, F. D., Kostin, A. M., Guillén-Gosálbez, G., & Jiménez, L. (2011). Multiobjective model for more sustainable fuel supply chains. A case study of sugar cane industry in Argentina. *Industrial and Engineering Chemistry Research*, *50*, 4939–4958.

Ministerio de Industria, Turismo y Comercio (Spain). (2010). Retrieved December 2011, from http://www.mityc.es/energia/petroleo/Precios/Informes/Informesmensuales/2010/diciembre2010.pdf.

Miret, C., Chazara, P., Montastruc, L., Domenech, S., & Negny, S. (2016). Design of bioethanol green supply chain?: Comparison between first and second generation biomass concerning economic, environmental and social criteria. *Computers & Chemical Engineering*, *85*, 16–35.

Operador del Mercado Eléctrico (OMEL) (Spain). (2011). Retrieved December 2010, from http://www.omel.es/inicio.

Osmani, Atif, & Zhang, Jun. (2014). Economic and environmental optimization of a large scale sustainable dual feedstock lignocellulosic-based bioethanol supply chain in a stochastic environment. *Applied Energy*, *114*, 572–587.

Pattiya, A. (2011). Thermochemical characterization of agricultural wastes from thai cassava plantations. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, *33*, 691–701.

Pérez-Fortes, M. (2011). *Conceptual Design of Alternative Energy Systems from Biomass*. Ph.D. thesis, UPC, Barcelona, Spain.

Pérez-Fortes, M., Laínez-Aguirre, J. M., Arranz-Piera, P., Velo, E., & Puigjaner, L. (2012). Design of regional and sustainable bio-based networks for electricity generation using a multi-objective MILP approach. *Energy*, *44*, 79–95.

Pérez-Fortes, M., Laínez-Aguirre, A., Bojarski, J. M., & Puigjaner, L. (2014). Optimization of pre-treatment selection for the use of woody waste in co-combustion plants. *Chemical Engineering Research and Design*, *92*, 1539–1562.

Perry, M., & Rosillo-Calle, F. (2008). Recent trends and future opportunities in UK bioenergy: Maximizing biomass penetration in a centralised energy system. *Biomass and Bioenergy*, *32*, 688–701.

PRé Consultants, B.V. (2004). SimaPro 6 LCA software. http://www.pre.nl/simapro/default.htm.

Puigjaner, L., Pérez-Fortes, M. & M. Laínez-Aguirre, J. (2015) Coordinated bioenergy supply chain-a process systems approach. Talianisches Kulturinstitut Berlin, Ambasciata d'Italia—Energy Policy and Climate Changes- Präsentation des Sonderhefts der Zeitschrift? Energies? und Podiumsdiskussion mit Prof. L. Puigjaner.

Puigjaner, L., Pérez-Fortes, M., & Laínez-Aguirre, J. M. (2015). Towards a carbon-neutral energy sector: Opportunities and challenges of coordinated bio-energy supply chains - a pse approach. *Energies*, *8*(6), 5613–5660.

Red Eléctrica de España (Spain). (2010). El sistema eléctrico español (2010). Technical report, Red Eléctrica de España (Spain). http://www.ree.es/sistema_electrico/pdf/infosis/inf_sis_elec_ree_2010.pdf.

Rentizelas, A. A., Tatsiopoulos, I. P., & Tolis, A. (2009). An optimization model for multi-biomass tri-generation energy supply. *Biomass and Bioenergy*, *33*, 223–233.

Serpagli, A., Nagori, G. P., Amoriggi, G., & Calvosa, C. (2010). *The use of cassava wastes to produce energy: outcomes of a feasibility study implemented in ghana*. Technical report, International Fund for Agricultural Development (IFAD), Rome, Italy.

Shastri, Yogendra, Hansen, Alan, Rodríguez, L., & Ting., K. C. (2011). A novel decomposition and distributed computing approach for the solution of large scale optimization models. *Computers and Electronics in Agriculture*, *76*(1), 69–79.

Tchapda, A. H., Pisupati, & S. V. (2014). A review of thermal co-conversion of coal and biomass/waste. *Energies*, *7*(3), 1098.

TTA. Trama Tecnoambiental. (2008). *Estudio de base y definición de proyectos piloto para un programa nacional de electrificación rural en paraguay*. Technical report, Inter-American Development Bank, Washingtonn, USA—Asunción, Paraguay.

Uslu, A., Faaij, A. P. C., & Bergman, P. C. A. (2008). Pre-treatment technologies, and their effect on international bioenergy supply chain logistics. Techno-economic evaluation of torrefaction, fast pyrolysis and pelletization. *Energy*, *33*, 1206–1223.

Vallvé, X., González, S., & Arranz-Piera, P. (2007). *Solar-diesel hybrid options for the peruvian amazon—lessons learnt from padre cocha*. Technical report 111/07, Energy Sector Management Assistance Program (ESMAP).

You, F., Tao, L., Graziano, D. J., & Snyder, S. W. (2012). Optimal design of sustainable cellulosic biofuel supply chains: Multiobjective optimization coupled with life cycle assessment and input-output analysis. *Aiche Journal*, *58*(4), 1157–1179.

Yue, D., You, F., & Snyder, S. W. (2014). Biomass-to-bioenergy and biofuel supply chain optimization?: Overview, key issues and challenges. *Computers & Chemical Engineering*. doi:10.1016/j.compchemeng.2013.11.016.

Chapter 25
A Composite-Curve-Based Biomass Procurement Planning Approach

WenZhao Wu, Daniel Kurniawan, WenBo Zhu and Christos T. Maravelias

Abstract The production of transportation fuels from biomass is a promising renewable alternative to traditional fossil fuels. To achieve low carbon footprint of the overall biofuel supply chain however an efficient biomass procurement plan is essential. To this end, we discuss a novel approach to biomass procurement planning. In terms of transportation, we propose a region-to-point modeling approach based on mathematical integration over the sourcing region that has unique characteristics such as shape, location, and productivity. Both algebraic and numerical solution methods are discussed. In terms of system-level procurement planning, we develop a composite-curve-based approach that incorporates the regional transportation modeling method, and aims at identifying the biomass procurement plan that minimizes the total procurement cost (including biomass purchasing, harvesting and transportation). The specific steps for the generation of the composite curve, as well as insights into the procurement planning problem are discussed. We complete the chapter with a case study illustrating the applicability of the proposed methods.

25.1 Introduction

The focus on the production of biofuels from biomass has been increasing recently due to environmental concerns such as climate change (DOE Bioenergy Technologies Office 2014). Since the production of biofuels involves relatively expensive feedstock and energy-intensive biomass transportation, any biomass-to-fuels strategy should include an efficient, both in terms of cost and environmental impact, biomass (feedstock) supply chain (SC).

W. Wu · D. Kurniawan · W. Zhu · C.T. Maravelias (✉)
Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA
e-mail: maravelias@wisc.edu

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_25

A manufacturing supply chain includes four functions: (1) procurement, i.e. delivery of biomass feedstocks from sourcing areas (i.e. farms) to the bio-refinery (referred to as "refinery" hereafter); (2) production, i.e. conversion of biomass to biofuels in the refinery; (3) distribution, i.e. delivery of the biofuels from the refinery to distribution centers; and (4) sales. The focus of this chapter is on the planning of biomass procurement.

Farms can produce various types of biomass, which is generally classified as either wastes, such as corn stover, bagasse and wood chips; or dedicated energy crops, such as switchgrass and miscanthus (Brown and Brown 2012). The biomass mainly consists of three components: cellulose, hemicellulose, and lignin. The biomass feedstock is first transported from the farms to the refinery, possibly through regional depots. In the refinery, the biomass is generally pre-treated to make it more susceptible to further conversion steps (Mosier et al. 2005). Then, cellulose and hemicellulose are converted to biofuels using various conversion methods, and lignin is used for heat and power generation (Brown and Brown 2012).

Multiple studies have looked into the planning of biomass supply chains taking into account economic and environmental aspects (Sharma et al. 2013; Josephine and Floudas 2014; Garcia and You 2015), including a range of different methodologies (Liu et al. 2010; Yue et al. 2014; Sharma et al. 2013), such as deterministic mixed-integer linear and nonlinear programming (An et al. 2011; Akgul et al. 2012; Andersen et al. 2013; Elia et al. 2011; Marvin et al. 2012a, b), multi-objective optimization (Santibanez-Aguilar et al. 2011; Liu et al. 2010), and optimization under uncertainty (Osmani and Zhang 2013; Tong et al. 2013).

Unlike fossil fuels, biomass, as a low-energy density resource, is sparsely distributed. The efficient biomass transportation thus requires biomass procurement planning methods (DOE/EERE 2013). In many studies, the farms are treated as points without shape or area (You et al. 2012), as shown in Fig. 25.1a. This is a reasonable assumption when the transportation distance is so large that the shape and size of the farms can be neglected. In this case, the transportation problem is modeled as a *point-to-point* (farm-to-refinery) problem. However, the shape and size of the farms cannot be neglected when the refinery is close to the farm, which

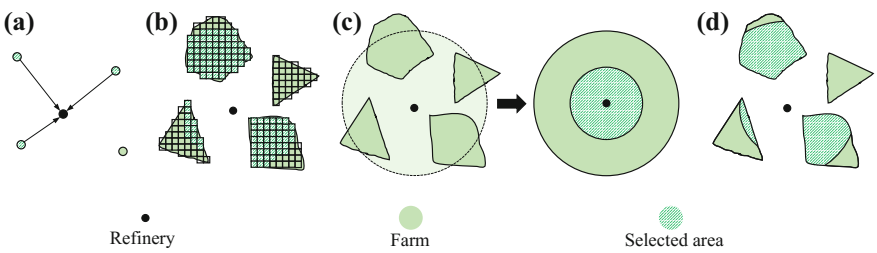

Fig. 25.1 Biomass transportation modeling methods. **a** Point-to-point; **b** region-to-point, with partitioning; **c** region-to-point, with "equivalent disk"; **d** region-to-point, our method with mathematical integration

means that the size of the farm is not significantly smaller than the transportation distances, which in turn means that the error in approximating the real transportation distance with the distance between the center of the farm and the biorefinery can be quite large. In this case, transportation should be treated as a *region-to-point* problem. One method to achieve this is by partitioning each farm into multiple small subregions (boxes) (Zhang et al. 2011), as shown in Fig. 25.1b. When the model is solved, the boxes with non-zero amount of biomass constitute the selected area of the farm from which biomass is transported. Thus, the region-to-point problem is converted to a point-to-point problem. Alternatively, an average biomass distribution density is first calculated among the farms in a given area, and then, as shown in Fig. 25.1c, it is assumed that the biomass is transported from an *equivalent disk* (Overend 1982; Uslu et al. 2008). The shapes and locations of the farms are neglected in this method which can again lead to large deviations between the calculated and actual distance.

Transportation options include roads, railways, and waterways. The most likely means is by road using haulage wagons, trucks, and tractors (Hamelinck et al. 2005; Miao et al. 2012). This is the only transportation method considered in this chapter. In terms of transportation cost, there are three major measurements for point-to-point distance: flow path distance (the exact traveling distance between the two points based on existing roads), rectilinear distance (the distance along two edges perpendicular to each other, connecting the two points), and straight line distance (the shortest distance between the two points; typically used in combination with a *tortuosity factor*, which accounts for the road infrastructure). We adopt the straight line distance measurement in conjunction with the tortuosity factor (Sultana and Kumar 2014).

The remainder of this chapter is structured as follows. In Sect. 25.2, we develop a region-to-point biomass transportation modeling method based on mathematical integration (including two examples). In Sect. 25.3, we develop a composite curve-based approach for the system level procurement planning optimization. In Sect. 25.4, we provide a case study to demonstrate our methods.

25.2 Region-to-Point Biomass Transportation Modeling Approach

A general representation of the biofuel SC system is shown in Fig. 25.1d. Each farm is regarded as a region with unique characteristics (shape, biomass productivity, and location relative to the refinery). Biomass is assumed to be distributed evenly in each farm; if the biomass is unevenly distributed, then the farm can be divided into multiple farms with evenly distributed biomass. Moreover, we assume demand of a single type of biomass. If multiple types of biomass are available, then the proposed method can be extended to a multi-curve method. We assume that the farms are approximated as convex polytopes. If they are non-convex regions, then

they can be divided into multiple convex polytopes. Once the vertices of the polytopes are given, the inequalities describing the polytopes can be obtained. Thus, the farms can be described by a set of linear constraints. In this section, we only consider a one-farm-one-refinery system. In the next section, we discuss how the one-farm system can be incorporated into a general multi-farm system. Note that in contrast to the traditional methods (Fig. 25.1b, c), we neither partition the farms into small boxes, nor assume an equivalent circle.

For simplicity, the refinery is located at the origin of the reference system. For an infinitesimal area on the farm shown in Fig. 25.2a, the following basic equations hold in a polar coordinate system:

$$dM = \mu R dR dQ \tag{25.1}$$

$$dC^{\text{tran}} = \omega \tau R dM \tag{25.2}$$

where,

M the amount of biomass (referred to as "mass" hereafter) transported from the farm (T/year; "T" = metric tonne);
R radius in the polar coordinate system (km);
Q angle in the polar coordinate system (rad);
C^{tran} transportation cost ($/year);
μ biomass productivity (T/(km$^2 \cdot$ year));
ω unit transportation cost ($/(km $\cdot$ T));
τ tortuosity factor, i.e. ratio of actual distance travelled to straight line distance

Since the biomass is assumed to be evenly distributed in a given farm, transporting the nearest biomass is economically most favorable. Thus, as shown by the dash-shaded area in Fig. 25.2b, the "selected area" D from which the biomass is transported is represented by the intersection of a circular sector of radius R and the farm. The mass transported from area D is then generally expressed as follows using polar coordinates:

$$M = \iint_D \mu R dR dQ \tag{25.3}$$

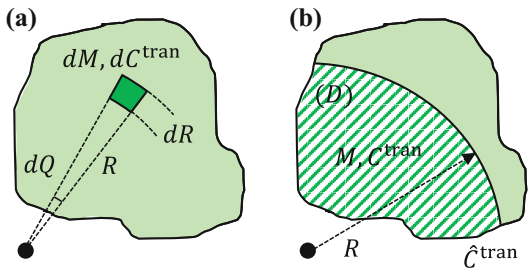

Fig. 25.2 Graphic representation of the region-to-point transportation modeling method. **a** Variables related to the infinitesimal area; **b** variables related to the selected area

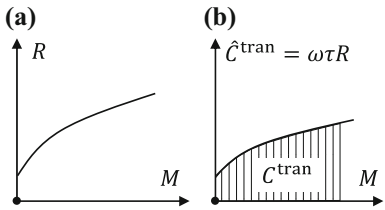

Fig. 25.3 $M-R$ curve and $M-\widehat{C}^{\text{tran}}$ curve. **a** $M-R$ curve; **b** $M-\widehat{C}^{\text{tran}}$ curve; C^{tran} is represented by the shaded area

Since the farm is described by a set of constraints, the selected area D is uniquely determined once the circular sector radius R is specified (see Fig. 25.2b). In this chapter, if Y is a function of X we write: $Y=y(X)$, where y denotes the corresponding function. The curve representing this function is called the X–Y curve. Thus, $M=m(R)$ can be calculated from Eq. 25.3. Since M monotonically increases with R (the larger the radius R of a circular sector, the more mass is selected), R is also a function of M, and thus $R=r(M)$ (see the M–R curve in Fig. 25.3a).

Based on Eq. 25.2 and the $R=r(M)$ expression obtained above, the transportation cost corresponding to mass M is calculated as follows:

$$C^{\text{tran}} = \int_0^M \omega\tau R dM = \int_0^M \omega\tau r(M) dM \tag{25.4}$$

We further define the following variable:

$$\widehat{C}^{\text{tran}} = \omega\tau R \tag{25.5}$$

representing the cost to transport a unit of mass at radius R (\$/T). Based on Eq. 25.5 and $R=r(M)$, function $\widehat{C}^{\text{tran}}=\hat{c}^{\text{tran}}(M)$ is obtained, as shown in Fig. 25.3b. Combining Eqs. 25.4 and 25.5, we obtain the following:

$$C^{\text{tran}} = \int_0^M \hat{c}^{\text{tran}}(M) dM \tag{25.6}$$

which means that C^{tran} can be calculated as the integration of $\hat{c}^{\text{tran}}(M)$ from 0 to M, as represented by the shaded area in Fig. 25.3b. The graphic representations of R, M, $\widehat{C}^{\text{tran}}$ and C^{tran} are summarized in Fig. 25.2b. Note that the $M-\widehat{C}^{\text{tran}}$ curve (Fig. 25.3b) can be regarded as the characteristic curve for a given farm, because $\widehat{C}^{\text{tran}}$ divided by $\omega\tau$ gives R, and the integration of $\widehat{C}^{\text{tran}}$ gives C^{tran}. Therefore, our major task is to generate the $M-\widehat{C}^{\text{tran}}$ curve.

To demonstrate how to generate the $M-\widehat{C}^{\text{tran}}$ curve, we describe how to algebraically calculate the double integral in Eq. 25.3, and then present the corresponding numerical solution method.

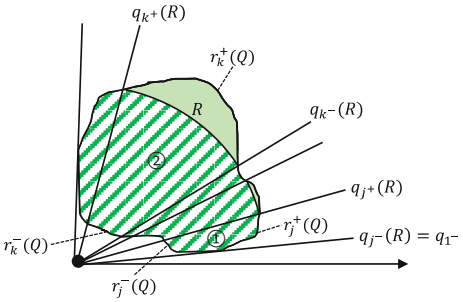

Fig. 25.4 Double integration in Eq. 25.3. For any R, the corresponding M is calculated by doing double integration in the polar coordinate system

As shown in Fig. 25.4, by connecting the refinery with the intersections between the arc of radius R and the farm border, we generally obtain two types of sections: sections where all the biomass is selected ($j \in \mathbf{J}$), and sections where a portion of the biomass is selected ($k \in \mathbf{K}$).

Sections $j \in \mathbf{J}$ (e.g. Section ① in Fig. 25.4) are determined by: (1) the lower and upper borders of the farm in the corresponding sections, with functions $r_j^-(Q)$ and $r_j^+(Q)$, respectively; and (2) the minimum and maximum angles restricting the sections, $q_{j^-}(R)$ and $q_{j^+}(R)$ respectively, which are the roots of $R = r_j^+(Q)$ (treating R as a constant), or known values determined by the farm border. For example, $q_{j^-}(R) = q_{1^-}$ in Section ①, where q_{1^-} is the minimum angle of the points on the farm.

Sections $k \in \mathbf{K}$ (e.g. Section ② in Fig. 25.4) are determined by: (1) the lower borders of the farm (with function $r_k^-(Q)$) and radius R in the corresponding sections; and (2) the minimum and maximum angles restricting the sections, $q_{k^-}(R)$ and $q_{k^+}(R)$ respectively, which are the roots of $R = r_k^+(Q)$, when treating R as a constant.

Equation 25.3 is then, in general, written as follows (in the polar coordinate system):

$$M = \sum_i \int_{q_{i^-}(R)}^{q_{i^+}(R)} \int_{r_i^-(Q)}^{r_i^+(Q)} \mu R dR dQ + \sum_j \int_{q_{j^-}(R)}^{q_{j^+}(R)} \int_{r_j^-(Q)}^{R} \mu R dR dQ \tag{25.7}$$

The right-hand-side of Eq. 25.7 is a function of R only. Therefore, $M = m(R)$. Since M monotonically increases with R, $R = r(M)$. Further, with $\widehat{C}^{\text{tran}} = \hat{c}^{\text{tran}}(R)$ from Eq. 25.5, $\widehat{C}^{\text{tran}} = \hat{c}^{\text{tran}}(R)$ is obtained.

In many cases, deriving an algebraic expression using Eq. 25.7 can be challenging. Therefore, we present a numerical solution method to obtain $M = m(R)$, and subsequently $\widehat{C}^{\text{tran}} = \hat{c}^{\text{tran}}(R)$. Specifically, as shown in Fig. 25.5a, the farm is first divided into n discrete partial annuli of width Δr (the smaller the more accurate is the numerical solution), and thus the average radius R of the inner and outer radii of the N^{th} partial annulus, and the corresponding $\widehat{C}^{\text{tran}}$ are calculated as follows:

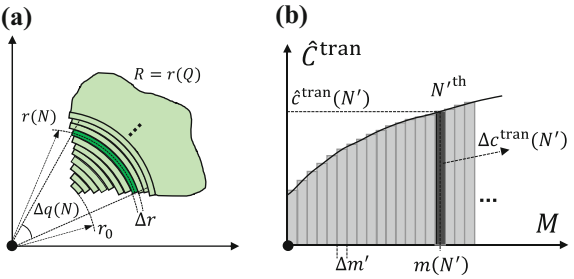

Fig. 25.5 Numerical solution methods for Eqs. 25.3 and 25.6. **a** Numerical solution method for Eq. 25.3; **b** numerical solution method for Eq. 25.6

$$R = r(N) = r_0 + (N-1)\Delta r + \frac{\Delta r}{2} \tag{25.8}$$

$$\widehat{C}^{\text{tran}} = \hat{c}^{\text{tran}}(N) = \omega\tau r(N) \tag{25.9}$$

where $r_0 =$ distance from the inner radius of the first partial annulus to the refinery. The mass ΔM for the N^{th} partial annulus is then calculated as follows:

$$\Delta M = \Delta m(N) = \mu r(N)\Delta r \Delta q(N) \tag{25.10}$$

where $\Delta q(N)$ is the angle corresponding to the N^{th} partial annulus, which is determined by the intersection between the farm border $R = r(Q)$ and the circle of radius $r(N)$; $\Delta q(N)$ is thus the difference between the two roots of $r(Q) = r(N)$. Next, the masses of the N partial annuli are summed up to obtain M:

$$M = m(N) = \sum_{1}^{N} \Delta m(N) \tag{25.11}$$

Equations 25.8–25.11 are implemented for all $1 \le N \le n$. Since M monotonically increases with N, N is a function of M. Thus, we obtain $\widehat{C}^{\text{tran}} = \hat{c}^{\text{tran}}(N) = \hat{c}^{\text{tran}}(M)$ for the given farm, represented by n pairs of M and $\widehat{C}^{\text{tran}}$ values. Subsequently, $\widehat{C}^{\text{tran}} = \hat{c}^{\text{tran}}(M)$ can be expressed as a fitted function such as a piece-wise linear function, and a polynomial function.

Similarly, we present the numerical solution method to calculating C^{tran} from Eq. 25.6, when deriving the algebraic expression is challenging. As shown in Fig. 25.5b, the space under the given $M - \widehat{C}^{\text{tran}}$ curve is divided into n' rectangles with width $\Delta m'$. Thus, C^{tran} is calculated with the following numerical reformulation:

$$\begin{cases} M = m(N') = (N'-1)\Delta m' + \frac{\Delta m'}{2} \\ \Delta C^{\text{tran}} = \Delta c^{\text{tran}}(N') = \Delta m' \hat{c}^{\text{tran}}(N') \\ C^{\text{tran}} = c^{\text{tran}}(N') = \sum_{1}^{N'} \Delta c^{\text{tran}}(N') \end{cases} \tag{25.12}$$

where $N^{'}$ denotes the number of the rectangles, and $\Delta c^{\text{tran}}\left(N^{'}\right)$ denotes the area of the $N^{'\text{th}}$ rectangle. Thus, the areal sum of the $N^{'}$ rectangles correspond to $c^{\text{tran}}\left(N^{'}\right)$.

Note that all the calculations discussed above are performed in the polar coordinate system. If the farm is originally described in a Cartesian system (e.g. see Example 2 in Sect. 25.2.2), then the following basic equations can be used to represent the description in the polar coordinate system.

$$X = R\cos Q, Y = R\sin Q \tag{25.13}$$

$$X^2 + Y^2 = R^2 \tag{25.14}$$

Next, we present two simple examples to demonstrate the general methods, where the algebraic and the numerical solution methods are applied, respectively.

25.2.1 Example 1: Ring Farm, Algebraic Solution

We consider a ring-shaped farm located between $r_1 = 3$ and $r_2 = 6$, with minimum and maximum angles $q_1 = \frac{\pi}{3}$ and $q_2 = \frac{2\pi}{3}$, respectively (see Fig. 25.6). Thus, the farm is described by the following constraints:

$$\begin{aligned} &3 \le R \le 6 \\ &\frac{\pi}{3} \le Q \le \frac{2\pi}{3} \end{aligned} \tag{25.15}$$

The following parameters are assumed: $\mu = 5000$ T/(km$^2 \cdot$ year); $\tau = 1$; $\omega = 0.1$ \$/(km $\cdot$ T). Thus, based on Eqs. 25.5, 25.6 and 25.7, we calculate the following.

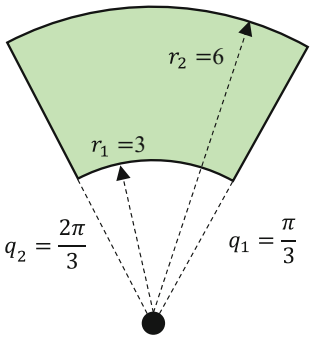

Fig. 25.6 Example 1: ring farm

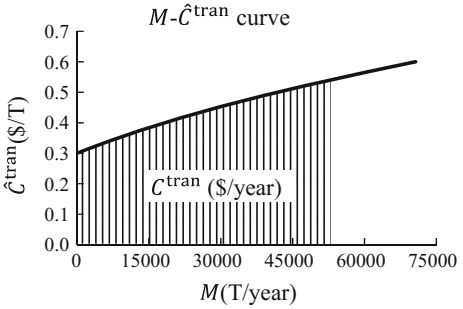

Fig. 25.7 $M-\widehat{C}^{\text{tran}}$ curve for the ring farm in Example 1

$$M=\int_{q_1}^{q_2}\int_{r_1}^{R}\mu RdRdQ=\int_{\frac{\pi}{3}}^{\frac{2\pi}{3}}\int_{3}^{R}\mu RdRdQ=\frac{2500\pi}{3}\left(R^2-9\right),\ 3\le R\le 6 \tag{25.16}$$

$$R=\sqrt{\frac{3M}{2500\pi}+9},0\le M\le 22500\pi \tag{25.17}$$

$$\widehat{C}^{\text{tran}}=\omega\tau R=0.1R,\ 3\le R\le 6 \tag{25.18}$$

$$\widehat{C}^{\text{tran}}=\frac{1}{10}\sqrt{\frac{3M}{2500\pi}+9},0\le M\le 22500\pi \tag{25.19}$$

$$C^{\text{tran}}=\int_0^M \hat{c}^{tran}(M)dM=\frac{500\pi}{9}\left(\frac{3M}{2500\pi}+9\right)^{\frac{3}{2}}-1500\pi,0\le M\le 22500\pi \tag{25.20}$$

Using Eqs. 25.17 and 25.18, $\widehat{C}^{\text{tran}}=\hat{c}^{\text{tran}}(M)$ is obtained in an algebraic form as shown in Eq. 25.19. The $M-\widehat{C}^{\text{tran}}$ curve is generated in Fig. 25.7.

25.2.2 Example 2: Square Farm, Numerical Solution

We consider a system where the vertex coordinates of a square-shape farm are (1, 1), (3, 1), (3, 3), and (1, 3), with the refinery located at (0, 0) in the Cartesian coordinates (see Fig. 25.8a). The same μ, τ and ω as in Example 1 (Sect. 25.2.1) are assumed. Applying the conversion in Eqs. 25.13 and 25.14, the corresponding vertex coordinates in the polar coordinates are obtained: $\left(\sqrt{2},\frac{\pi}{4}\right)$, $\left(\sqrt{10},\cot^{-1}3\right)$, $\left(3\sqrt{2},\frac{\pi}{4}\right)$, and $\left(\sqrt{10},\tan^{-1}3\right)$; the $x=1$, $x=3$, $y=1$, and $y=3$ lines become

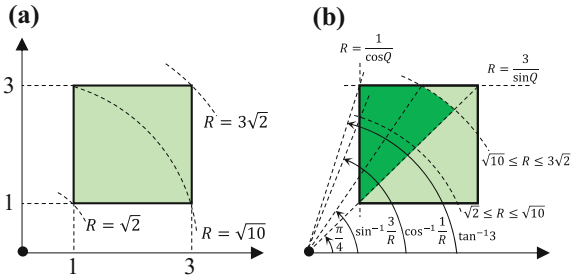

Fig. 25.8 Example 2: square farm. **a** Farm layout; **b** calculation demonstration

$R = \frac{1}{\cos Q}$, $R = \frac{3}{\cos Q}$, $R = \frac{1}{\sin Q}$, and $R = \frac{3}{\sin Q}$, respectively, in the polar coordinate system. Thus, the farm is described with the following constraints:

$$\begin{aligned} \frac{1}{\cos Q} \le R \le \frac{3}{\cos Q} \\ \frac{1}{\sin Q} \le R \le \frac{3}{\sin Q} \end{aligned} \tag{25.21}$$

Based on Eq. 25.7, and the symmetry of the farm, M is expressed as follows (see Fig. 25.8b):

$$M = \begin{cases} 2\mu \int_{\frac{\pi}{4}}^{\cos^{-1}\frac{1}{R}} \int_{\frac{1}{\cos Q}}^{R} R dR dQ, & \sqrt{2} \le R \le \sqrt{10} \\ 2\mu \left(\int_{\sin^{-1}\frac{3}{R}}^{\tan^{-1}3} \int_{\frac{1}{\cos Q}}^{\frac{3}{\sin Q}} R dR dQ + \int_{\frac{\pi}{4}}^{\sin^{-1}\frac{3}{R}} \int_{\frac{1}{\cos Q}}^{R} R dR dQ \right), & \sqrt{10} \le R \le 3\sqrt{2} \end{cases} \tag{25.22}$$

Obtaining the algebraic expression for $R = r(M)$ and further $\widehat{C}^{\text{tran}} = c^{\text{tran}}(M)$ from Eq. 25.22 is challenging. Therefore, we apply the numerical solution method discussed previously, and the $M - \widehat{C}^{\text{tran}}$ curve is generated as shown in Fig. 25.9, which is plotted with 400 discrete points. The curve function can be expressed as a

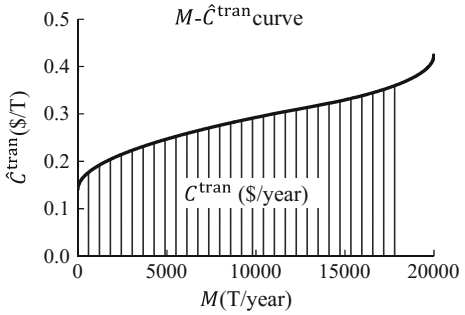

Fig. 25.9 $M - \widehat{C}^{\text{tran}}$ curve for the square farm in Example 2, plotted with 400 points

fitted function (piece-wise linear function, polynomial function, etc.), based on these points. Note that obtaining the algebraic expression is challenging even for a simple square farm with symmetry in this example. Therefore, the numerical solution method is likely to be used for most practical cases.

25.3 Composite Curve-Based Procurement Planning Optimization

In general, we assume a refinery to be surrounded by multiple farms $i \in \mathbf{I}$, with unique shapes, locations, biomass productivities, tortuosity factors, and unit biomass purchasing and harvesting costs (see Fig. 25.1d). A given total biomass demand β must be supplied from the farms to the refinery to satisfy biofuel production requirement. The goal is to minimize the total procurement cost (including biomass purchasing, harvesting, and transportation). The corresponding procurement plan includes the mass to transport from each farm, and the selected areas. The optimization problem can then be stated as follows:

$$\min \sum_i C_i^{\text{supp}} \tag{25.23}$$

$$\sum_i M_i = \beta \tag{25.24}$$

$$\widehat{C}_i^{\text{supp}} = \hat{c}_i^{\text{supp}}(M_i) = \hat{c}_i^{\text{tran}}(M_i) + \varepsilon_i \tag{25.25}$$

$$C_i^{\text{supp}} = \int_0^{M_i} \hat{c}_i^{\text{supp}}(M_i) dM_i \tag{25.26}$$

where M_i denotes the mass transported from farm i; C_i^{supp} denotes the procurement cost for farm i; $\widehat{C}_i^{\text{supp}}/\widehat{C}_i^{\text{tran}}$ denotes the marginal per-mass supply/transportation cost for the infinitesimal mass for farm i (\$/T); and ε_i denotes the unit biomass purchasing and harvesting cost for farm i (\$/T). Equation 25.24 specifies the total mass demand; Eq. 25.25 correlates $\widehat{C}_i^{\text{supp}}$ with $\widehat{C}_i^{\text{tran}}$ and ε_i; and Eq. 25.26 calculates the procurement cost C_i^{supp}. Based on Eq. 25.25, the $M_i - \widehat{C}_i^{\text{supp}}$ curve (referred to as biomass supply "$M - C$ curve") is obtained by shifting up the $M_i - \widehat{C}_i^{\text{tran}}$ curve (see Fig. 25.3b) in Sect. 25.2, by ε_i (Fig. 25.10).

Note that in the subsequent presentation, the index i is only used when the variables need to be distinguished among farms. Otherwise, it is omitted.

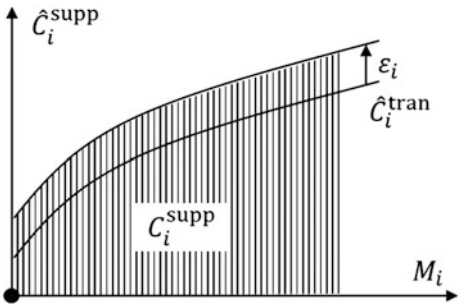

Fig. 25.10 $M_i - \widehat{C}_i^{\text{supp}}$ curve, obtained by shifting the $M_i - \widehat{C}_i^{\text{tran}}$ curve up by ε_i

25.3.1 Generation of the Composite Curve and Identification of the Optimum Procurement Plan

In the pinch design method for heat exchanger networks (Linnhoff and Hindmarsh 1983), the $H - T$ curves of all the hot (or cold) streams are combined to generate a hot (or cold) composite curve. Then, the identification of the minimum heating load is achieved using the composite curve. Using the method in Sect. 25.2, and the modification shown in Fig. 25.10, we can generate an $M - C$ curve for each farm. How do we utilize them to identify the optimum procurement plan? Our goal is to combine the individual $M - C$ curves into a composite curve to account for the procurement of the entire system.

The composite curve is generated in the following four steps (Fig. 25.11):

(1) *Step 1*. Plot the shifted $M - C$ curves for each farm on a single diagram so that the maximum M value of the previous curve is aligned with the zero M value of the next curve (Fig. 25.11a). The x-axis (corresponding to M) is relative, while the y-axis (corresponding to $\widehat{C}^{\text{supp}}$) is absolute. Thus, one curve begins where the previous ends, and the mass transported from each farm is represented by the mass difference, instead of the absolute mass on the X-axis.

(2) *Step 2*. Divide the y-axis into $\widehat{C}^{\text{supp}}$ intervals (defined by the horizontal dashed lines in Fig. 25.11b), where the interval boundaries mark the lower and upper $\widehat{C}^{\text{supp}}$ values for each curve.

(3) *Step 3*. In each $\widehat{C}^{\text{supp}}$ interval, draw a new curve that represents the individual curves present in the interval. To illustrate, we consider the individual curves (corresponding to farms f, g…) are expressed by functions f, g…, respectively (e.g. as shown in Fig. 25.11c). The function for the composite curve in the current interval is denoted as h. Then, for any given M, the following equation must hold:

$$M - m_{\text{f0}} = \left\{f^{-1}[h(M)] - m_{\text{f0}}\right\} + \left\{g^{-1}[h(M)] - m_{\text{g0}}\right\} + \cdots \tag{25.27}$$

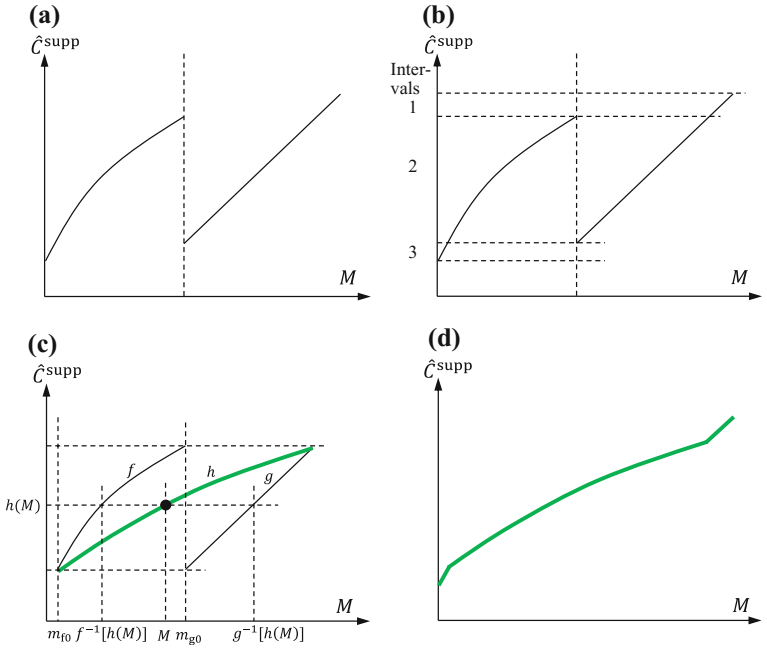

Fig. 25.11 Composite curve generation. **a** Step 1: shifting individual curves; **b** Step 2: dividing $\widehat{C}^{\text{supp}}$ into intervals; **c** Step 3: generating a composite representation in each interval (in this case, only interval 2 is demonstrated); **d** Step 4: generating the final composite curve. The composite curve is marked with *thick curves*

where, $m_{f0}, m_{g0} \ldots$ represent the starting mass values of the shifted curves in the current interval; $f^{-1}, g^{-1} \ldots$ represent the inverse function of $f, g \ldots$. Thus, the composite curve is generated by finding the function h that holds Eq. 25.27 true for all M, which can be implemented numerically.

(4) *Step 4*. Generate the final composite curve by eliminating the original curves, as shown in Fig. 25.11d. Note that since $\widehat{C}^{\text{supp}}$ on the individual curves monotonically increases with M, the same is true for the composite curve.

The optimum procurement plan can be identified from the generated composite curve once a total demand β is specified. For instance, as shown in Fig. 25.12, the $\widehat{C}^{\text{supp}}$ corresponding to β on the composite curve uniquely determines the optimum mass transported from farms f and g to be M_1 and $M_2 - m_{g0}$, respectively. The minimum procurement cost is represented by the area below the composite curve, and between $M = 0$ and $M = \beta$, calculated from Eq. 25.26. The selected area (characterized by R) of each farm is then calculated using the $R_i = r(M_i)$ relation (discussed in Sect. 25.2).

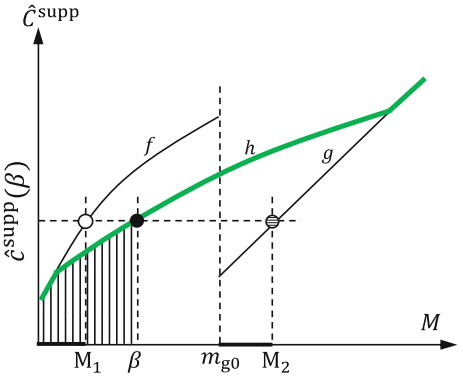

Fig. 25.12 Identification of the optimum procurement plan and the minimum procurement cost

25.3.2 Remarks

Note that $\frac{dM}{d\widehat{C}^{\text{supp}}}$ of our $M-C$ curves (i.e. one divided by the slope of the curves) reflects the mass that can be delivered at a given cost, which we denote as $\alpha\left[\mathrm{T}^2/(\mathrm{yr}\cdot\$)\right]$. Specifically, based on Eqs. 25.3 and 25.25, α can be generally expressed as follows:

$$\alpha=\frac{dM}{d\widehat{C}^{\text{supp}}}=\frac{\mu d\int\int_D RdRdQ}{\omega\tau dR} \tag{25.28}$$

For instance, α for the ring-shape farm in Sect. 25.2.1, is expressed as:

$$\alpha=\frac{\mu d\int_{q_1}^{q_2}\int_{r_1}^{R}RdRdQ}{\omega\tau dR}=\frac{\mu(q_2-q_1)R}{\omega\tau} \tag{25.29}$$

Equation 25.29 suggests that the following factors will increase α:

(i) lower unit transportation cost ω and tortuosity factor τ;
(ii) larger biomass productivity μ;
(iii) larger $(q_2-q_1)R$, which represents the length of the arc (of radius R) in the farm region.

Furthermore, in addition to the α indicator, the actual procurement preferences among farms also depend on the exact location, shape, and biomass purchasing and harvesting cost. These additional characteristics are mainly captured by the $\widehat{C}^{\text{supp}}$ value at zero mass on the $M-C$ curve. A general expression can be deducted by setting $M=0$ in Eqs. 25.5 and 25.25:

$$\left[\widehat{C}^{\text{supp}}\right]_{M=0}=\omega\tau r_0+\varepsilon \tag{25.30}$$

where r_0 represents the closest point on the farm to the refinery (which corresponds to $M = 0$). Therefore, the smaller r_0 and ϵ are, the smaller the $\widehat{C}^{\text{supp}}$ value at zero mass is, and thus the more preference is placed toward this farm. The composite curve depicts a combined representation of such preferences, for all the farms.

25.4 Case Study

We consider a system with 12 farms surrounding the refinery located at the origin (see Fig. 25.13). Each farm is depicted as a convex polytope with unique biomass productivity μ, tortuosity factor τ, unit biomass purchasing and harvesting cost ϵ, and vertex coordinates (given in Tables 25.1 and 25.2). A unit transportation cost $\omega = 0.2\,\$/(\text{km} \cdot \text{T})$ is assumed. From the vertex coordinates, the constraints describing the farms, and the functions of all the edges of the farm are determined.

Using the numerical method presented in Sect. 25.2 (see Fig. 25.5a) and the curve shifting discussed in Fig. 25.10 of Sect. 25.3, the individual $M - C$ curves for all the 12 farms are generated. Then, with the method presented in Sect. 25.3.1, the composite curve is generated. The individual and composite curves are shown in Fig. 25.14.

First, we consider two cases where the total biomass demand β is assumed to be 11800 T/year and 13300 T/year, respectively. After the generation of the composite curve, the optimum procurement plan is identified (corresponding to the dashed lines in Fig. 25.14) using the method discussed in Sect. 25.3.1. The specific procurement plan, including the mass transported, the procurement cost, and the selected area (characterized by radius R) for each farm, are given in Table 25.3. The graphic representation of the procurement plan is shown in Fig. 25.15. The total

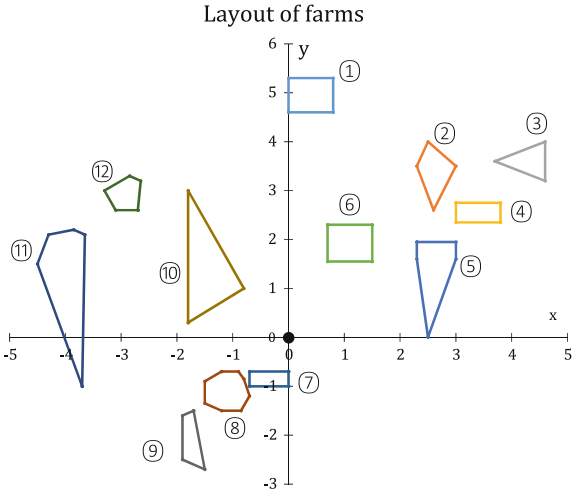

Fig. 25.13 Layout of farms for the case study. The *dot* at the origin represents the refinery. The farm numbers are labeled accordingly

Table 25.1 Parameters for farms 1–6

	Farm 1	Farm 2	Farm 3	Farm 4	Farm 5	Farm 6
μ (T/(km^2 · year))	5000	1700	4000	3000	2000	1700
T	1.5	5	3	4	2	3
ε ($/T)	18	16	17	17	18	17
Coordinates (km)	(0, 4.6)	(2.3, 3.5)	(3.7, 3.6)	(3, 2.75)	(2.3, 1.95)	(0.7, 2.3)
	(0, 5.3)	(2.5, 4)	(4.6, 4)	(3.8, 2.75)	(3, 1.95)	(1.5, 2.3)
	(0.8, 5.3)	(3, 3.5)	(4.6, 3.2)	(3.8, 2.35)	(3, 1.6)	(1.5, 1.55)
	(0.8, 4.6)	(2.6, 2.6)		(3, 2.35)	(2.5, 0)	(0.7, 1.55)
					(2.3, 1.6)	

Table 25.2 Parameters for farms 7–12

	Farm 7	Farm 8	Farm 9	Farm 10	Farm 11	Farm 12
μ (T/(km^2 · year))	4500	4300	5000	3500	3500	5500
T	4	5	3	4	3	3
ε ($/T)	16	16	17	18	16	17
Coordinates (km)	(0, −0.7)	(−1.2, −0.7)	(−1.7, −1.5)	(−1.8, 0.3)	(−3.7, −1)	(−2.7, 2.6)
	(0, −1)	(−0.9, −0.7)	(−1.5, −2.7)	(−1.8, 3)	(−3.65, 2.1)	(−2.65, 3.2)
	(−0.7, −1)	(−0.8, −0.85)	(−1.9, −2.5)	(−0.8, 1)	(−3.85, 2.2)	(−2.85, 3.3)
	(−0.7, −0.7)	(−0.7, −1.2)	(−1.9, −1.6)		(−4.3, 2.1)	(−3.3, 3)
		(−0.85, −1.5)			(−4.5, 1.5)	(−3.1, 2.6)
		(−1.2, −1.5)				
		(−1.5, −1.35)				
		(−1.5, −0.9)				

procurement cost for the 11800 T/year demand is 215444 $/year, rendering a unit cost of 18.26 $/T; the total procurement cost for the 13300 T/year demand is 244311 $/year, rendering a unit cost of 18.37 $/T. Note that the unit cost has increased with the increase in demand because the composite curve monotonically increases, and thus the unit cost $\frac{C^{\text{supp}}}{M}$ in the $0 \le M \le 11800$ range is smaller than that in the $11800 \le M \le 13300$ range.

Note that, to satisfy the biomass demand of 11800 T/year, all the mass in Farms 6, 7, 8, 9 and 11, and a fraction of the mass in Farms 5 and 10 are transported to the refinery. The 13300 T/year demand is not a significant increase in comparison. Farms 5 and 10 contain enough extra mass to satisfy the incremental demand of 1500 T/year. However, the optimum procurement plan shown in Fig. 25.15b

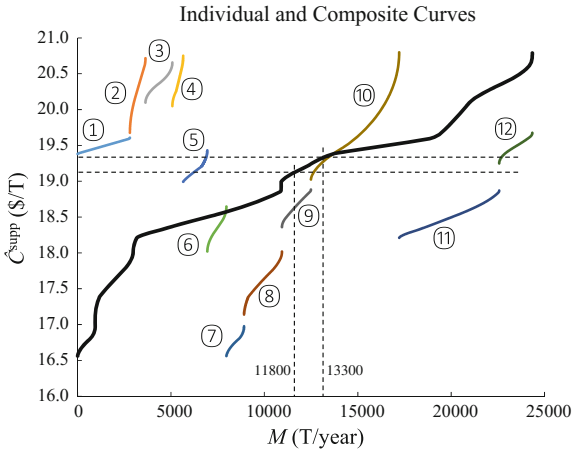

Fig. 25.14 Individual $M - C$ curves and the composite curve for the case study. The composite curve is marked *thick*. The demands, 11800 and 13300 T/year, and the corresponding $\widehat{C}^{supp}$ are marked with *dashed lines*. The corresponding farm numbers are labeled accordingly

Table 25.3 Supply strategies for the 11800 T/year and 13300 T/year demands. M denotes the mass transported from the corresponding farm, not the absolute value on the $M - C$ diagram. For the farms with zero transported mass, the corresponding R indicates the distance from the closest point on the farm to the refinery

	β = 11800 T/year			β = 13300 T/year		
	R (km)	M (T/year)	C^{supp} ($/year)	R (km)	M (T/year)	C^{supp} ($/year)
Farm 1	4.60	0	0	4.60	0	0
Farm 2	3.68	0	0	3.68	0	0
Farm 3	5.16	0	0	5.16	0	0
Farm 4	3.81	0	0	3.81	0	0
Farm 5	2.87	689	13141	3.37	1243	23786
Farm 6	2.75	1020	18704	2.75	1020	18704
Farm 7	1.22	945	15829	1.22	945	15829
Farm 8	2.02	2030	35760	2.02	2030	35760
Farm 9	3.14	1550	28913	3.14	1550	28913
Farm 10	1.44	212	4049	1.68	1000	19226
Farm 11	4.79	5354	99048	4.79	5354	99048
Farm 12	3.75	0	0	3.91	158	3045
Total	NA	11800	215444	NA	13300	244311

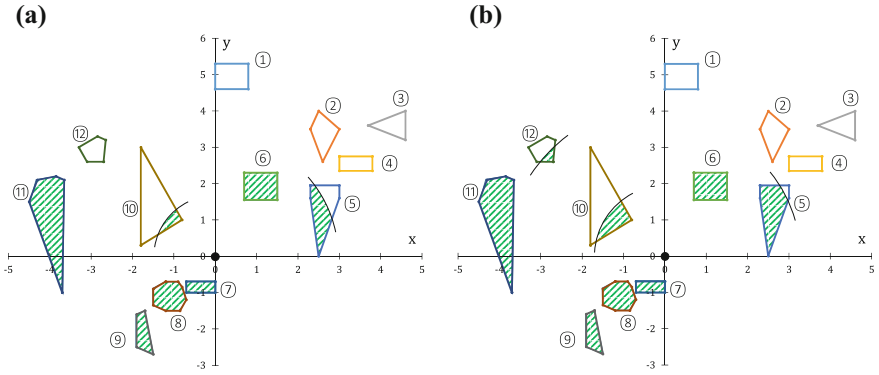

Fig. 25.15 Graphic representation of the supply strategies for the demand of 11800 T/year (**a**) and 13300 T/year (**b**), respectively. The corresponding farm numbers are labeled accordingly

suggests that we switch to Farm 12 after the radii R of Farms 5 and 10 reach 3.37 km and 1.68 km, respectively, as shown in Table 25.3. Only an additional 554 T/year (a total of 1243 T/year) is procured from Farm 5; only an additional 788 T/year (a total of 1000 T/year) is procured from Farm 10; the remaining 158 T/year is procured from Farm 12. This is because the α values for Farms 5 and 10 (see the corresponding $M - C$ curves in Fig. 25.14) are not large enough, and their minimum $\widehat{C}^{\text{supp}}$ values corresponding to $M = 0$, are not small enough. Thus, $\widehat{C}^{\text{supp}}$ of Farm 5 at $M > 1243$, and that of Farm 10 at and $M > 1000$ become larger than the $\widehat{C}^{\text{supp}}$ of Farm 12 at $M = 0$. Therefore, 158 T/year biomass is selected from Farm 12 to satisfy the 1500 T/year additional demand, minimizing total procurement cost.

Next, we consider a total demand of 11800 T/year biomass, but a new refinery location is assumed at (1, 1) in the layout map shown in Fig. 25.13. For simplicity, we change the system of reference such that the refinery in this case is located at the origin. Thus, the coordinates of the farms in the new system are obtained by subtracting 1 from all x and y coordinates in Tables 25.1 and 25.2. Using the same methods, we obtain the corresponding individual and composite curves shown in Fig. 25.16a.

To satisfy the given 11800 T/year demand (see Fig. 25.16a), the optimum procurement plan is identified as shown in Fig. 25.16b and Table 25.4. The total procurement cost is 221246 $/year, rendering a unit cost of 18.75 $/year, which is higher than the cost assuming the same demand, but with the original refinery location.

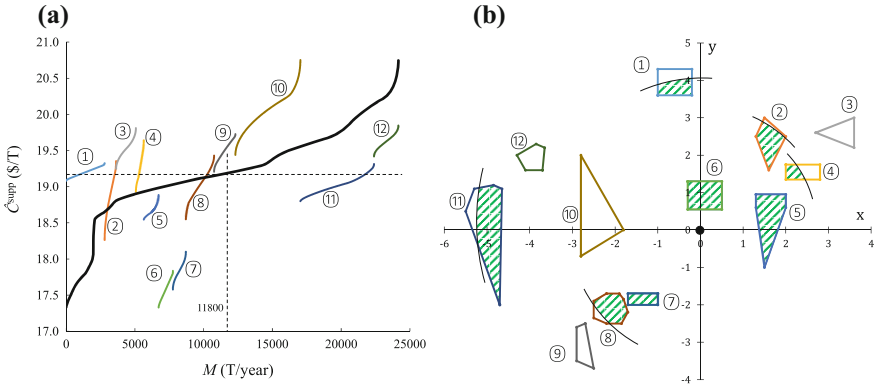

Fig. 25.16 Individual $M-C$ curves and the composite curve, and the corresponding optimum procurement plan for the 11800 T/year demand, with the new refinery location. **a** Individual $M-C$ curves and the composite curve; the composite curve is marked thick; **b** graphic representation of the procurement plan. The corresponding farm numbers are labeled accordingly

Table 25.4 Procurement plan for the 11800 T/year demand, with the new refinery location. The notes in Table 25.3 applies here

	R (km)	M (T/year)	C^{supp} ($/year)		R (km)	M (T/year)	C^{supp} ($/year)
Farm 1	3.97	1269	24300	Farm 7	2.62	945	16856
Farm 2	3.19	751	14138	Farm 8	3.19	1620	30697
Farm 3	3.75	0	0	Farm 9	3.68	0	0
Farm 4	2.74	245	4678	Farm 10	1.80	0	0
Farm 5	2.21	1078	20114	Farm 11	5.32	4854	92197
Farm 6	1.39	1039	18266	Farm 12	4.03	0	0

What is the qualitative explanation to the procurement plan changes in response to the refinery location change? Comparing Figs. 25.15a and 25.16b, we see that less mass is transported from Farms 8, 9, 10 and 11, while more mass is transported from Farms 1, 2, 4 and 5. Overall, the selected areas have shifted from the bottom left to the top right on the map. Note that when the refinery location shifts from (0, 0) to (1, 1) on Fig. 25.15a, Farms 1-6 become closer to the refinery, while Farms 7–12 move further away. As a result, as shown in Fig. 25.17, the individual $M-C$ curves for Farms 1–6 shift down, while the curves for Farms 7–12 shift up, indicating that Farms 1–6 are more preferable in the new system. Therefore, to maintain the 11800 T/year total demand, more mass is procured from Farms 1–6, and less is procured from Farms 7–12.

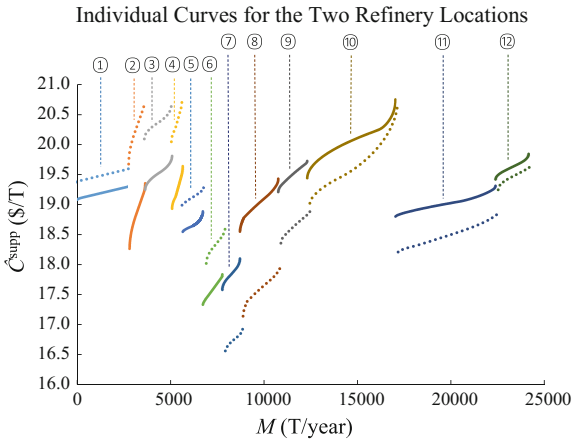

Fig. 25.17 Individual $M - C$ curves corresponding to the two refinery locations. Curves corresponding to the original location is represented by *dotted curves*, while curves corresponding to the new location is represented by *solid curves*. The corresponding farm numbers are marked accordingly at the top

25.5 Summary

In this chapter, we discussed a new method for biomass procurement planning. We first developed a new modeling method for the region-to-point transportation problem, based on mathematical integration. Specifically, we presented methods for deriving an $M - \widehat{C}^{tran}$ curve for a single farm. Then, we discussed how the $M - \widehat{C}^{tran}$ curves can be effectively incorporated into a general multi-farm procurement planning approach. To account for biomass purchasing and harvesting cost, $\widehat{C}^{tran}$ is first modified to $\widehat{C}^{supp}$. Then the individual $M - C$ curves for all farms are "composed" together to generate a composite curve. Using this curve, the minimum procurement cost together with the optimal procurement plan are identified for a given total biomass demand.

The proposed composite-curve-based method allows us to more accurately calculate transportation distance, and thus transportation costs and GHG emissions due to transportation. It also provides some key insights into the design of biofuel SCs. Future work will extended the method to handle multi-feedstock and multi-period planning problems, as well as systems with contracts, storage facilities, and integration with mathematical programming approaches.

Acknowledgments This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494).

References

Akgul, O., Shah, N., & Papageorgiou, L. G. (2012). Economic optimisation of a UK advanced biofuel supply chain. *Biomass and Bioenergy, 41*, 57–72.

Andersen, F. E., Diaz, M. S., & Grossmann, I. E. (2013). Multiscale strategic planning model for the design of integrated ethanol and gasoline supply chain. *AIChE Journal, 59*(12), 4655–4672.

An, H., Wilhelm, W. E., & Searcy, S. W. (2011). A mathematical model to design a lignocellulosic biofuel supply chain system with a case study based on a region in Central Texas. *Bioresource technology, 102*(17), 7860–7870.

Brown, R. C., & Brown, T. R. (2012). *Why are we producing biofuels*. Ames, Iowa: Brownia LLC.

DOE Bioenergy Technologies Office. (2014). *Multi-year program plan*. Washington DC, USA: DOE.

DOE/EERE (2013). *Feedstock supply and logistics: biomass as a commodity,* US: Department of Energy, Office of Energy Efficiency and Renewable Energy

Elia, J. A., Baliban, R. C., Xiao, X., & Floudas, C. A. (2011). Optimal energy supply network determination and life cycle analysis for hybrid coal, biomass, and natural gas to liquid (CBGTL) plants using carbon-based hydrogen production. *Computer Chemical Engineering, 35*(8), 1399–1430.

Garcia, D. J., & You, F. (2015). Supply chain design and optimization: challenges and opportunities. *Computers and Chemical Engineering, 81*, 153–170.

Hamelinck, C. N., Suurs, R. A., & Faaij, A. P. (2005). International bioenergy transport costs and energy balance. *Biomass and Bioenergy, 29*(2), 114–134.

Josephine, E. A., & Floudas, C. A. (2014). Energy supply chain optimization of hybrid feedstock processes: a review. *Chemical and Biomolecular Engineering, 5*, 147–179.

Linnhoff, B., & Hindmarsh, E. (1983). The pinch design method for heat exchanger networks. *Chemical Engineering Science, 38*(5), 745–763.

Liu, P., Georgiadis, M. C., & Pistikopoulos, E. N. (2010). Advances in energy systems engineering. *Industrial and Engineering Chemistry Research, 50*(9), 4915–4926.

Marvin, W. A., et al. (2012a). Economic optimization of a lignocellulosic biomass-to-ethanol supply chain. *Chemical Engineering Science, 67*(1), 68–79.

Marvin, W. A., Schmidt, L. D., & Daoutidis, P. (2012b). Biorefinery location and technology selection through supply chain optimization. *Industiral and Engineering Chemistry Research, 52*(9), 3192–3208.

Miao, Z., et al. (2012). Lignocellulosic biomass feedstock transportation alternatives, logistics, equipment configurations, and modeling. *Biofuels, Bioproducts and Biorefining, 6*(3), 351–362.

Mosier, N., et al. (2005). Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology, 96*, 673–686.

Osmani, A., & Zhang, J. (2013). Stochastic optimization of a multi-feedstock lignocellulosic-based bioethanol supply chain under multiple uncertainties. *Energy, 59*, 172–457.

Overend, R. P. (1982). The average haul distance and transportation work factors for biomass delivered to a central plant. *Biomass, 2*, 75–79.

Santibanez-Aguilar, J. E., et al. (2011). Optimal planning of a biomass conversion system considering economic and environmental aspects. *Industrial and Engineering Chemistry Research, 50*(14), 8558–8570.

Sharma, B., Ingalls, R. G., Jones, C. L., & Khanchi, A. (2013). Biomass supply chain design and analysis: basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews, 24*, 608–627.

Sultana, A., & Kumar, A. (2014). Development of tortuosity factor for assessment of lignocellulosic biomass delivery cost to a biorefinery. *Applied Energy, 119*, 288–295.

Tong, K., Gong, J., Yue, D., & You, F. (2013). Stochastic programming approach to optimal design and operations of integrated hydrocarbon biofuel and petroleum supply chains. *ACS Sustainable Chemistry and Engineering, 2*(1), 49–61.

Uslu, A., Faaij, A. P., & Bergman, P. C. (2008). Pre-treatment technologies, and their effect on international bioenergy supply chain logistics. Techno-economic evaluation of torrefaction, fast pyrolysis and pelletisation. *Energy, 33*, 1206–1223.

You, F., Tao, L., Graziano, D., & Snyder, S. W. (2012). Optimal design of sustainable cellulosic biofuel supply chains: multiobjective optimization coupled with life cycle assessment and input-output analysis. *AIChE Journal, 58*, 1157–1180.

Yue, D., You, F., & Snyder, S. W. (2014). Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computer Chemical Engineering, 66*, 36–56.

Zhang, F., Johnson, D. M., & Sutherland, J. W. (2011). A GIS-based method for identifying the optimal location for a facility to convert forest biomass to biofuel. *Biomass and Bioenergy, 35*, 3951–3961.

Chapter 26
An Integrated Energy Storage System Based on Hydrogen Storage

Dan Gao, Dongfang Jiang and Naiqiang Zhang

Abstract The interconnection between a renewable power generation facility and a power grid poses challenges because of volatility and intermittent characteristics. Energy storage is one of the best solutions for this problem. The object of the present work is to evaluate the features and performances of energy storage system (ESS) with the aim to determine the best available ESS technology. For each one of the storage solutions presented, we have compared key parameters such as: efficiency, lifetime, energy density, capacity, and capital and response time. The paper presents an integrated ESS based on hydrogen storage, especially hydrogen energy technologies for hydrogen production, storage and utilization. Possibilities for integrated ESS coupled wind power to generate hydrogen using electrolyzer with hydrogen-oxygen combined cycle to generate power are discussed, wherein energy efficiency in the range of 49–55 % can be achieved. The results show that the proposed integrated system cannot be constrained by geological conditions and availability of materials, and appears to be an appropriate tool for the development of renewable power. Moreover, a case study is conducted for a special wind power plant. The integrated system is designed based on the daily wind load. Energy efficiency and preliminary economic comparison studies for the integrated system operated in two modes show that up to 50 % average net efficiency. Therefore, the integrated ESS can be useful to mitigate the bottleneck of renewable power development.

Keywords Integrated energy storage system · Wind power · Energy efficiency · Hydrogen-oxygen combined cycle

D. Gao (✉) · D. Jiang · N. Zhang
MOE's Key Lab of Condition Monitoring and Control for Power Plant Equipment,
School of Energy, Power and Mechanical Engineering,
North China Electric Power University, Beijing 102206, China
e-mail: gaodan@ncepu.edu.cn

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_26

26.1 Introduction

The development of renewable energy has experienced rapid growth over the past few years. In the next 50 years all sustainable energy systems will have to be based on the rational use of traditional resources and greater use of renewable energy. The rapid development of wind power as a representative of renewable energy eases the pressure on the traditional power grid and improves the energy structure. It not only brings environmental and economic benefits but also many uncertain challenges.

The traditional power grid has a significant influence on the current status and situation of the wind power connection to the grid (Kiviluoma and Meibom 2010). One of the main disadvantages of wind turbines, however, is that naturally variable wind speed causes power fluctuation problems at the load side, and wind power capacity is limited by low voltage ride through (LVRT) and often runs off-grid when the large grid fails. Besides, long-term storage of power generated by wind turbines is considered as a limitation on the part of remote locations. Moreover, a large number of wind farms were carelessly installed because of the impact of the local government's policy direction. These wind farms increased the pressure on the grid, and the "rubbish electricity" phenomenon occurred. To resolve the series of problems caused by the rapid development of wind power and bridge the huge gap present between energy supply and energy demand, the concept of energy storage has been gaining momentum in recent years. Energy storage technology is a system that equalizes electricity generation and load demand. The storage system operates to store energy during off-peak periods and runs the generator to provide stable power during on-peak periods. The energy storage system (ESS) was based on the integration of energy storage technology. ESS generally consists of two parts, energy storage devices and power conversion systems. A major goal of energy storage is to achieve the transformation of an energy medium for energy storage and release. ESS not only provide energy security with less environmental impact, but also improves operational performance of energy systems. Besides, it can meet the short-term fluctuating energy demand requirements and supply energy during transitory power disturbances or surges.

This chapter proposes an ESS based on the study of the current methods of energy storage technologies that employ hydrogen as an energy medium and that have sustainable development features. This study also designs a wind-hydrogen and hydrogen-oxygen combined cycle system, analyzes the wind load curve, and models the system.

26.2 Summary for ESS

Energy storage technology provides a simple solution to the balance of electricity supply and demand. The history of energy storage system began in the early 20th century with the emergence of a variety of systems with the capability to store

electrical energy in the form of charges and allowed to be discharged when the energy is needed. This was first achieved by application of lead–acid battery ESS in the power supply system of New York City. The energy was stored during the day, and then the generator was operated at night. Battery power was used for street lighting. The various storage methods were proposed as the technologies continue to grow. Energy storage technology has four types based on the different methods of energy conversion: mechanical (e.g. flywheel, compressed air energy storage and pumped hydroelectricity storage), electrical (e.g. capacitors, superconducting magnetic energy storage and super capacitors), thermal (e.g. low and high temperature energy storage systems) and chemical energy (e.g. electrochemical, thermochemical and chemical storage devices) technologies.

26.2.1 Common ESS

26.2.1.1 Compressed Air Energy Storage (CAES)

CAES is based on conventional gas turbine technology (Cavallo 2001). A simple CAES model is shown in Fig. 26.1. Energy is stored by compressing air in the pipe, solid airtight facilities, or an airtight underground storage cavern. The first commercially operated CAES was a 290 MW unit built in the Huntorf plant in Germany in 1978 (Díaz-González et al. 2012). The second one was a 110 MW unit that operated in Alabama. The unit could get into the grid within 14 min. These units provide valuable experience for the later development of a CAES power station. To improve efficiency and reduce operation costs, air leaks (self-discharge) must be kept to an absolute minimum. The use of CAES with air injection concept and supercharging with inlet chilling was discussed and analyzed (Jubeh and Najjar 2012), the result showed that the mean values of power generated, energy ratio, and primary efficiency of CAES-IC system are about 7 % higher, 3.3 % lower, and 1 % higher than the corresponding values with CAES-AI system. The hybrid wind-diesel system with CAES was also proposed (Ibrahim et al. 2011). Besides, Compressed air storage under high pressure in cylinders (up to 300 bars with carbon fiber structures) is a good solution for small and medium scale applications. Using an electric compressor that can be turned into a generator during retrieval, the system has an overall efficiency of 50 %.

26.2.1.2 Superconducting Magnetic Energy Storage (SMES)

In early 1911, the Dutch physicist discovered the superconductor in a trial process (Yekini Suberu et al. 2014). However, in the 1970s, scientists proposed SMES, energy is stored in the magnetic field of superconducting coil without resistive loss. The system operates by injecting a DC current into the superconducting coil, which stores the energy in magnetic field. When the energy is needed, the current is

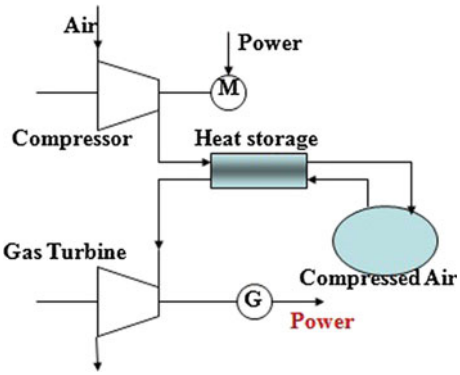

Fig. 26.1 A simple compressed air energy storage model

generated using the energy stored in the magnetic field. A simple SMES model is shown in Fig. 26.2. It became well known in ESS because of its good technical characteristics. The storage energy without loss in the long term and its conversion efficiency is much higher than that of other storage systems by up to 95 %. SMES has short response time in which it can provide high power in very short time, hence it is extensively used in wind power generation systems but not for providing long term energy storage (Etxeberria et al. 2010). Other applications of SMES include corrections of voltage sags and dips at industrial facilities (1 MW units) and stabilization of ring networks (2 MW units). At present, Low-temperature SMES (1 MJ/MW to 5 MJ/MW) have been produced, and 100 MJ SMES have been operated in high-voltage power grids. Another study (Trianni et al. 2013) identified and evaluated barriers to industrial energy efficiency through the investigation of 48

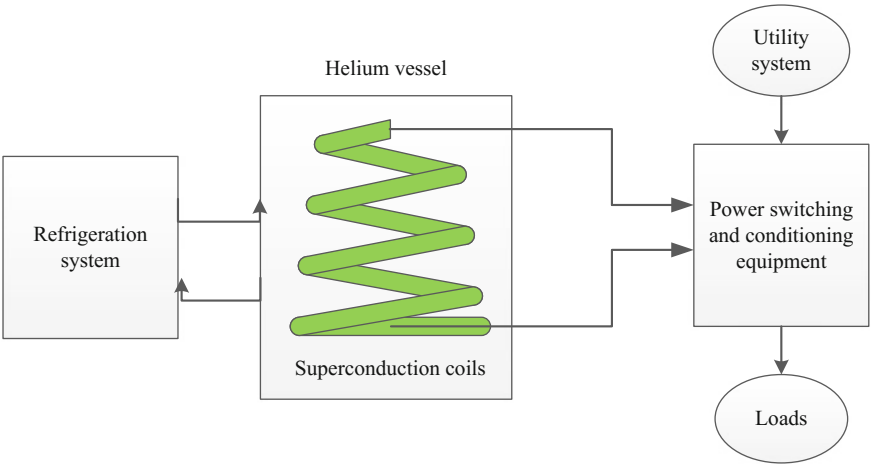

Fig. 26.2 A simple SMES model

manufacturing Small and Medium-sized Enterprises (SMES) in Northern Italy, the research provide interesting suggestions both for enterprises and energy policy-makers.

26.2.1.3 Battery Energy Storage System (BESS)

BESS uses chemical energy as intermediate energy media through reversible electrochemical reactions between two substances to store and release electrical energy. The system is composed of an A/D and D/A conversion system, battery pack, control equipment, and other ancillary equipment. Figure 26.3 illustrates a simple BESS. Lead batteries, nickel-cadmium batteries, hydride-nickel batteries, and sodium-sulfur (NaS) storage batteries have appeared in succession. A lead acid (LA) battery is the first kind of rechargeable battery in existence for both household and some major commercial applications. It is relatively cheap ($300–600/kWh), highly reliable and efficient (70–90 %). LA has a useful lifespan of approximately 5 years or 250–1000 charge/discharge cycles but depends on the depth of discharge (Ayodele and Ogunjuyigbe 2015). Ni-Cd which belongs to the family of rechargeable batteries has an effectively high energy density, good life cycle, sustainable efficiency, good system performance at low temperature, with characteristic wide range of sizes and ratings. Ni-Cd batteries are robust and proven substitute to lead-acid batteries. NaS battery is an inexpensive, high energy density, better efficiency, enhanced energy storage capacity device operating with active liquid substances. It is highly economical due to its affordable construction materials with the tendency for the materials to be recycled for reuse. This battery system is widely used in Japan and United States. Japan has installed a 500 kW/5MWh vanadium redox battery ESS with a net efficiency of up to 85 %. Regenesys Technologies in the UK adopted polysulfide–bromide batteries to build a 15 MW/120 MWh energy storage power station with a net efficiency of approximately 75 %. In China, the 10 kW vanadium

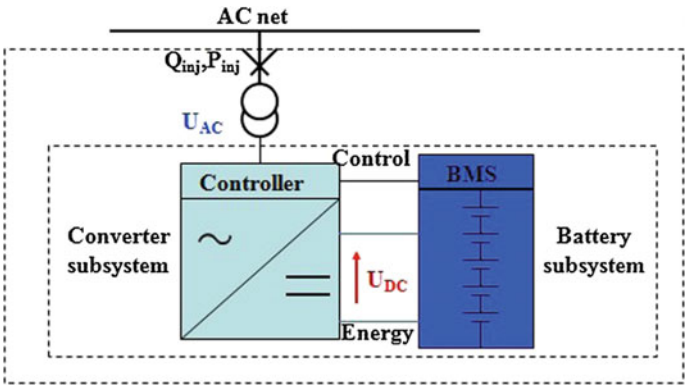

Fig. 26.3 A simple battery energy storage systems model

flow BESS was developed by the Dalian Institute of Chemical Physics of the CAS. The China Electric Power Research Institute exploited a 100 kW level ESS used in a wind power farm in 2008 and assessed the operation reliability and durability of the system (Gao et al. 2014). In July 2009, a hybrid ESS of photovoltaic (PV) power generation and 10 kW flow battery were built in the Vanguard village in Chongming Island in Shanghai. Currently, security, high cost, and capacity limitations are the bottlenecks of the BESS. System maintenance and material problems also limit the development of BESS.

26.2.1.4 Pumped Hydro Storage (PHS)

The pumped storage power station technology is a mature technology. It uses the power of water, a highly concentrated renewable energy source. This technology is currently the most used for high power applications. Figure 26.4 shows the geographical restrictions of a pumped storage power station. It is composed of reversible pump water turbines and upper and lower reservoirs. During periods when demand is low, these stations use electricity to pump the water from the lower reservoir to the upper reservoir. When demand is very high, the water flows out of the upper reservoir and activates the turbines to generate high value electricity for peak hours. Among ESS, the pumped storage power station is the most sophisticated and most extensively used storage method and system. At present, the system is extensively used in peak load shifting, and emergency backup. The system can respond quickly in a few minutes to maintain the stability of the grid. The conversion efficiency is about 65–80 %, mainly depending on equipment characteristics (Ibrahim et al. 2008). The storage capacity depends on the height between the upper reservoir and lower reservoir and the volume of water. However, for wind farms, solar PV power generation field, and other new energy producing areas, the

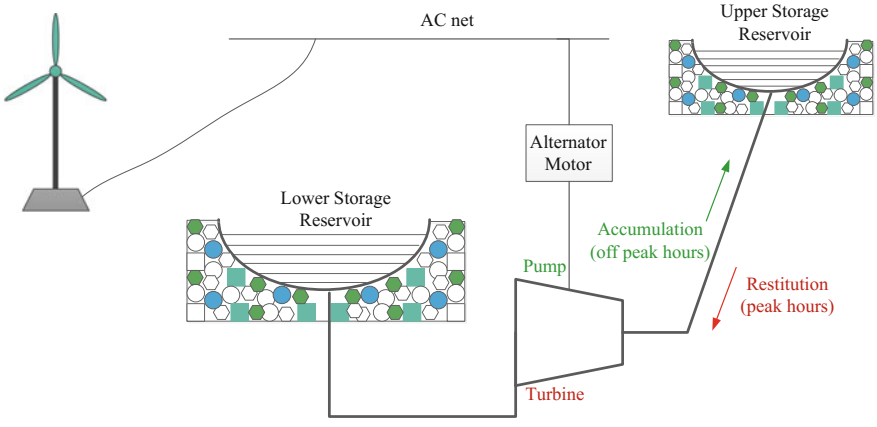

Fig. 26.4 A simple pumped storage power station model

majority of new energy power plants do not have the conditions to establish a pumped storage power station because of the limits of geographical conditions.

26.2.1.5 Flywheel Energy Storage (FES)

Flywheels have been used to store kinetic energy for thousands of years. A mass rotates on two magnetic bearings in order to decrease friction at high speed, coupled with an electric machine. The entire structure is placed in a vacuum to reduce wind shear. Figure 26.5 shows the simple mode of FES. Energy is transferred to the flywheel when the machine operates as motor, charging the energy storage device. The FES is discharged when the electric machine regenerates through the drive. The kinetic energy stored in a flywheel is proportional to the mass and the square of its rotating speed, the maximum energy storage density depends on the tensile strength of the flywheel material. Two types of flywheel are available in the energy market, the conventional steel rotor flywheel intended for low speed cycling (<6000 rpm) and the advanced composite flywheel for high speed operations (10^{-4}–10^{-5} rpm). Conventional low speed flywheels are suitable for the uninterruptable power supply. To store energy in an electrical power system, high-capacity flywheels are needed. Friction losses of a 200-tons flywheel are estimated at about 200 kW. Using this hypothesis and instantaneous efficiency of 85 %, the overall efficiency would drop to 45 % after 24 h (Becherif et al. 2015). From a practical point of view, electromechanical batteries are more useful for the production of energy in isolated areas. For example, some systems have been installed to supply areas of

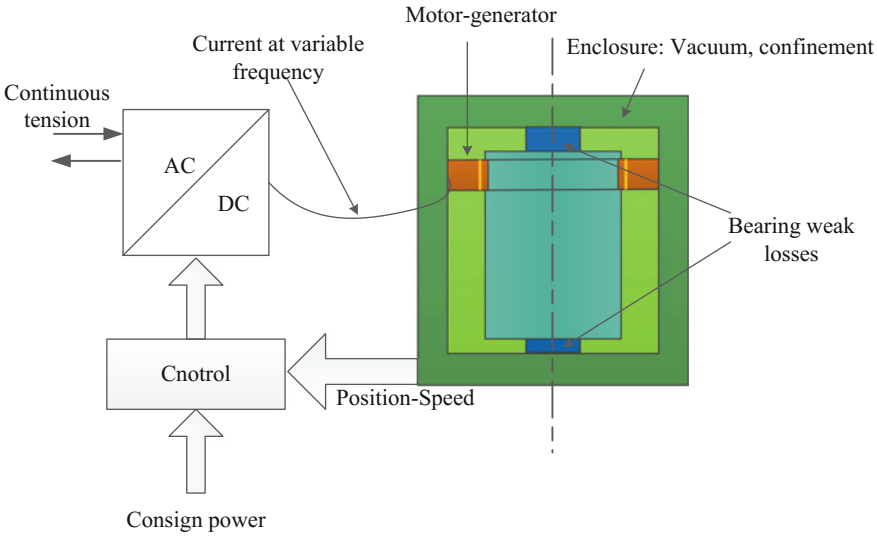

Fig. 26.5 A simple flywheel energy storage model

scattered houses, as well as the islands of Scotland and Wales (Ulleberg et al. 2010). In recent years, the development of both long- and short-term flywheel energy storage systems have become increasingly popular. The reliability or cyclic ability are very high and the life span between 15 and 20 years. Besides, the high power density and low-cost maintenance are the advantages of FES. Of course, there are some disadvantages, including: high rate of self-discharging over time, high materials cost of FES. Thus, flywheel systems can find suitability mostly to short-term energy storage applications requiring instant power delivery for meeting frequent load fluctuations as well as in power conditioning sectors.

26.2.1.6 Other Energy Storage Systems

Capacitors store energy as electric charge between two plates metal or conductive separated by an insulating material known as a dielectric when a voltage differential is applied across the plates. Figure 26.6 shows the model of capacitors (Kousksou et al. 2014). The instantaneous recharge capability and long cycle lifetime can be regarded as the major advantages of capacitors, which are highly suitable for utility small-scale power control applications. However, the low energy density limits the capacitor's suitability for large-scale applications because it requires a larger area of dielectric medium, which is unrealistic and expensive. Supercapacitors have very high power output and energy storage systems now under trial reach approximately 50–100 kW. In most applications, the energy stored will supply the load only for a few seconds to minutes. The number of charge and discharge cycles is for all

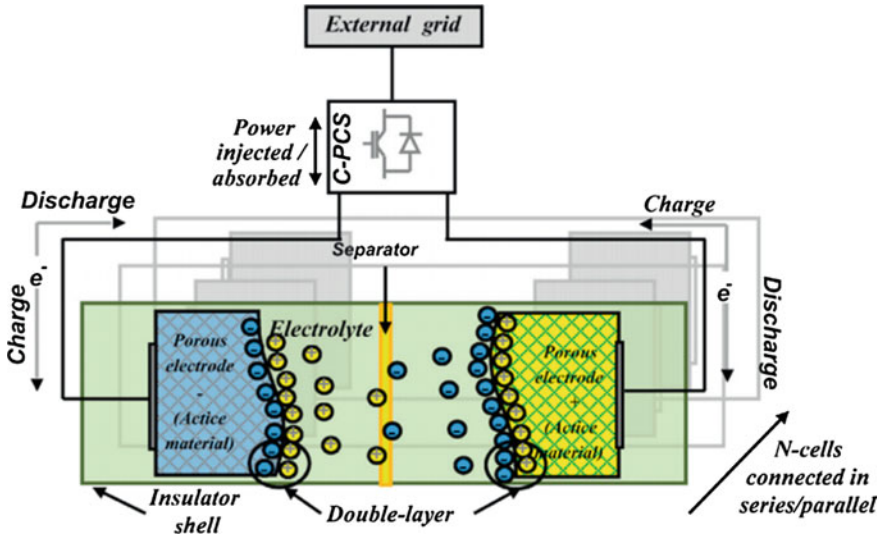

Fig. 26.6 The capacitor storage system

practical purposes nearly unlimited but the energy throughput in fast cyclic operation is limited. Control circuit to balance the individual voltages of each supercapacitor is necessary for safe and reliable operation if supercapacitors are connected in series to achieve a high output voltage. Supercapacitors generally are very durable, that is to say 8–10 years, 95 % efficiency, and 5 % per day self-discharge, which means that the stored energy must be used quickly.

The concepts of storing thermal energy in the form of sensible heat, latent heat, and reversible thermochemical reactions have been put into practice over years for achieving energy redistribution and energy efficiency on short-term or diurnal and long-term basis. The first is low temperature thermal storage, The heat energy can be stored and retrieved using a heat storage material like chilled water, phase change materials, or ice-thermal storage options. The other method is medium temperature thermal storage system, which utilizes heat storage materials whose operating temperature is generally higher than that of the human comfort range, In the case of solar water heating or air heating applications, the heat storage materials would function at an elevated temperature for storing the requisite heat energy.

A series of storage means are also used in power systems. A study (Ferreira et al. 2013) provides a qualitative methodology to select the appropriate technology or mix of technologies for different applications of energy storage. The multiple comparisons based on different characteristics distinguish this paper from others about ESS. However, considering material, capacity, geographical, and economic constraints, ESS cannot avoid constraints in practice. Therefore, a new ESS that can match wind power, solar power, and other new energy generating units must be developed. The multiregional, multi-capacity building as well as the rapid charge and discharge state must be supported. The safety and stability of network systems must also be ensured. An energy self-efficient building using integrated renewable energy was proposed (Marino et al. 2013), with two different configurations: one with solar PV and the other with combined solar PV and wind power.

26.2.1.7 Evaluation of Energy Storage Systems

Each storage technology has unique characteristics and is different in terms of its appropriate application field and energy storage scale. A comprehensive analysis of each storage technology needs to be performed before a decision can be made about the storage technology that is most suitable. Figure 26.7 shows the comparison of efficiency and lifetime. The highest efficiency of ESS is SMES, FES and Li-ion show the higher efficiency than other ESS. Energy storage systems or devices functioning on a continuous or partial basis have inherent process energy losses or from the components associated to the system. The efficiency of the storage system can be enhanced by regulating the power-transfer chain losses during charging and discharging processes. Mechanical energy storage system including PHS and CAES have a long cycle life, the lifetime of BESS are not as high as other systems owing to chemical deterioration with the operating time.

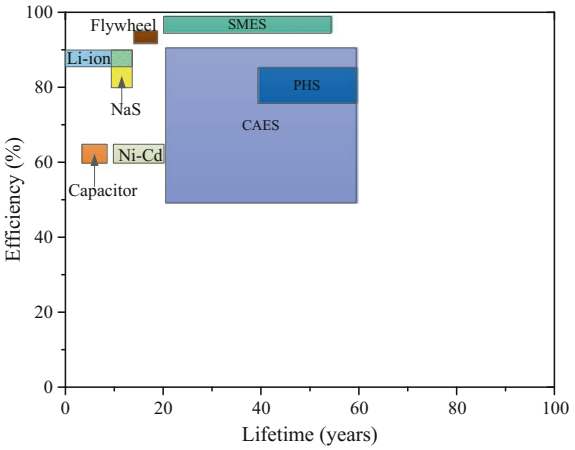

Fig. 26.7 Comparison of efficiency and lifetime of ESS technologies

Figure 26.8 shows the comparison of energy density and capacity. It can be seen PHS has the largest capacity among the ESS, but its energy density is the lowest, the operation experience shows a mass of 1 ton falling 100 m only generates 0.272 kWh.

Figure 26.9 shows the comparison of capital and response time. It can be seen the capital cost per kW of SMES, FES and Capacitor are lower than other energy storage technologies. SMES has the lowest capital cost per kW among the developed technologies.

The advantages and the challenges of various storage technologies are presented in Table 26.1 (Ayodele and Ogunjuyigbe 2015) and they are summarily based on the following criteria: ramp rate of the technology, response delay time, duration of storage, maturity of technology, installation cost, efficiency of the technology, environmental impact, human health implication, complexity of the charging devices and the suitability of the site topology. It can be observed from the table no single storage technology could provide high energy and power capacity, low capital cost, long lifetime and high response time.

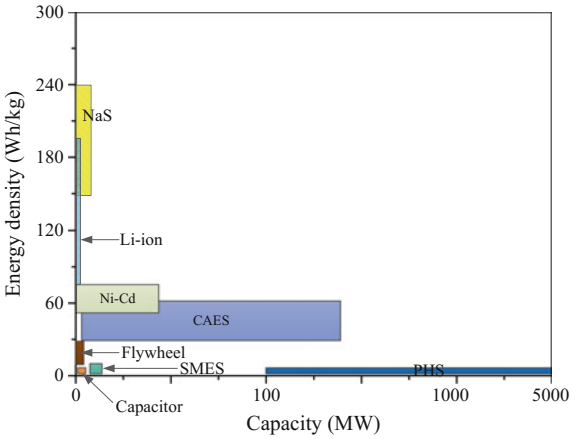

Fig. 26.8 Comparison of energy density and capacity of ESS technologies

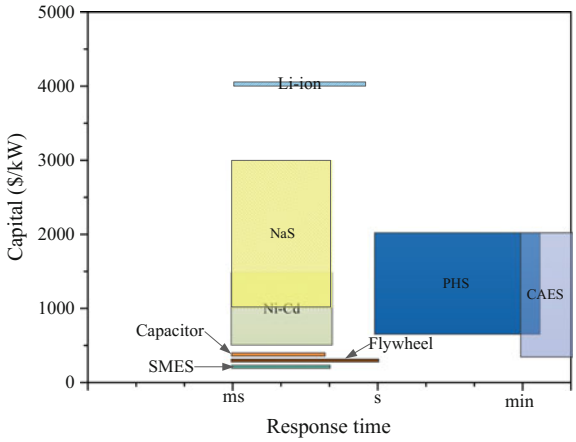

Fig. 26.9 Comparison of capital and response time of ESS technologies

Table 26.1 Advantages of storage technologies and their challenges

Storage technology	Advantages	Challenges
CAES	High energy and power capacity, moderate capital cost, moderate long storage duration, quick start up, moderate efficiency, long life	Adverse environmental impact, requires burning of fossil fuels, difficult to site
SMES	High energy and power capacity, short response time, long life time and very high efficiency	Requires to be kept at low temperature, low energy density, high capital cost, potential health issue, discharges quickly
NaS battery	High energy and power capacity, good response time, no self-discharge, high efficiency, long life time, minimum maintenance	High production cost, concerns for safety, operate with in specific temperature
Ni-Cd battery	High energy density, long life, tolerant to temperature and deep discharge	Environmental concern, high capital cost, memory effect, need advance charging and discharging monitoring device
Li-ion battery	Short access time, high energy density, high efficiency, long life, possible in small devices	High cost and technical issues, requires ideal charging condition
PHS	Moderate efficiency, long storage duration, large capacity, long life, low cycle cost	Specific topological site, large land use, high capital cost and adverse effect on environment
Super capacitor	Moderate efficiency, long life	Low energy density, limited power system application, high discharge rate

26.2.2 ESS Based on Hydrogen Storage

Hydrogen is the lightest, simplest and most abundant element of the universe, which can be produced basically by water reduction with carbon or electrolysis and be used as storage medium for electricity. Hydrogen is not a primary source of energy. However, it becomes an attractive energy carrier when split from these other elements by using a source of energy. Hydrogen, as clean energy carrier, is considered to be the clean fuel of future particularly for energy storage and transport. Figure 26.10 shows the hydrogen economy as a network composed of three functional steps: production, storage and use. There are basic technical means to achieve each of these steps, but none of them can yet complete with fossil fuels in cost, performance, or reliability (Carton et al. 2012).

The advantages of hydrogen energy are: energy security, sustainability, less pollution and economic viability by potentially shaping the future global energy markets. Therefore, Hydrogen is a worldwide accepted clean energy carrier as it is source independent and has a high energy content per mass compared to petroleum as listed in Fig. 26.11. The date are from (Cau et al. 2014).

Many researchers have observed and modeled hydrogen storage systems based on fuel cell technology, the study (Carton and Olabi 2010) outlined a review of energy storage options for Ireland including the use of hydrogen and fuel cell and discussed the key strategy for improving low-temperature proton exchange membrane fuel cell performance and durability. The paper (Midilli et al. 2005) reviewed the basics and fundamentals related to hydrogen as a fuel and its importance to the development of a sustainable future. The study (Chutichai et al. 2013) focused on the performance analysis of an integrated biomass gasification and proton exchange membrane fuel cell system and found that the electrical efficiency of the integrated system is 22 %, and when waste heat recovery is considered, the total efficiency of the system is 51 %. The study (Doherty et al. 2010) developed a model of a biomass-solid oxide fuel cell system capable of predicting performance under diverse operating conditions. Some researchers paid considerable attention to the thermodynamic cycle using hydrogen energy. The study (Bose et al. 2013) proposed a dual fuel mode in a diesel engine, with hydrogen injected into the intake manifold and ignited with diesel injected in the conventional manner. The study also attempted to explain the application of the fuzzy logic-based Taguchi analysis to optimize the performance parameters. The study (Niknam et al. 2013) evaluated

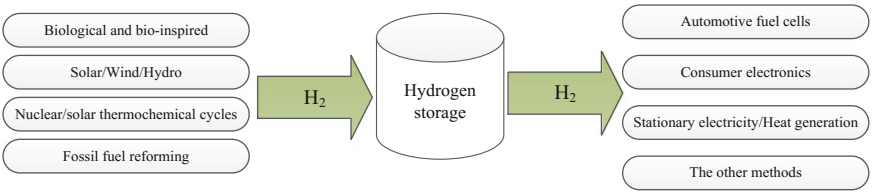

Fig. 26.10 Comparison between hydrogen and different fuel as source of energy

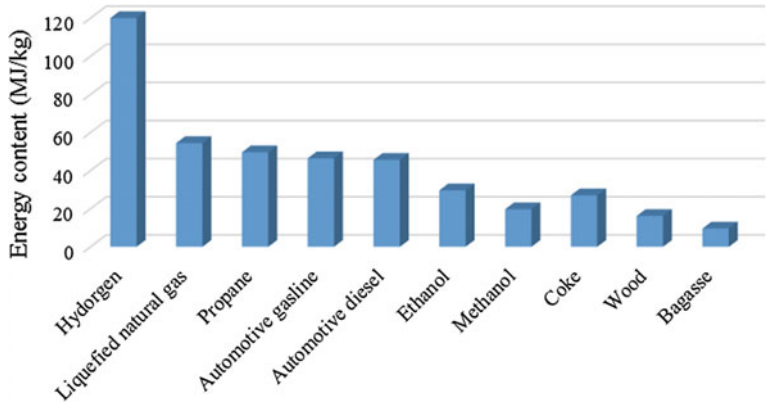

Fig. 26.11 Comparison between hydrogen and different fuel as source of energy

a plant with combined heat and power energy consumption and its environmental performance in different fuels with hydrogen percentages of 0, 5, and 8 %. The data analysis shows that modifying engine settings and increasing hydrogen amount significant affect the production of energy. Consequently, the combined heat and power fuelled with mixtures provides a lower thermal power. The study (Ulleberg et al. 2010) proposed the probabilistic energy management of a renewable micro-grid with hydrogen storage using the self-adaptive charge search algorithm.

26.2.2.1 Hydrogen Production

There are many processes for hydrogen production, mainly include: hydrogen production using fuel processing technologies and from alternative resources such as biomass and water. Figure 26.12 shows the hydrogen production technologies.

Fuel processing technologies convert a hydrogen containing material such as gasoline, ammonia, or methanol into a hydrogen rich stream. Fuel processing of

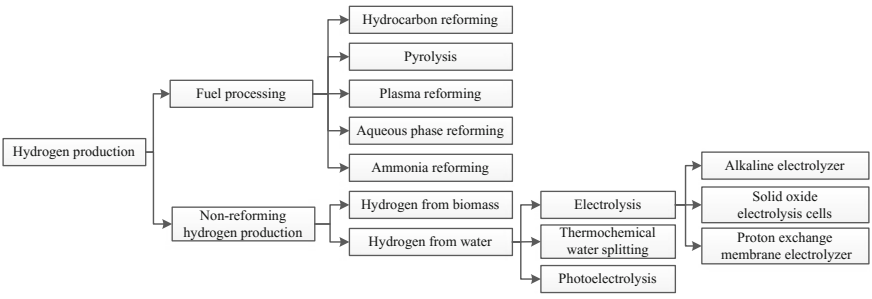

Fig. 26.12 Hydrogen production technologies

methane is the most common hydrogen production method in commercial use today. The technologies include: hydrocarbon reforming, plasma reforming, aqueous reforming, and pyrolysis. There are three primary techniques used to produce hydrogen from hydrocarbon fuels: steam reforming, partial oxidation, and autothermal reforming. The reforming process produces a gas stream composed primarily of hydrogen, carbon monoxide and carbon dioxide. Endothermic steam reforming of hydrocarbons requires an external heat source. Pyrolysis is another hydrogen-producing technology where the hydrocarbon is decomposed (without water or oxygen present) into hydrogen and carbon. Pyrolysis can be done with any organic material and is used for the production of hydrocarbons and carbon nanotubes and spheres (DemİRbaŞ 2005). In plasma reforming the overall reforming reactions are the same as conventional reforming; however, energy and free radicals used for the reforming reaction are provided by a plasma typically generated with electricity or heat (Matsui et al. 2005). Aqueous phase reforming is under development to process oxygenated hydrocarbons or carbohydrates to produce hydrogen (Momirlan and Veziroglu 2002). These reactors often operate at pressures up to 25–30 MPa and temperatures ranging from 220 to 270 °C. Ammonia reforming has been proposed primarily for use with fuel cells for portable power applications. It is an inexpensive fuel that, due to its use in fertilizer production, has an extensive distribution system including thousands of miles of pipeline (Rodrigues et al. 2014).

Hydrogen is also produced from biomass and water. Biomass is available from a wide range of sources such as animal wastes, municipal solid wastes, crop residues, short rotation woody crops, agricultural wastes, sawdust and aquatic plants. The current biomass technologies include: gasification, pyrolysis, conversion to liquid fuels by supercritical extraction, liquefaction, hydrolysis, etc. followed in some cases by reformation, and biological hydrogen production.

The electrolysis of water enables atoms to be split into oxygen atoms and hydrogen atoms. Gases are produced separately at the electrodes and have high purity. It is essentially the conversion of electrical energy to chemical energy in the form of hydrogen, with oxygen as a useful by-product. The most common electrolysis technology is alkaline based, but more proton exchange membrane (PEM) electrolysis and solid oxide electrolysis cells (SOEC) units are developing. Alkaline electrolyzers are typically composed of electrodes, a microporous separator and an aqueous alkaline electrolyte of approximately 30 wt% KOH or NaOH. The typical current density is 100–300 mA cm^{-2} and alkaline electrolyzers typically achieve efficiencies 50–60 % based on the lower heating value of hydrogen. PEM electroyzers build upon the recent advances in PEM fuel cell technology. It use Pt black, iridium, ruthenium, and rhodium for electrode catalysts and a Nafion membrane which not only separates the electrodes, but acts as a gas separator. PEM electrolyzers have low ionic resistances and therefore high currents of >1600 mA cm^{-2} can be achieved while maintaining high efficiencies of 55–70 %. At present, electrolyzers can be operated at pressure between atmospheric pressure and 200 bars. For instance, advanced alkaline electrolysis may advance the conversion efficiency to 90 % at optimal pressure. Generally, the higher the pressure of the processes, the higher the efficiency obtained. A study (Gutiérrez-Martín et al. 2010)

reported that the current water electrolysis equipment price ranges from 1,000 dollars/kW to 2,500 dollars/kW and the efficiency ranges from 60 to 70 %. Another study (Sherif et al. 2005) stated that, to make wind power more competitive, the future target price should be 400 dollars/kW, with efficiency of up to 75 % and capacity of up to MW level. Table 26.2 summarizes the technologies of hydrogen produced form water (Holladay et al. 2009).

26.2.2.2 Hydrogen Storage

Currently, the common method is to develop storage with sufficient capability. There are four main technologies for hydrogen storage, hydrogen pressurization and the hydrogen adsorption in metal hydrides, the adsorption of hydrogen on carbon nanofibres and the liquefaction of hydrogen.

Pressurized hydrogen technology relies on high materials permeability to hydrogen and to their mechanical stability under pressure. Currently steel tanks can store hydrogen at 200–700 bar but present very low ratio of stored hydrogen per unit weight (Hadjipaschalis et al. 2009). For the 200 bar system, a standard 16-steel cylinder package (40L each) allows for a hydrogen storage of 0.8 wt%. For the standards of 350 and 700 bar, which allows a storage capacity of 6 and 6.7 % by weight (Di Profio et al. 2009). Storage capability increase with higher pressures but stronger materials are then required. Research is currently under way to materials that are adequate for use in high pressures. The study (Zheng et al. 2012) indicates that, according to the US Ford Company, a high-pressure hydrogen storage vessel with pressure of up to 70 MPa, aluminum alloy liner, and a resin layer of carbon fiber reinforced outsourcing is extensively used in hydrogen energy-related industries. A scholar (Hua et al. 2011) proposes storage tank design pressures of 350 bar (5,000 psi) and 700 bar (10,000 psi). Metal hydride storage is based on the excellent hydrogen absorption properties of these compounds. During the formation of the metal hydride, hydrogen molecules are split and hydrogen atoms are inserted in spaces inside the lattice of suitable metals and/or alloys. These compounds have a low equilibrium pressure at room temperature in order to prevent leaks and guarantee containment integrity and a low degree of sensitivity to impurities in the hydrogen stored. For example, $LaNi_5H_6$ as the

Table 26.2 Technology summary table

Technology	Feed stock	Efficiency	Maturity
Alkaline electrolyzer	H_2O, electricity	50–60 %	Commercial
PEM electrolyzer	H_2O, electricity	55–70 %	Near term
Solid oxide electrolysis cells	H_2O, electricity, Heat	40–60 %	Med. term
Thermochemical water splitting	H_2O, heat	Not available	Long term
Photoelectrochemical water splitting	H_2O, sunlight	12.4 %	Long term

representative of low temperature hydrides, MgH_2 as the high temperature hydrides and $NaAlH_4$ as alanate. Metal hydride compounds exhibit rather low mass absorption capacities and do require thermal management system. This is because the absorption of hydrogen is an exothermic reaction while desorption of hydrogen is endothermic. Liquid hydrogen storage technology use is currently limited. This is due to the properties and cost of the materials used in the manufacturing of the container/tank and the extreme temperatures that are required for such storage. The typical temperature required to maintain hydrogen in a liquid state is around −253 °C. It requires amounts of energy equal to about one third of the energy in liquefied hydrogen and the hydrogen self-discharge of the tank may reach 3 % daily which translates to 100 % self-discharge in 1 month. Carbon nanotubes claimed to have a high storage capacity, but in this chapter it was not introduced so much because its high cost and self-discharge.

The study (Satyapal et al. 2007) conducted a detailed comparison of hydrogen storage and analyzed the efficiency of the hydrogen energy system. A study (Blechinger et al. 2014) used a model to analyze the hydrogen-based renewable energy systems. Table 26.3 shows the current storage methods (Gao et al. 2014). This paper develops high-pressure tanks as a method for hydrogen storage.

26.2.2.3 Hydrogen Utilization

Fuel Cells

Fuel cells are one of the main facilitating technologies for the development of future hydrogen economy. A fuel cell can be categorized as an indirect ESS. The constructional feature is similar to that of BESS but differ in their mode of operations in the sense that the device consumes fuel from external supply system to generate electricity. In a fuel cell, water, heat and electricity are produced from a prevailing electrochemical reaction between the reactants. The chemical sequence is such that the reactants flow in and reaction products flow out, while the electrolyte stays in the cell. Figure 26.13 illustrates a simple FC.

NASA was the first to deploy the applications fuel cells, United Technology Corporation (UTC) and Francis Thomas Beacon have developed various capacities

Table 26.3 Hydrogen storage methods

Storage method	Compressed hydrogen	Liquid hydrogen	Metal hydrides
Equipment	High-pressure tanks	Low-pressure and adiabatic tanks	Specific metallic alloys
Main parameters	700 ~ 1000 bar	−253 °C	≥ 250 °C
Characteristics	Feasibility and manageability	Considerable energy cost and high net loss	Inter-mass of metal has to be installed and moved

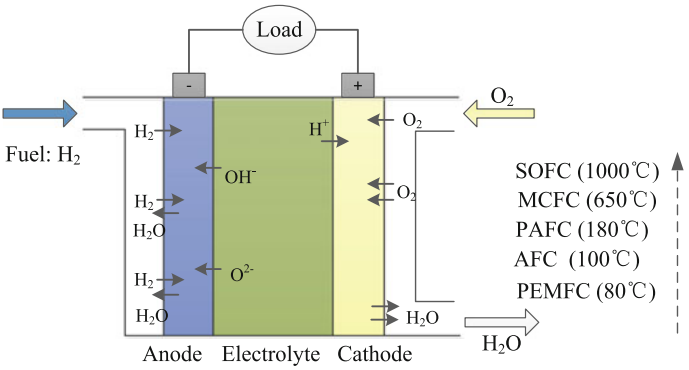

Fig. 26.13 Illustration of fuel cells

of fuel cells for applications in recent years. Varieties of fuel cells have been used in the last two decades for different applications, mostly for replacing internal combustions engines, providing power in stationary systems and portable power consumptions as well as automobile technologies for auxiliary power supply. Output energy efficiency of a fuel cell is in the range of 40–65 %, it is higher than any other ESS, but the cost is estimated between 550 and 8800 $/kW (Hu et al. 2015). The fuel cells which used the hydrogen as single fuel including: alkaline fuel cell (AFC), proton exchange membrane fuel cell (PEMFC), phosphoric acid fuel cell (PAFC). Molten carbonate fuel cell (MCFC) used a humidified mixture of hydrogen and CO as fuel. Solid-oxide fuel cell (SOFC) used a humidified mixture of hydrogen and CO, or single CH_4 as fuel. Direct methanol fuel cell (DMFC) used methanol and phosphoric acid fuel cell (PAFC) used ethanol as fuel (Kimiaie et al. 2014). Figure 26.14 shows the comparison of efficiency and operating temperature of FCs.

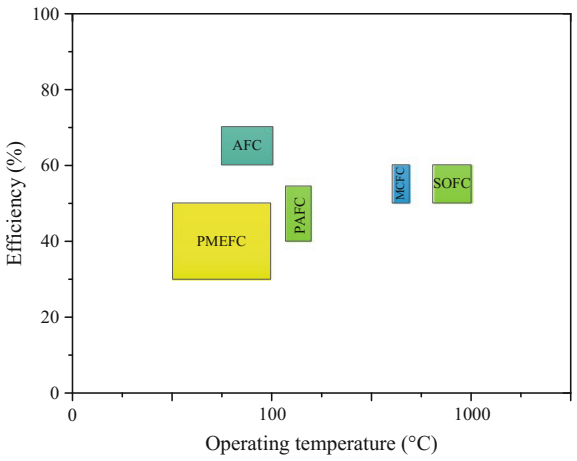

Fig. 26.14 Comparison of efficiency and operating temperature of FCs

The efficiency of AFC is highest, it can reach 60–70 %, and the operating temperature is between 70 and 100 °C (Dunn 2002).

Hydrogen fuel cell vehicles have in recent times emerged as a zero tailpipe-emission substitute to the battery electric vehicle because of its emission free nature. The electrochemical reaction approach for electricity production within the hydrogen fuel cell is such that hydrogen fuel passes through the anode while the oxidant (oxygen) passes over the cathode. This cross movement allows the formation of hydrogen ions and electrons at the anode and thereby guarantees the electrons produced to flow through an external circuit for electrical energy production. Another type of fuel cell used hydrogen as fuel is PEMFC, which is the promising technologies for clean and efficient power generation in the 21st century. The power delivery capacity of PEMFCs ranges from 0.1 to 100 kW and the operating efficiency is usually in the range of 40–50 %. PEMFC design simplicity and robustness, the weight is low and the power density is high. It has been successfully used in stationary power generators, light power vehicles, power bicycles and hybrid power buses (Lin et al. 2005). MCFC use lithium-sodium or Lithium-potassium carbonate salts as the electrolyte at high temperature (about 650 °C) (Folkesson et al. 2003). The fuel gas used in this system is a humidified mixture of H_2 and CO, the oxidant is a mixture of O_2 and CO_2, which may contain water vapor (Dicks 2004). MCFCs have the potential to utilize various kinds of fuels unlike other FCs. In Japan and United States, MCFC are at present undergoing some technological modifications to handle fuels like syngas, natural gas and biogas for power generation. There is also another advantage where the platinum catalyst can be replaced with the nickel type without any documented effects. MCFC have also being used in many hybrid power systems. Quite a lot of MCFCs hybrid systems with fuel-to-electricity efficiencies above 70 % have been conceptualized with some under development at the moment. In hybrid system with turbines, MCFC can produce 55–90 % electricity while the remainder capacity can be attributed to turbines. The chapter mainly reviews the hydrogen fuel cells. The characteristics of hydrogen fuel cells are shown in Table 26.4.

Hydrogen-Oxygen Combined Cycle

The hydrogen-oxygen combined cycle takes full advantage of traditional gas-steam combined cycle equipment, improves the fuel recycling method, and obtains a more economical and efficient power generation system. Comparatively, the cycle operated and is achieved more easily than fuel cell systems that also use hydrogen as fuel, especially considering the size of the system, equipment costs, and lifetime. Figure 26.15 shows a simple hydrogen-oxygen combined cycle.

Pure hydrogen and oxygen as fuel of the system combust in the combustion chamber and then obtain high temperature and high pressure superheated steam at

Table 26.4 Characteristics of hydrogen fuel cells

Type	Operating temperature (°C)	Electrolyte	Charge carrier	Catalyst anode	Efficiency (%)	Qualified power (kW)	Fuel
AFC	70–100	Potassium hydroxide	H^+	Ni	60–70	10–100	H_2
PEMFC	50–100	Ion exchange membrane	H^+	Pt	30–50	0.1–500	H_2
PAFC	150–220	Immobilized liquid phosphoric acid	H^+	Pt	40–55	5–10000	H_2
MCFC	650–700	Immobilized liquid molten carbonate	CO^{2-}	Ni	50–60	100–300	Reformate or CO/H_2
SOFC	800–1000	Ceramic	O^{2-}	Ni	50–60	0.5–100	Reformate CO/H_2 or CH_4

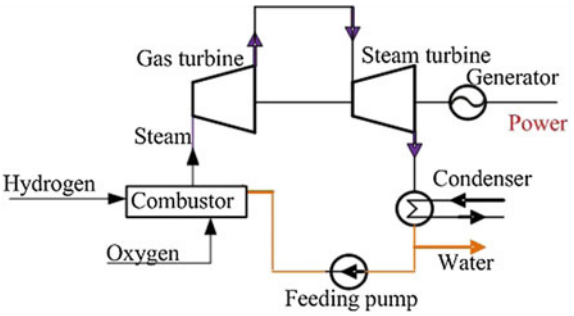

Fig. 26.15 Diagram of a simple hydrogen-oxygen combined cycle

about 1,500 °C. The superheated steam expands through the gas turbine first before entering the steam turbine to expand further. With feedback to the condenser after full expansion and heat transfer, part of the condensed water is supplied to the combustion chamber through the pipe as an inert substance after treatment, preventing the high temperature from damaging the combustion chamber. The rest of the water passes the electrolyzer and produces hydrogen continuously. The system is fully closed, and the recycled material used is water. This system does not produce pollution and is therefore environment friendly. Researchers around the world have made further studies on the hydrogen-oxygen combined cycle. Literature (Najjar 2000) analyzed and discussed the combined cycle of hydrogen and oxygen. The study (Gao et al. 2014) shows that at a temperature of 1,700 °C and a pressure of 30 MPa, efficiency can reach 62 %. Research has also been conducted on heat recovery and reheats parameters.

26.3 Process Configuration for Integrated ESS and Comparison with Other ESS

Based on electrolyzer, hydrogen and oxygen storage devices, and hydrogen-oxygen combined cycle, an integrated ESS was introduced. Figure 26.16 shows the diagram of the integrated storage system process. The system selects hydrogen as the intermediate medium when the power price is low. Electrical energy from hydrogen is obtained by electrolysis of the heated water in the electrolyzer. Energy conversion in this manner is clean, pollution-free, and easy to control. When the network overloads, hydrogen and oxygen are converted into electrical energy by the integrated system to supply bulk power.

Aspen Plus is introduced in this part for modeling and optimizing the corresponding systems. Energy efficiency is analyzed based on the result of Aspen Plus. The key assumptions for energy efficiency analysis in this part are listed in Table 26.5.

The three systems analyzed in the following sections are simple integrated ESS, integrated ESS with feed water heater, and integrated ESS with feed water heater and steam reheater.

26.3.1 Simple Integrated ESS

The simple integrated ESS is shown in Fig. 26.6. The energy flow chart for this system is presented in Fig. 26.17.

Based on the chart, the net energy efficiency for this system is 49.2 %. Power outputs from gas and steam turbines are 29.2 and 21.6 %. Energy loss in the system

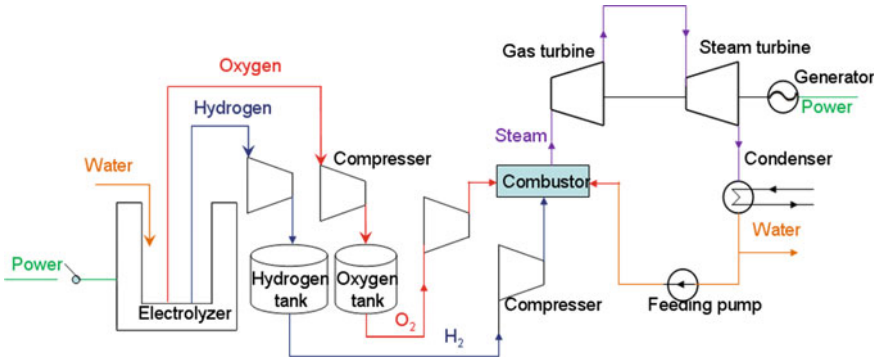

Fig. 26.16 Diagram of simple integrated energy storage system

Table 26.5 Key assumptions for energy efficiency analysis (Ibrahim et al. 2011; Lin et al. 2005)

Parameter	Value
Electrolysis efficiency	85 %
Hydrogen storage dynamic energy efficiency	91.5 %
Steam turbine efficiency	90 %
Gas turbine efficiency	88 %
Mechanical efficiency	98 %
Generator efficiency	99 %
Combustor efficiency	99 %
Heat exchanger efficiency	99 %
Temperature of (high pressure) combustor	1773 K
Out temperature of water heater	623 K
Feed water temperature	293.15 K
Pressure of (high pressure) combustor	30 MPa
Average pressure of hydrogen/oxygen tank	15 MPa
Maximum pressure of hydrogen/oxygen tank	30 MPa
Out pressure of gas turbine	4 MPa
Out pressure of high pressure steam turbine	0.05 MPa
Out pressure of low pressure steam turbine	0.005 MPa
H_2 mass flow ratio of combustion 2 to combustion1 in steam reheat system	0.424

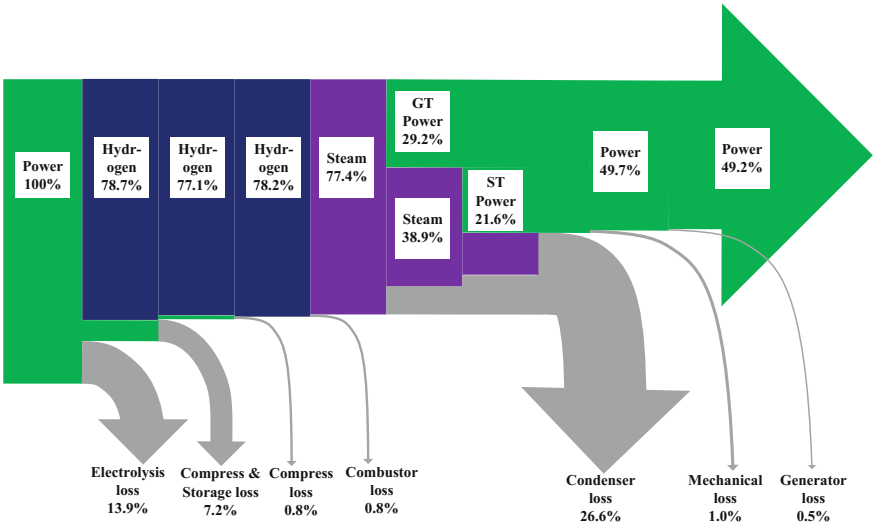

Fig. 26.17 Energy flow chart of a simple integrated energy storage system

consists of electrolyzer loss, first compressor and storage loss, second compressor loss, combustion loss, and condenser loss, and their values are 13.9, 7.2, 0.8, 0.8, and 26.6 %. Mechanical loss is 1 % and generator loss is 0.5 %.

26.3.2 Integrated ESS with Feed Water Heater

Figure 26.18 shows the integrated ESS with feed water heater. The electrolyzer and the hydrogen storage systems are the same as those in the simple integrated system. The hydrogen-oxygen combined cycle is improved through feed water heating.

Part of the water fed into the combustor is warmed on the cool side by steam from the gas turbine. Before water is fed into the combustor, feed water is heated by the water heater from the cool side. The superheated steam on the hot side enters the steam turbine to expand further after the heat exchange. The feed water heater not only reduces the temperature span that exists between the gas and steam turbines but also raises the temperature of the feed water and improves the energy efficiency of the systems.

The energy flow chart of this system is shown in Fig. 26.19. The results indicate that the net energy efficiency of this system is 49.7 %, or 0.5 % higher than that of the simple system. The power output of gas turbine is higher than that of the simple system but is lower than that of the steam turbine because 14.8 % of the energy is fed back to the combustor by feed water. The largest energy loss is caused by the condenser, approximately 26.1 %, or 0.5 % lower than that in the simple system. Energy loss before the hydrogen-oxygen combined cycle also consists of electrolyzer loss, first compressor and storage loss, second compressor loss, and combustion loss, and their values are 13.9, 7.2, 0.8, 0.8, and 26.6 %, similar to the simple system.

26.3.3 Integrated ESS with Feed Water Heater and Steam Reheater

The integrated system with feed water heater and steam reheater is shown in Fig. 26.20. Aside from the water heater, combustor 2 and high-pressure steam turbine (HPST) are introduced into this system. Feed water is heated by the steam

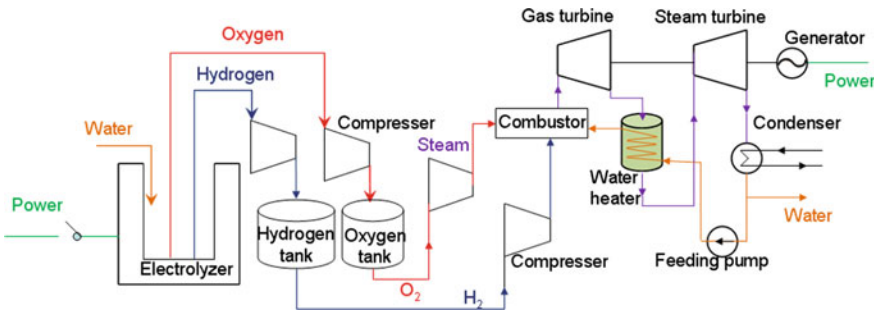

Fig. 26.18 Diagram of integrated ESS with feed water heater

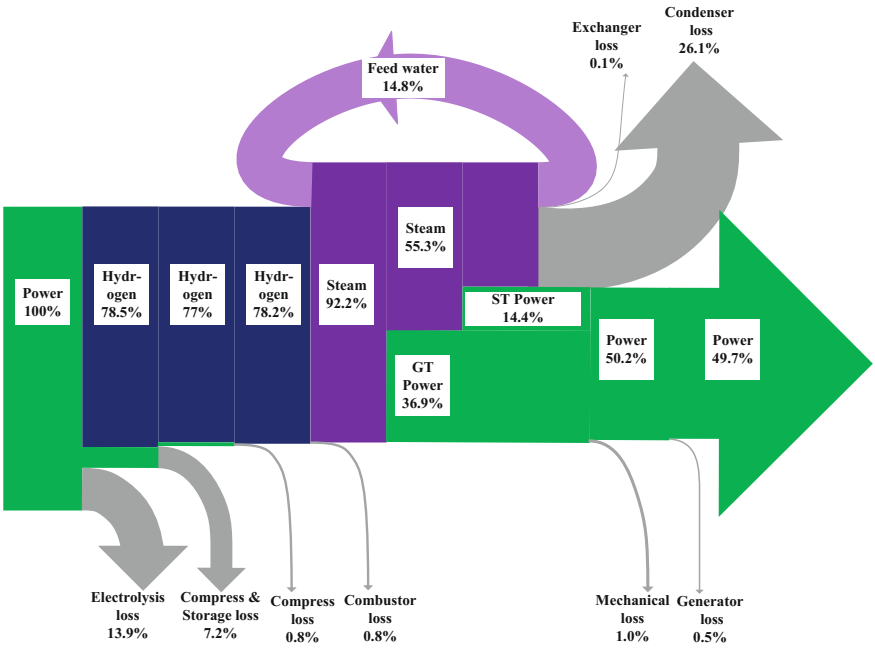

Fig. 26.19 Energy flow chart of an integrated ESS with feed water heater

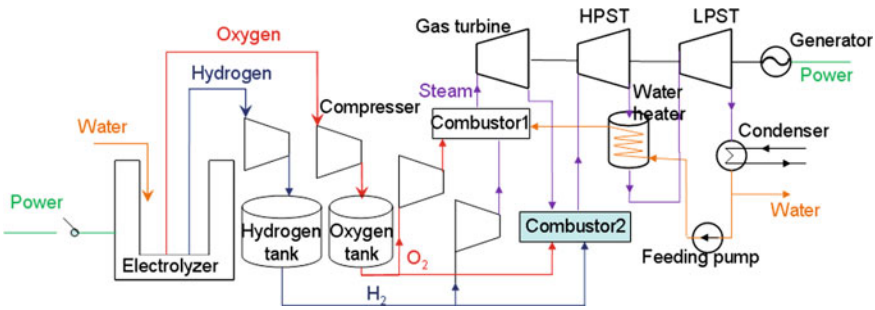

Fig. 26.20 Diagram of integrated ESS with feed water heater and steam reheater

from the HPST and then it enters into combustor 1. Steam from the gas turbine is imported into combustor 2, mixed and combusted with hydrogen and oxygen, and then enters into the HPST. The combustion temperature of combustor 2 is 1,500 °C, the same as combustor 1.

The energy flow chart of this system is shown in Fig. 26.21. The results indicate that the net energy efficiency of this system is 54.6 %, or 5.4 % higher than that of the simple system, which is caused by the feed water heater and steam reheater. The

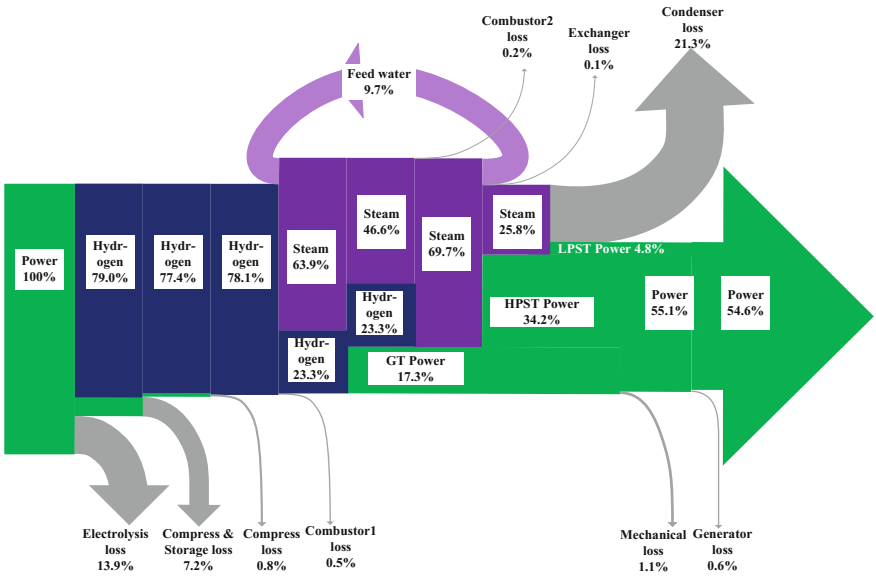

Fig. 26.21 Energy flow chart of an integrated ESS with feed water heater and steam reheater

power outputs of the gas turbine, HPST, and low-pressure steam turbine are 17.3, 34.2, and 4.8 %. HPST produces majority of the power. Energy loss is significantly lower than that in the previous two systems, which is approximately 21.3 %. Energy loss before the hydrogen-oxygen combined cycle is approximately 21.1 %, the same as that in the previous two systems.

The results obtained in the energy efficiency analysis clearly demonstrate that the net energy efficiency of the integrated ESS is 49–55 % and that feed water heater and steam reheater are viable for integrated ESS to improve efficiency.

26.3.4 Comparison Between Integrated ESS and Other ESS

This study analyzes several ESS such as CAES, SMES, BASS, pumped hydro storage, and integrated ESS. The comparison between ESS and other ESS is presented in Table 26.6. The comparison indicates that the integrated system has certain advantages in most aspects. The integrated system can be large scale, has flexible operation and rapid response, and is environment friendly. Disadvantages of the integrated system include low efficiency and some technological limitations, especially for the gas turbine that burns hydrogen. As such, research on the development stage of the integrated system is still under way. The system can be simulated through Aspen Plus. This paper makes a case analysis with wind power.

Table 26.6 Comparison of ESS

Energy storage system	Pumped storage	CAES	SMES	BESS	Integrated storage
Intermediary	Water	Air	Electromagnetism	Electricity	Hydrogen
Energy efficiency	70–85 %	50–80 %	>95 %	70–95 %	49–55 %
Advantages	Mature, large scale, environment friendly	Large scale	Rapid response, flexible unit	Fast delivery, low operating cost	Large scale, flexible unit, environment friendly, rapid response
Disadvantages	Long lead time, geographical limitation	Geographical limitation, low efficiency	Short storage time, high CAPEX	Small scale, medium CAPEX	Low efficiency, hydrogen-based gas turbine
Development stage	Commercialize	Demonstration	Demonstration	Commercialized	Research

26.3.5 Integrated ESS Based on Hydrogen Storage Design

As shown in Fig. 26.22, the wind farm is connected to the ESS. The stability actions of the ESS can offset the variability and intermittence of wind power. On-site hydrogen production can protect against regional network overloading or provide a short-term backup supply during slack periods. Hydrogen can be stored at a utility level and then distributed electrically along transmission lines during peak demand periods. Wind power can be used with high efficiency and is environment friendly.

A previous article (Winter 2009) argues that the relationship between energy input and hydrogen output is a nonlinear function, which is influenced by the installed capacity, size of the electrolyzer, and operation time. When designing the entire system, optimal equipment is required, which can convert the maximum energy and enable work with the highest efficiency. We chose the largest and the best performing model as the basic unit with a 2.13 MW capacity. As shown in the wind load curve analysis, the total power of the electrolyzer can be limited to 95 MW, with 45 units connected in parallel. The number of operating units depends on the current wind load curve. Each electrolyzer consists of 230 cells (0.6 m^2) and operates with 5.15 kA at maximum power. Thus, the nonce density is 8.58 kA/m^2, and the potential between electrodes is 1.8 V. The efficiency of hydrogen production is 85 %. Providing a tank with enough capacity to store water is important. Water is provided for the electrolyzer. Part of the water is recycled from the hydrogen-oxygen combined cycle, and the remaining parts are fed by disposal water. Suitably sized containers are used because of the special feature of the hydrogen yield. Having more reserves to distribute the surplus energy in a greater number of days and to match with the production planning is possible. Downstream of the electrolyzer, a set of compressors is required to increase the exit hydrogen pressure to 200 atm. Then, hydrogen is transported through gas pipelines to spherical deposits 9.55 m in diameter (340 m^3), which can store 5,800 kg of hydrogen. Three tanks are connected with each other, with the first two tanks being the main storage containers to maintain constant pressure during unloading. The third is the backup tank to be used

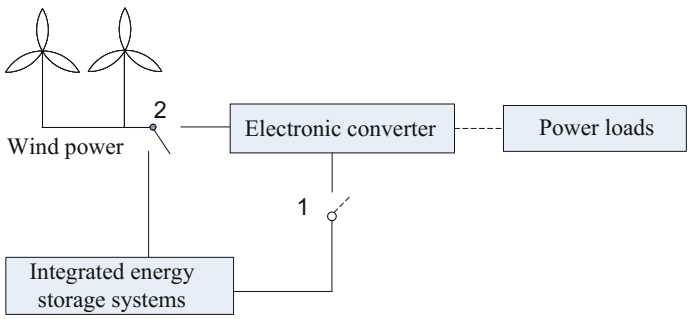

Fig. 26.22 Diagram of installed ESS

during an emergency. The hydrogen-oxygen combined cycle with feed water heater and steam reheater is the conversion component (Gao et al. 2014). The size of the system is established by multiplying the nominal wind power and the efficiency of the electrolyzers ($100 \times 0.85 = 85$ MW). However, theoretically, this power is the maximum energy available in the form of hydrogen because of the uncertainty of various factors. We selected different generators for two operating modes based on the hydrogen-oxygen combined cycle.

26.3.6 Case Studies for Different Operating Modes

The challenge of an integrated system is that a continuous balance must be maintained between generation and demand. In this study, two extremes of system operation modes are proposed, namely, the intermittent operating mode and the continuous operating mode. The intermittent mode is an extension of the current practice, with hydrogen storage reaching high values. The hydrogen-oxygen combined cycle starts to work at a high hydrogen pressure level. The continuous operating mode is relatively limited by the wind power load. The hydrogen-oxygen combined cycle begins to run at the basic level of hydrogen storage.

(a) Intermittent operating mode

Given the variability and unpredictability of wind, hydrogen storage values fluctuate in spots in this mode. The electrolyzer systems initially run intelligentized. When the two hydrogen storage tanks reach their full capacity, the hydrogen-oxygen combined cycle system begins to work to provide stable power coupled with hydrogen and oxygen consumption. The two main tanks are connected by pipes. The pressure is regulated by a compressor, and hydrogen produced by the electrolyzer is introduced into the tank continuously. The systems operate at the pressure level of 10–20 MPa. The storage container cannot maintain the pressure in the set point. The hydrogen-oxygen combined cycle shuts down until the tanks reach their full capacity again. Energy efficiency was analyzed based on the result of Aspen Plus. The results show that the net energy efficiency for this system is 50.1–50.9 % and the hydrogen-oxygen combined cycle starts up and shuts down more than once at the installed capacity of 47 MW as a result of the limitation of hydrogen production and storage. However, the systems can convert the stored energy into electricity effectively during on-peak times.

(b) Continuous operating mode

In this mode, after achieving the full storage capacity of hydrogen and oxygen, we run the hydrogen-oxygen combined cycle system at a lower load capacity of 20 MW relative to the intermittent operating mode. The differential pressure

between the communicating hydrogen storage container and the hydrogen-oxygen combined cycle system is regulated by a compressor. The system operates at the pressure level of 10–20 MPa. The fuel is provided by one tank. When the tank pressure cannot maintain the running pressure, a second tank assumes the task. The hydrogen produced by the electrolyzer is stored in the first tank. The system continuously operates until the two tanks cannot contain the instantaneous fuel input. The system operates throughout the day at a capacity of 20 MW and can continuously convert wind power to stable electricity. The result shows that the two tanks can maintain the pressure of 10–20 MPa before 3:30 p.m. After this time, the third tank as the backup tank is used because of its high hydrogen production and low consumption.

(c) Performance comparison and economic analysis of intermittent and continuous operating modes

We compared the performances of the two operating modes presented in the previous section. It shows that the performance of the hydrogen-oxygen combined cycle depends largely on the capacity of hydrogen storage and power generation. As shown in Fig. 26.23, the daily average efficiency of the combined cycle at the intermittent operating mode is 3.6 % points higher than that at the continuous operating mode. While operating at the intermittent mode has a lightly higher energy efficiency to meet the peak load, frequent starts up and shuts down involved with the intermittent mode can significantly reduce the lifetime of the equipment. As for base load, the continuous operating mode is superior to the intermittent mode.

An economic analysis of the intermittent operating mode and continuous operating mode is shown in Table 26.7. Results show that the continuous operating mode has relatively lower investment cost, but its annual income is also lower than the intermittent operating mode.

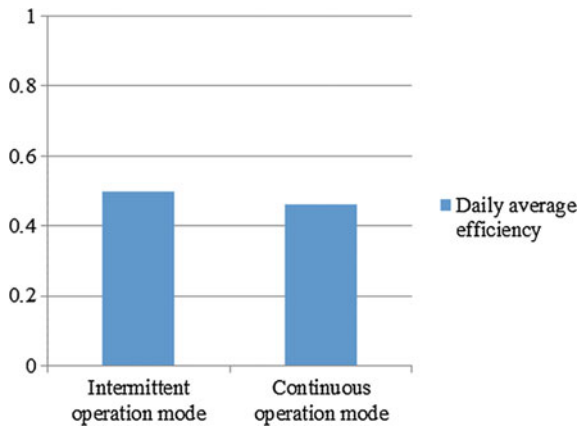

Fig. 26.23 Daily average efficiency of two different operation modes

Table 26.7 Preliminary economic analysis for different operation modes

Items	Intermittent operation mode	Continuous operation mode
Number of hydrogen storage tanks	2	3
Number of oxygen storage tanks	1	1
Installed capacity	47 MW	20 MW
Power generation time	12.5 h	24 h
Peak and valley electric prices	0.03 $/kWh (23:00–7:00), 0.04 $/kWh (7:00–9:00, 12:00–17:00, 22:00–23:00), 0.05 $/kWh (9:00–12:00, 17:00–22:00)	
Total equipment cost of integrated system	48 M$	35 M$
Total capital cost of integrated system	80 M$	59 M$
Capital cost per installed capacity	1700 $/kW	2450 $/kW
Generating income	13 M$/y	10 M$/y

26.4 Conclusion

Wind energy contributes to power generation systems in China by supplying electricity to meet the increase in demand. Growth pattern is foreseen to have a much faster pace in the future. This paper presents an ESS integrated by an electrolyzer and hydrogen-oxygen combined cycle. The integrated ESS is an environment-friendly system that cannot be constrained by geography and materials. Based on the energy flow chart, the energy efficiency of integrated ESS is 49–55 %, with steam parameters of 1,500 °C and 30 MPa. Based on the case analysis, this paper analyzes and optimizes the operating modes of wind-hydrogen and hydrogen-oxygen combined cycle systems. The major advantage of storage systems is the reduction in wind variability and intermittent performance. The average net efficiency of the integrated ESS can be up to 50 %, and the capital cost of the integrated system is about 2,000 $/kW. Hydrogen storage is closely linked to the system's capacity. Wind power can be utilized in a more scientific, reasonable, and efficient way through integrated systems.

References

Ayodele, T. R., & Ogunjuyigbe, A. S. O. (2015). Mitigation of wind power intermittency: Storage technology approach. *Renewable and Sustainable Energy Reviews, 44*, 447–456.

Becherif, M., Ramadan, H. S., Cabaret, K., Picard, F., Simoncini, N., & Bethoux, O. (2015). Hydrogen energy storage: new techno-economic emergence solution analysis. *Energy Procedia, 74*, 371–380.

Blechinger, P., Seguin, R., Cader, C., Bertheau, P., & Breyer, C. (2014). assessment of the global potential for renewable energy storage systems on small islands. *Energy Procedia, 46*, 325–331.

Bose, P. K., Deb, M., Banerjee, R., & Majumder, A. (2013). Multi objective optimization of performance parameters of a single cylinder diesel engine running with hydrogen using a Taguchi-fuzzy based approach. *Energy, 63*, 375–386.

Carton, J. G., & Olabi, A. G. (2010). Wind/hydrogen hybrid systems: Opportunity for Ireland's wind resource to provide consistent sustainable energy supply. *Energy, 35*, 4536–4544.

Carton, J. G., Lawlor, V., Olabi, A. G., Hochenauer, C., & Zauner, G. (2012). Water droplet accumulation and motion in PEM (proton exchange membrane) fuel cell mini-channels. *Energy, 39*, 63–73.

Cau, G., Cocco, D., Petrollese, M., Knudsen, K. R. S., & Milan, C. (2014). Energy management strategy based on short-term generation scheduling for a renewable microgrid using a hydrogen storage system. *Energy Conversion and Management, 87*, 820–831.

Cavallo, A. J. (2001). Energy storage technologies for utility scale intermittent renewable energy systems. *Journal of Solar Energy Engineering, 123*, 387–389.

Chutichai, B., Authayanun, S., Assabumrungrat, S., & Arpornwichanop, A. (2013). Performance analysis of an integrated biomass gasification and PEMFC (proton exchange membrane fuel cell) system: Hydrogen and power generation. *Energy, 55*, 98–106.

D Az-Gonz Lez, F., Sumper, A., Gomis-Bellmunt, O., & Villaf Fila-Robles, R. (2012). A review of energy storage technologies for wind power applications. *Renewable and Sustainable Energy Reviews, 16*, 2154–2171.

Demirbaş, A. (2005). Bioethanol from cellulosic materials: A renewable motor fuel from biomass. *Energy Sources, 27*, 327–337.

Di Profio, P., Arca, S., Rossi, F., & Filipponi, M. (2009). Comparison of hydrogen hydrates with existing hydrogen storage technologies: Energetic and economic evaluations. *International Journal of Hydrogen Energy, 34*, 9173–9180.

Dicks, A. L. (2004). Molten carbonate fuel cells. *Current Opinion in Solid State and Materials Science, 8*, 379–383.

Doherty, W., Reynolds, A., & Kennedy, D. (2010). Computer simulation of a biomass gasification-solid oxide fuel cell power system using Aspen Plus. *Energy, 35*, 4545–4555.

Dunn, S. (2002). Hydrogen futures: toward a sustainable energy system. *International Journal of Hydrogen Energy, 27*, 235–264.

Etxeberria, A., Vechiu, I., Camblong, H., & Vinassa, J. M. (2010). *Hybrid Energy Storage Systems for renewable Energy Sources Integration in microgrids: A review. IPEC, 2010 Conference Proceedings*, 27–29 Oct. 2010 (pp. 532–537).

Ferreira, H. L., Garde, R., Fulli, G., Kling, W., & Lopes, J. P. (2013). Characterisation of electrical energy storage technologies. *Energy, 53*, 288–298.

Folkesson, A., Andersson, C., Alvfors, P., Alak la, M., & Overgaard, L. (2003). Real life testing of a Hybrid PEM fuel cell bus. *Journal of Power Sources, 118*, 349–357.

Gao, D., Jiang, D., Liu, P., Li, Z., Hu, S., & Xu, H. (2014). An integrated energy storage system based on hydrogen storage: Process configuration and case studies with wind power. *Energy, 66*, 332–341.

Guti Rrez-Mart, N. F., Confente, D., & Guerra, I. (2010). Management of variable electricity loads in wind—hydrogen systems: The case of a Spanish wind farm. *International Journal of Hydrogen Energy, 35*, 7329–7336.

Hadjipaschalis, I., Poullikkas, A., & Efthimiou, V. (2009). Overview of current and future energy storage technologies for electric power applications. *Renewable and Sustainable Energy Reviews, 13*, 1513–1522.

Holladay, J. D., Hu, J., King, D. L., & Wang, Y. (2009). An overview of hydrogen production technologies. *Catalysis Today, 139*, 244–260.

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. *Applied Energy, 137*, 913–924.

Hua, T. Q., Ahluwalia, R. K., Peng, J. K., Kromer, M., Lasher, S., McKenney, K., et al. (2011). Technical assessment of compressed hydrogen storage tank systems for automotive applications. *International Journal of Hydrogen Energy, 36*, 3037–3049.

Ibrahim, H., Ilinca, A., & Perron, J. (2008). Energy storage systems—characteristics and comparisons. *Renewable and Sustainable Energy Reviews, 12*, 1221–1250.

Ibrahim, H., Youn, S. R., Basbous, T., Ilinca, A., & Dimitrova, M. (2011). Optimization of diesel engine performances for a hybrid wind–diesel system with compressed air energy storage. *Energy, 36*, 3079–3091.

Jubeh, N. M., & Najjar, Y. S. H. (2012). Power augmentation with CAES (compressed air energy storage) by air injection or supercharging makes environment greener. *Energy, 38*, 228–235.

Kimiaie, N., Wedlich, K., Hehemann, M., Lambertz, R., Muller, M., Korte, C., et al. (2014). Results of a 20 000 h lifetime test of a 7 kW direct methanol fuel cell (DMFC) hybrid system—degradation of the DMFC stack and the energy storage. *Energy and Environmental Science, 7*, 3013–3025.

Kiviluoma, J., & Meibom, P. (2010). Influence of wind power, plug-in electric vehicles, and heat storages on power system investments. *Energy, 35*, 1244–1255.

Kousksou, T., Bruel, P., Jamil, A., El Rhafiki, T., & Zeraouli, Y. (2014). Energy storage: Applications and challenges. *Solar Energy Materials and Solar Cells, 120*(Part A), 59–80.

Lin, M., Cheng, Y., Lin, M., & Yen, S. (2005). Evaluation of PEMFC power systems for UPS base station applications. *Journal of Power Sources, 140*, 346–349.

Marino, C., Nucara, A., Pietrafesa, M., & Pudano, A. (2013). An energy self-sufficient public building using integrated renewable sources and hydrogen storage. *Energy, 57*, 95–105.

Matsui, Y., Kawakami, S., Takashima, K., Katsura, S., & Mizuno, A. (2005). Liquid-phase fuel re-forming at room temperature using nonthermal plasma. *Energy and Fuels, 19*, 1561–1565.

Midilli, A., Ay, M., Dincer, I., & Rosen, M. A. (2005). On hydrogen and hydrogen energy strategies: I: current status and needs. *Renewable and Sustainable Energy Reviews, 9*, 255–271.

Momirlan, M., & Veziroglu, T. N. (2002). Current status of hydrogen energy. *Renewable and Sustainable Energy Reviews, 6*, 141–179.

Najjar, Y. S. H. (2000). Gas turbine cogeneration systems: a review of some novel cycles. *Applied Thermal Engineering, 20*, 179–197.

Niknam, T., Golestaneh, F., & Shafiei, M. (2013). Probabilistic energy management of a renewable microgrid with hydrogen storage using self-adaptive charge search algorithm. *Energy, 49*, 252–267.

Rodrigues, E. M. G., Godina, R., Santos, S. F., Bizuayehu, A. W., Contreras, J., & Catal, O. J. P. S. (2014). Energy storage systems supporting increased penetration of renewables in islanded systems. *Energy, 75*, 265–280.

Satyapal, S., Petrovic, J., Read, C., Thomas, G., & Ordaz, G. (2007). The U.S. department of energy's national hydrogen storage project: Progress towards meeting hydrogen-powered vehicle requirements. *Catalysis Today, 120*, 246–256.

Sherif, S. A., Barbir, F., & Veziroglu, T. N. (2005). Wind energy and the hydrogen economy—review of the technology. *Solar Energy, 78*, 647–660.

Trianni, A., Cagno, E., Worrell, E., & Pugliese, G. (2013). Empirical investigation of energy efficiency barriers in Italian manufacturing SMEs. *Energy, 49*, 444–458.

Ulleberg, Ø., Nakken, T., & Et, A. (2010). The wind/hydrogen demonstration system at Utsira in Norway: Evaluation of system performance using operational data and updated hydrogen energy system modeling tools. *International Journal of Hydrogen Energy, 35*, 1841–1852.

Winter, C.-J. (2009). Hydrogen energy—abundant, efficient, clean: A debate over the energy-system-of-change. *International Journal of Hydrogen Energy, 34*, S1–S52.

Yekini Suberu, M., Wazir Mustafa, M., & Bashir, N. (2014). Energy storage systems for renewable energy power sector integration and mitigation of intermittency. *Renewable and Sustainable Energy Reviews, 35*, 499–514.

Zheng, J., Liu, X., Xu, P., Liu, P., Zhao, Y., & Yang, J. (2012). Development of high pressure gaseous hydrogen storage technologies. *International Journal of Hydrogen Energy, 37*, 1048–1057.

Chapter 27
State Monitoring and Fault Diagnosis of Wind Turbines

Wang Lingmei, Meng Enlong, Shen Jianlin, Guo Dongjie, Yin Shaoping, Jiao Jinxiu, Li Ruize and Yu Yuehan

Abstract In this chapter, after the comprehensive exposition of the basic knowledge of state monitoring and fault diagnosis of wind turbine, a new targeted method for fault diagnosis is proposed. Starting from the non-stationary characteristics of fault signal of transmission system of wind turbine, a new improved method of combining wavelet packet and envelope spectrum for fault diagnosis is put forward. Finally, the author introduces the design idea and method of the overall structure and function of remote real-time state monitoring and fault diagnosis system of large-scale wind turbine in detail. Based on the field research of the author and his team into 3 MW wind turbine of a group company, the method is researched and developed for the state monitoring and fault diagnosis system of large-scale wind turbine, according to fault types and features of key equipment components of drive system of wind turbine as well as the actual needs of users. The system has been proved to have excellent performance in the field test. During introducing the

W. Lingmei (✉) · M. Enlong · S. Jianlin · Y. Shaoping · J. Jinxiu · L. Ruize · Y. Yuehan
Dadongguan Campus, Shanxi University, Taiyuan, Shanxi, China
e-mail: wanglingmei08@163.com

M. Enlong
e-mail: tymj52@163.com

S. Jianlin
e-mail: Shenjianlin@126.com

Y. Shaoping
e-mail: yin297@126.com

J. Jinxiu
e-mail: jjx8451539@163.com

L. Ruize
e-mail: 1017612859@qq.com

Y. Yuehan
e-mail: 634624574@qq.com

G. Dongjie
SPIC SHANXI NEW ENERGY CO., LTD., Taiyuan, Shanxi, China
e-mail: 418827949@qq.com

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_27

design idea and method of the overall structure and function of remote real-time state monitoring and fault diagnosis system of large-scale wind turbine, the author focuses on the algorithm design of fault diagnosis software of the data acquisition instrument.

27.1 What Is the State Monitoring of Wind Turbine?

As the most mature technology in the field of new energy power generation, with the most large-scale development potential and commercial development prospects, coupled with its prominent role in optimization of energy industry structure, improving the ecological environment, promoting sustainable development of the global economy and society and other aspects, wind power generation has become the strategic choice for the development of low carbon economy in the world. In the past 20 years, the development of wind power is very rapid around the world, and global cumulative installed capacity of wind power increases year by year. At the same time, the high cost of operation and maintenance of wind turbine affects the economic benefits of wind farm. Because of the restrictions of the wind resource distribution, wind farms are mainly located in the desert, mountain areas and grassland, far away from the city and monitoring center. Due to the harsh natural environment, the complexity of the generator and power electronic device and other factors, wind power generation equipment is very easily damaged. In order to improve operation and maintenance of wind farm, as well as the operating efficiency and reliability of wind turbine, it is essential to monitor operation status of each wind turbine. The wind farm requires reliable remote monitoring and unmanned operation control system as a necessary mean of normal operation and fault diagnosis of wind power equipment. The condition monitoring technology, through real-time observation of wind turbine components, can detect hidden faults in time and take effective measures to avoid the occurrence of major accidents, and changing regular maintenance and breakdown maintenance to predictive maintenance can effectively reduce operation and maintenance costs, and improve the economic efficiency of wind power (Wang 2014a).

Condition monitoring and fault diagnosis technology as a new subject, based on information theory, cybernetics, system theory and reliability theory, using modern testing instruments and computer technology as technical measures and combined with special laws of various diagnostic object, is gradually formed. At present, at home and abroad there are a number of manufacturers and universities engaged in research work related to condition monitoring, and the research on the condition monitoring and fault diagnosis of wind turbine is emerging in large numbers. But the technology is not mature yet, and the industrialization has not been fully realized. Therefore, this chapter will tell the story of the latest research developments and research results in this area.

Condition monitoring refers to judging the running state of the equipment, raising timely the alarm on the abnormal or fault state and providing information

and data for further fault analysis and performance evaluation, by using various (online or offline) monitoring and analysis instruments and a variety of detection, analysis and processing methods, through collecting, processing and analyzing of parameters of the running state of the equipment, which is an important basis for effective fault diagnosis.

State monitoring technology of wind turbine is mainly divided into the following categories: vibration analysis, oil monitoring, thermal imaging technology, process parameters monitoring, and performance parameter check; and also includes physical state examination of the material, strain measurement, acoustic monitoring, electrical effect, visual inspection, sensor self diagnosis technology, of which the strain measurement, acoustic emission and vibration monitoring can be used to detect the trend of blade fault; the vibration analysis of different sensors is mainly used for the condition monitoring of the gearbox; acoustic emission technology can also be used to detect mutation of the surface stress of the components caused by friction, especially to early detect pitting corrosion, cracks and other potential failure of the gear; and temperature monitoring can be used to check the operation of generator stator and rotor winding and torque measurements can be used for the fault detection of transmission system. Due to the highest probability of vibration fault occurrence and vibration signal containing the maximum amount of information, the vibration analysis method is the most widely used and effective method in the fault diagnosis of rotating equipment. In the actual operation of condition monitoring and fault diagnosis of rotating equipment, it usually carries out comprehensive analysis, mainly based on vibration analysis method with the above corresponding methods as well as the process and operation parameters.

27.2 What Is the Fault Diagnosis of Wind Turbine?

Fault diagnosis and condition monitoring are two different concepts. Condition monitoring is to monitor the equipment's operating parameters and development trend and to collect the key data, while fault diagnosis is to judge its working condition when the unit deviates from normal operation state and puts forward the corresponding processing methods. Condition monitoring is the foundation, prerequisites and necessary means of fault diagnosis. And fault diagnosis is to comprehensively use monitoring data and information to make decisions.

Fault diagnosis of wind turbine belongs to the category of the fault diagnosis of rotating machinery. Its content mainly includes fault diagnosis of the main shaft, gear box and generator on the main drive chain of wind turbine. The fault diagnosis technology of wind turbine actually refers to the technology using certain technologies to judge the state of wind turbine, diagnose and assess wind turbine's faults and its development and change without disintegration of the wind turbine, according to the accumulation of human experience and data. If we can accurately diagnose and estimate the wind turbine's faults and its developments and changes, then we can develop the optimal maintenance strategies.

For condition monitoring and fault diagnosis system of wind turbine, ideally, it should monitor and diagnose all potential failures of wind turbine. It should raise the alarm to the owners or operators before components of the wind turbine malfunction, give maintenance recommendations for faults of wind turbine, and also command system to stop operation in case of sudden major fault.

Specifically, the wind turbine should be able to collect enough wind turbine's state information. It should include the online monitoring of sufficient sensors, such as vibration sensor, displacement sensor, temperature sensor, pressure sensor, acoustic sensor, etc., and also include offline detection of the oil analysis of lubricating oil, blades testing, leaf coloring, ultrasonic detection and so on. It should arrange corresponding sensors on components of wind turbine with potential faults in such a way that installation of sensor should try to make the fault signal passed directly to the sensor without blocking and spacing, and number of sensors should make sure the enough signal is collected. The signal should be identified after obtained; and such identification is to compare signal to be checked with sample signal, and the result, which is above or below the desired target range, is considered to be a failure. There are some difficulties in the alarm limit set, since setting accurate alarm value requires both experience of unit operation and running values of unit when each working condition is normal. In the ideal diagnostic system, the setting of alarm value is important and difficult as well. When all kinds of collected signals exceed the alarm value, the collected signals should be processed and analyzed; and the methods for analysis and processing may vary according to different signal. Fourier transform or wavelet transform could be used for vibration signals and electrical signals. After being processed, the signal is diagnosed with manual work or computer intelligence. Finally, after the decision-making process, to control and operate the wind turbine correctly, for example temporary maintenance or shutdown maintenance, in order to protect wind turbine and wind farm as well as the safe operation of power grid. In addition, the system also provides the function of remote alarm and remote access. Owners can enter the fault diagnosis system of the wind turbine through the Internet or LAN to monitor the operation of each wind turbine on the wind farm and receive alarm signal from the system through the web mail or other means.

27.3 Why to Do Condition Monitoring and Fault Diagnosis?

With the installation and grid connection of large-scale wind turbines, due to special geographical location where wind turbines are impacted constantly by the harsh natural environment condition as well as SCADA system failing to sound the early alarm on faults of the wind turbine, many potential problems in the wind turbine can not be found in time and wind turbines often operate with potential faults, which causes more and more wind turbines to malfunction one after another after being put into operation for several years (Meng et al. 2012) (Fig. 27.1).

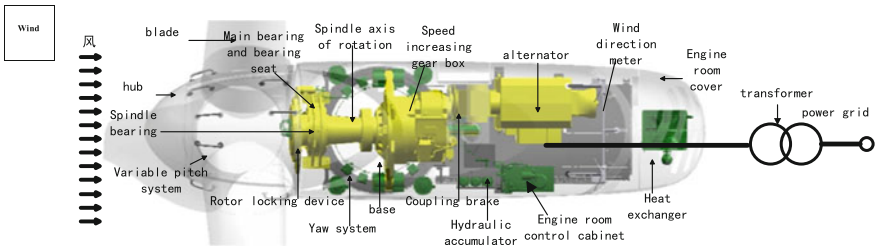

Fig. 27.1 Schematic diagram of the structure principle of wind turbine

As shown above, the basic principle of wind power generation is that the wind blows the wind wheel blades to make the wind wheel rotate, and then wind energy is converted to mechanical energy driving the generator to rotate through a series of transmission mechanism to realize the transformation of wind energy to mechanical energy to electrical energy, and ultimately into the power grid for the use of electricity load. Grid-connected wind turbine is generally horizontal axis type, mainly comprising wind wheel, pitch, yaw, transmission, hydraulic, control, generator and safety systems; and each system may inevitably have various faults due to long operation and aging. In order to find out which parts are the most prone to failure, which components' failure cause the longest shutdown, and what failures have the biggest impact on wind power production, many institutions at home and abroad have conducted long-term tracking and statistics.

Table 27.1 shows the statistical data of 2,151 sets of faulted wind turbine of three European countries by foreign research institutions. The statistical results show: the fault of wind power system is mainly located in electrical system, wind wheel system and gearbox, and the fault of gearbox is the one resulting in the longest shutdown of wind turbine. Long shutdown means great economic losses. Data shows that an unplanned one-day shutdown of one set of 1.5 MW land wind turbine will cause about 5,000 yuan economic losses, and the cost may hit around 1 million yuan for the use of large lifting and transporting equipment and sending personnel to dismantle, transport and fix in case of failure of gearbox of wind turbine; and the cost of dismantling of offshore wind power will be higher since it may need to use ships, cranes and even helicopters. Thus, this problem absolutely can not be ignored by enterprises. It is also fully explained that the gearbox-centered transmission system fault has the greatest impact on the electric power production, which should be paid great attention to (Meng et al. 2012).

The function of condition monitoring system is to monitor continuously key components of wind turbine and carry out pre-diagnosis, so that the operator could do corresponding maintenance to avoid significant losses.

The related studies also show that the on-line condition monitoring and fault diagnosis of transmission system through collecting vibration signal data, can help operation and maintenance personnel on the wind farm to grasp the health status of the transmission system of the wind turbine in time and reduce unnecessary on-site inspection work to achieve active maintenance of wind turbine, which not only

Table 27.1 Statistical data of wind field failure in Germany, Finland and Sweden

Country	Germany	Finland	Sweden
Average number of faults in wind turbines (annual)	0.4 times/year	1.4 times/year	2.4 times/year
Average down time of wind turbine (annual)	52 h/year	237 h/year	149 h/year
Average down time caused by single type fault	170 h/year	172 h/year	62.7 h/year
Components in which most of the fault occurred	1 Electrical system	1 Hydraulic system	1 Electrical system
	2 Sensors	2 Wind wheel	2 Control system
	3 The wind wheel	3 Gear box	3 Hydraulic system
Components in which multiple faults cause the longest down time	1 Transmission chain	1 Gear box	1 Generator
	2 Yaw system	2 Wind wheel	2 Gear box
	3 Gear box	3 Structure	3 Transmission chain
Components in which single failure causes the longest downtime	1 Gear box	1 Gear box	1 Generator
	2 Control system	2 Wind wheel	2 Gear box
	3 Electrical system	3 Hydraulic system	3 Transmission chain

ensures the safety of the wind power production, but also greatly reduces the cost of operation and maintenance, having a positive role in promoting the benign development of the wind power industry, either from an economic perspective (reducing maintenance costs) or from a security perspective (reducing accident rate). Therefore, the condition monitoring and fault diagnosis of the transmission system of the wind turbine is an urgent need for the development of wind power industry, and it is also an inevitable choice, which is of great practical significance to improve the efficiency of the wind turbine and the reliability, safety and economy of the wind power operation.

27.4 Arrangement of Measuring Points for Condition Monitoring

Reasonable choice of position of condition monitoring of wind turbine is of vital importance to ensure that the effective signal can be collected to reflect the vibration of equipment. According to the actual operating conditions of the wind turbine at home and abroad, usually the gear and bearing of gearbox, as well as generator and its bearing are more prone to failure, therefore, in conventional monitoring, the front and rear bearings of the gearbox and the front and rear bearings of the generator could be chosen as measurement points, respectively. From the perspective of signal band, different faults and frequency bands have different

sensitivity in the test direction, therefore it should collect low frequency signal (5 times more than power frequency) of rotating machinery from vertical, horizontal and axial 3 directions; for high frequency signal (above 1 kHz), due to less sensitive to the direction, it could only detect the vertical or horizontal direction. This is because the direction of the low frequency signal is stronger, while the high frequency signal is not sensitive to the direction.

To carry out accurate vibration analysis, the following test point selection principle must be followed: firstly, it should be able to reflect the real vibration of parts; secondly, it should be as close as possible to bearing area of bearing, and ensure that there is a solid metal between the bearing and the sensor; thirdly, there should be enough contact area between installing position and vibration sensor to ensure the sensor is installed firmly, not easy to fall off. The mechanical transmission part of the wind turbine is mainly composed of the spindle, gear box, and generator. Place vibration acceleration sensors on the bearing of each part to collect their acceleration signal, and also install rotating speed sensor on the blade wheel of the main shaft to collect rotating speed signal of the wind turbine.

For the readers' reference, Table 27.2 shows positions of state monitoring and locations of corresponding measuring points finally determined by our team, through going to the site to inspect the monitored object repeatedly, according to the "Guidelines for the Vibration State Monitoring of Wind Turbine" set by China's National Energy Administration.

27.5 Fault Diagnosis Methods

Because of the complexity of MW class wind turbine, its fault signal is nonlinear, time-varying and uncertain. The accurate and effective mathematical model can not be obtained in the general diagnosis of various faults of wind turbines. There are kinds of common fault diagnosis methods, including the traditional diagnosis methods, mathematical diagnosis methods and intelligent fault diagnosis methods. Most of the traditional diagnosis methods are based on the data analysis of the state monitoring technology, and are often used in combination with other methods to diagnose the fault. The mathematical diagnosis method mainly includes pattern recognition, time series model diagnosis based on probability and statistics, fault diagnosis based on distance criterion, fuzzy diagnosis, grey system diagnosis, fault tree analysis, wavelet analysis, chaotic analysis and fractal geometry. Intelligent fault diagnosis is the fault diagnosis technology based on artificial intelligence theory, and is applicable to the system with the characteristics of complexity, nonlinearity etc.; at present, it is widely used in the fault diagnosis of various fields.

This section mainly introduces a fault diagnosis method with good diagnosis effect: based on the combination of improved wavelet packet algorithm and demodulation and modulation signal of envelope spectrum analysis method.

The operation characteristics of wind turbine, such as random, intermittent, and variable load, make fault vibration signal of wind turbine transmission system

Table 27.2 Arrangement of measuring points and sensors

Parts	Measuring point	Installation location	Sensor	Frequency range (Hz)	Sensitivity (mv)
Main shaft	1	Bearing housing of front bearing of main shaft	Low frequency acceleration sensor	0.1–7000	500
	2	Bearing housing of rear bearing of main shaft	Low frequency acceleration sensor	0.1–7000	500
Gear box	3	Bearing housing of low speed end box of gear box	Low frequency acceleration sensor	0.1–7000	500
	4	Box parts of one-stage planet gear of gear box	High frequency acceleration sensor	0.5–14000	100
	5	Box parts of two-stage planet gear of gear box	High frequency acceleration sensor	0.5–14000	100
	6	Box parts of fixed axis gear train of gear box	High frequency acceleration sensor	0.5–14000	100
	7	Bearing housing of high-speed end box of gear box	High frequency acceleration sensor	0.5–14000	100
Generator	8	Front bearing housing of generator	High frequency acceleration sensor	0.5–14000	100
	9	Rear bearing housing of generator	High frequency acceleration sensor	0.5–14000	100

having the characteristics of non-stationary, nonlinear and complex modulation components, leading to the mining and extraction difficulty of fault characteristic quantity of wind turbine transmission system, and making it difficult to identify the fault types and to determine the specific fault location. Wavelet transform can automatically expand and contract video window, has good time and frequency localization ability with excellent time-frequency resolution, and is considered to be the ideal tool to analyze the non-stationary signal. The envelope spectrum analysis method of modulated and modulated signal, which is now widely used, is an

effective method for diagnosis of gear and bearing fault of transmission system; the core of realizing envelope spectrum algorithm is the band pass filter and Hilbert transform; and the effect of the band-pass filter determines the effectiveness and accuracy of the results of the envelope analysis (Wang 2014b). The multi-resolution of wavelet packet makes it have the function of band pass filter, so the combination of the wavelet packet and envelope spectrum is suitable for the fault diagnosis of non-stationary signal of transmission system of wind turbine. However, due to the non ideal characteristics of high and low pass decomposition and reconstruction filter of the wavelet, if the fault signal of transmission system of wind turbine is directly decomposed and reconstructed by wavelet packet, a serious band confusion will be generated, and using the disorder node as the components of the signal to conduct fault diagnosis will lead to error diagnosis conclusion. Therefore, in order to improve the accuracy of the fault analysis of the transmission system of the wind turbine by using the envelope spectrum algorithm of wavelet packet, aiming at the defects existing in the traditional wavelet packet transform, this paper presents an analysis method by combining improved wavelet packet with envelope spectrum to identify fault feature. The basic idea of the analysis method is to decompose and restructure the fault signal of gear and bearing of transmission system of wind turbine by using the improved wavelet packet at first; then select the appropriate node signal to realize band pass filtering; thirdly, conduct envelope demodulation on the selected node signal with Hilbert transform; fourthly and finally analyze the demodulated signal (Shen et al. 2014). All programs in this paper are written in M language based on MATLAB platform.

Wavelet packet algorithm is developed from wavelet analysis theory, but can provide a very high resolution compared with wavelet decomposition, which is a more precise time-frequency decomposition method. Its greatest feature is that it decomposes not only the low frequency sub band of the signal, but also the high frequency sub band of the signal further, so it can provide a more precise signal decomposition than wavelet transform. Wavelet packet meet the requirements of people for further improvement of the spatial resolution. On some occasions where a higher frequency resolution is required, wavelet packet has the incomparable advantages compared with the wavelet. In addition, wavelet packet is self-adaptive, and can select self-adaptively frequency band matched with signal spectrum according to signal characteristics. The excellent characteristics can enable people to find the most suitable optimal basis for signal analysis. Therefore, wavelet packet has a more excellent, wide range of practical application value.

In the division of binary frequency bands of wavelet, the wavelet filter theoretically should have an ideal cut-off characteristic, and signal should be continuously halved and marked on the specified node according to wavelet filter; however, wavelet filter in fact does not have the absolutely ideal cut-off characteristics because the frequency window energy of the wavelet function is not concentrated enough (Shen et al. 2014), the consequence of which is that each node contains an adjacent sub band component with the phenomenon of cross frequency domain and amplitude distortion in accordance with the law of frequency division of binary scale. On the other hand, in the transformation process, the operation of dot

interlacing sampling and zero insertion are used; after the completion of the convolution operation of the wavelet filter, the high and low frequency nodes contain frequency components each other and high frequency node component and high frequency node signal contained in the low-frequency nodes after dot interlacing sampling will cause the sampling frequency of each high frequency sub band signal is twice less than signal frequency and does not satisfy the Shannon sampling theorem, resulting in frequency aliasing and frequency band interleaving; conducted in the process of wavelet reconstruction, the operation of dot interlacing zero insertion makes frequency aliasing and frequency band interleaving more serious due to the formation of the image portion of the true frequency component.

When the discrete wavelet packet transform is used to decompose and extract the coefficient of wavelet packet node of signal, it can be used to store and recover signals since frequency band aliasing and interleaving have no effect on this; but it will be quite dangerous if directly using the decomposed wavelet packet coefficients as components in a certain frequency range of the original signal to conduct fault diagnosis, because the frequencies of these components are not all true, and the more complex the signal components are, the more serious the frequency disorder in the results are; and using the wrong node as the components of the signal to conduct fault diagnosis analysis may result in wrong diagnosis, and the adverse effects are immeasurable if using such conclusion to guide the practical engineering application.

The result of the frequency band confusion is closely related to the basic operation of wavelet packet and the wavelet packet's six basic operations, including the convolution operation of high pass filter with wavelet decomposition, convolution operation of low pass filter with wavelet decomposition, convolution operation of high pass filter with wavelet reconstruction, convolution operation of low pass filter with wavelet reconstruction, operation of dot interlacing sampling and zero insertion, are basically responsible for this error. In order to improve or solve the inherent defects of wavelet packet algorithm, we have made such an idea: after the convolution operation, if we filter the operation results according to the ideal frequency band range, filter out the extra frequency band components contained in each sub node, and then can eliminate the confusion of frequency components. In addition, as we know that dot interlacing sampling and zero insertion are two reverse operations and the adverse effects of the two operations are reverse as well, we can use the reverse folding of these two operations, then there will not be folded false signals in the reconstructed signal.

Accordingly, this section, starting from six key factors leading to frequency band disorders in the basic operations of the wave packet, improves the wavelet packet algorithm. The core idea to realize the filter in this section is to use the decomposition of fast Fourier transform (FFT) and synthetic action of fast Fourier inverse transform (IFFT) as well as the reverse folding of dot interlacing sampling and zero insertion to eliminate the confusion of frequency band. Compared with the conventional wavelet packet transform, the biggest difference of the improved node-restructured wavelet packet transform is the addition of the two filter operators; the two operators essentially carry out the time frequency conversion of signals

with the help of time and frequency domain transfer function of fast Fourier transform, then set the frequency beyond frequency band to zero after transform according to the division theory of wavelet packet frequency band to complete the filtering, finally convert the frequency domain signal back to time domain signal by using the fast Fourier inverse transform (Wang 2014a). We call the two operators C and D operation, respectively. For the definition of C and D operation, and the specific implementation process of the improved algorithm, please refer to the conclusion of Shen et al. (2014).

When damages occur in the change gear, rollaway nest of the bearing and rolling body in the mechanical transmission system, periodic impact caused by faulty components has broadband characteristics; the impact often causes resonance in some of components of the system; and the periodic impact can modulate the resonance signal. Fault signal is expressed in the frequency domain as sideband structure, centered on the resonance frequency and spaced by characteristic frequency near the center. The sideband is created, when the impact caused by the failure modulates the resonant amplitude of the signal. In order to analyze the complex side band frequency near the resonance frequency center in the spectrum of frequency, we need to make resonance demodulation of the side band to obtain the envelope curve of the modulation signal, do the Fourier transform of the envelope curve, draw the frequency diagram, and identify the modulation frequency that may contain fault frequency in the spectrum.

At present, the most commonly used method for demodulation of this kind of side band modulation signal is Hilbert resonance envelope demodulation. The principle of this method is to use Hilbert transform to construct the analytic signal, then the amplitude of the analytic signal is the envelope signal. And let the signal f (t) be a modulation signal, then its Hilbert transform is defined as the formula (27.5.1):

$$f(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(\tau) / t - \tau d\tau \tag{27.5.1}$$

The analytic signal H (T) of this constructed signal is the formula (27.5.2):

$$h(t) = f(t) + j f_h(t) \tag{27.5.2}$$

The amplitude envelope a (t) of the analytic signal is calculated according to the formula (27.5.3):

$$a(t) = |h(t)| = \sqrt{f(t)^2 + f_h(t)^2} \tag{27.5.3}$$

Do Fourier transform of the amplitude envelope and then draw the spectrum diagram, i.e. envelope spectrum.

The transmission system of wind turbine has a huge gearbox. When gear box is working, the rotational speed of high speed end gear and bearing may reach up to 1,000 rpm. When damages occurs in the variable speed gear, rollaway nest of bearing and rolling body in the transmission system, a great impact energy will be created if the damaged gears continue to engage and the damaged bearing parts continue to operate. The impact energy will arouse the inherent high frequency vibration of bearing and gear. Fault signal is expressed as pivot turn frequency, or composite signal of low frequency fault of bearing modulating high-frequency meshing frequency or natural vibration frequency through frequency. It is difficult to identify the modulation frequency in the spectrum after the FFT transform. Therefore, it is very necessary to apply envelope spectrum analysis to demodulate the modulated signal in the wind power fault analysis. But, on the other hand, the premise of realizing the envelope spectrum algorithm is to choose suitable high-frequency frequency band, that is to say, it is necessary to carry out band pass filter of signal in advance; and the effect of band pass filter determines whether the results of the envelope spectrum analysis are effective. The wavelet itself possesses excellent characteristics of band-pass filtering, and wavelet packet has more powerful multi-resolution than Marat wavelet; so the effect of using wavelet package instead of wavelet for envelope spectrum analysis in the previous segment is better. The improved node-restructured wavelet package algorithm proposed in this paper eliminates the defect of frequency band disorder, enhances the resolution of wavelet packet, and has a more accurate frequency resolution compared with the traditional wavelet packet. Therefore, the improved node-restructured wavelet packet, in combination with Hilbert envelope spectrum, is very suitable for the fault diagnosis of such non-stationary signal as transmission system of wind turbine. Based on this, this paper will elaborate the fault diagnosis method by combining the improved node-restructured wavelet packet with envelope spectrum. The principle of this method is to use the improved wavelet package algorithm with the defects of frequency band disorder eliminated to filter out the frequency band, of which the center frequency is not equal to high-frequency meshing frequency or natural vibration frequency, separate out high frequency restructured frequency band containing the modulation frequency, demodulate them through Hilbert envelope to filter out high-frequency natural vibration frequency, and ultimately demodulate out the low-frequency fault frequency. For the implementation flow of the fault diagnosis method and the specific description of the method model, please refer to the conclusions of Shen et al. (2014).

The fault diagnosis method of the improved wavelet packet combined with envelope spectrum introduced in this section has achieved good results in the experimental verification both in MATLAB simulation platform and in the trouble experiment set of transmission system of wind turbine independently developed by our team.

27.6 Condition Monitoring and Fault Diagnosis System of Wind Turbine

Online condition monitoring and fault diagnosis system of wind turbine has integrated multiple functions, such as signal acquisition, online monitoring and signal analysis; and it could realize online monitoring of vibration, temperature, pressure and electrical parameters of the transmission system and generator system, and analyze and process the collected data variously to correctly locate the faults of each system.

Our team has studied and developed the condition monitoring and fault diagnosis system of large-scale wind turbine, through the field research into 3 MW wind turbine of a certain group company, according to the fault types and characteristics of the key components of transmission system of the wind turbine as well as the actual demands of the users. In order to make readers have a systematic and comprehensive understanding of the condition monitoring and fault diagnosis system of wind turbine, this section will discuss the design and application of the condition monitoring and fault diagnosis system of the wind turbine by taking the specific design and field testing of the system as a carrier.

27.6.1 The Whole Design of the Condition Monitoring and Fault Diagnosis System for Large Scale Wind Turbine

As shown in the figure, the system is mainly composed of three parts: data acquisition and analysis system (including vibration sensors and data acquisition equipment), on-site monitoring system, and remote monitoring system (Fig. 27.2).

Among them, the data acquisition and analysis system is responsible for collecting, processing, intelligent automatic analysis and automatic diagnosis of the status signal of monitored components of each wind turbine synchronously in real time, remotely sending the raw data and analysis results to on-site monitoring system and accepting the remote control and parameter setting of on-site monitoring system; on-site monitoring system is responsible for remotely controlling the data acquisition and analysis system, showing characteristic curve diagram of the actual operation state of monitored equipment based on status data sent back from the data acquisition and analysis system, raising real-time automatic alarm to the fault wind turbine and its fault location, and storage and transmission of fault raw data collected by the data acquisition system and results of diagnosis and analysis; and the remote monitoring system adopts advanced fault diagnosis theory and method to analyze and diagnose the equipment condition of the wind turbine; when receiving alarm information or help from a wind turbine on a wind farm, professional and technical personnel will use professional analysis method to analyze the failure,

Fig. 27.2 Condition monitoring and fault diagnosis system of large scale wind turbine

locate the failure, predict the development trend of the failure and give professional analysis report and suggestion which will be sent back to the on-site monitoring system to instruct the actual work on the site (Wang 2014b).

27.6.2 Data Acquisition and Analysis System

Sensor selection

Sensor is a device for monitoring the system to obtain the original signal; according to the purpose of monitoring the vibration condition of the equipment of wind power generator and the configuration of the monitoring system, choosing vibration sensor correctly and reasonably is a key step to complete the condition monitoring and fault diagnosis of wind turbine, which requires various characteristics of the selected sensors must meet the requirements of the monitoring system.

This section selects the sensor and shielding signal transmission line as shown in Fig. 27.3 to make the further elaboration.

The specific parameters of the sensor are shown in Table 27.3.

Selection of monitoring points

Reasonable choice of the monitored components of wind turbine is very important to ensure that the effective signal reflecting the vibration of equipment could be collected. Since, in the fourth section of this chapter, we have elaborated the selection and arrangement of monitoring points, this section will not repeat them.

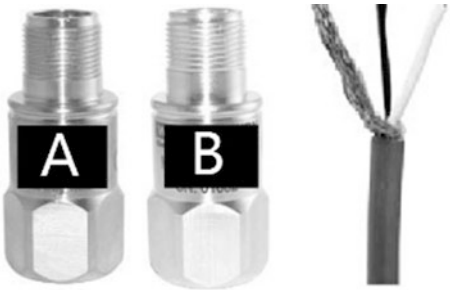

Fig. 27.3 A type and B type vibration sensor and signal transmission line of a company

Table 27.3 Vibration sensor parameters

Model	A	B
Sensitivity	100 mV/g	500 mV/g
Sensitivity error	±10 %	±5 %
Frequency response (±3db)	0.5 Hz–14 kHz	0.2 Hz–14 kHz
Frequency response (±10 %)	0.7 Hz–9 kHz	0.5 Hz–9 kHz
Frequency response (±5 %)	1.0 Hz–7 kHz	0.7 Hz–5 kHz
Spectrum noise @10 Hz	7 μg/$\sqrt{\text{Hz}}$	2.5 μg/$\sqrt{\text{Hz}}$
Acceleration range	80 g peak value	10 g peak value
Resonance frequency	30 kHz	30 kHz
Temperature range	−50 °C to 120 °C	−50 °C to 120 °C
Size (six head × length):	18 mm × 45 mm	22 mm × 53 mm

Design of acquisition instrument and algorithm design of fault diagnosis software Since the output speed of Megawatt-class wind turbine is more than 1,000 rpm after being accelerated by speed increaser, the sampling frequency required for measuring the high-frequency vibration of the wind turbine component is up to tens of thousands of Hertz when monitoring its vibration; such a high sampling frequency leads to a huge amount of data for instantaneous sampling, makes real-time analysis and transmission more difficult and is also a great challenge to realize the on-line and real-time of the whole system. In order to realize the real-time and synchronous acquisition and analysis of multi-point mass data and ensure real-time requirements of the system, we participate in the design of a high performance data acquisition instrument, which has the function of real-time high-speed synchronous acquisition, analysis and transmission of massive data, determines the location and severity of the fault in accordance with the threshold alarm index given by the on-site monitoring center and gives the diagnosis results by real-time, synchronous and parallel

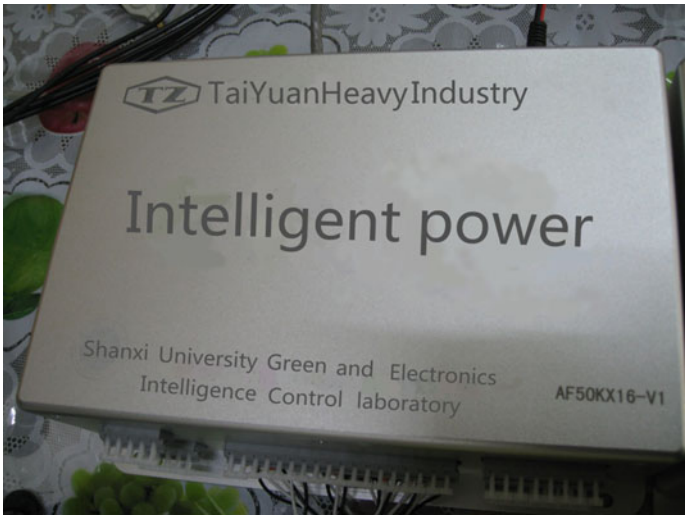

Fig. 27.4 High performance data acquisition instrument

computing of the collected time and frequency domain index of multi-channel data, and remotely send the diagnosis result and the original data to the on-site monitoring center in real time to complete the functional requirements for monitoring and diagnosis at last. The independently designed acquisition instrument is shown in the following Fig. 27.4 (Wang et al. 2013a).

Because the monitored components are different and the fault diagnosis algorithm of each component is different as well, this section designs different fault diagnosis algorithms for different monitored components. Then, taking the vibration monitoring of the main shaft as an example, we will simply describe the algorithm design of the fault diagnosis system software. The algorithm software for monitoring the main shaft in this section specifically includes 8 time domain analysis indexes and 4 frequency domain indexes; time domain indexes are: mean value, effective value, peak value, waveform index, peak index, pulse index, margin index, and kurtosis index; and frequency domain indexes are: bearing inner ring, bearing outer ring, bearing holder and fault characteristic frequency of ball bearing (Wang 2014b). For the calculation formula of time domain index given in the algorithm, please refer to the conclusion given by Shen (2014).

The fault diagnosis algorithm for monitoring and alarming the component fulfils the function of the analysis and diagnosis of fault based on the above indexes. This part of work is completed by the designed acquisition instrument. The detailed algorithm process will not be repeated due to the project's patent.

27.6.3 Field Monitoring System

Operating environment and architecture of the field monitoring system

In order to facilitate the data exchange between remote monitoring system and field monitoring system and make it easy for remote experts to provide professional remote diagnostic services, after fully taking into account the actual needs of the field and remote monitoring, the system structure of B/S architecture is adopted and designed. Field monitoring system, through field Browser monitoring interface, controls front-end data acquisition and analysis system to work, displaying real-time operating status of each wind turbine on the wind farm. The server of the field monitoring system uses SQL Server database to receive, send and store the original data, state information, analysis results and control commands.

Interface design of state monitoring and fault diagnosis system

According to the actual demand, the interface of the condition monitoring system with independent intellectual property rights is designed. The interface has the following functions: display the state of wind turbine in real time according to the data transmitted by the front end; display the status of the monitored components of wind turbine; display the curve diagram of condition monitoring indexes of monitored components of wind turbine; display the date and time and location of failure, diagnostic results, treatment recommendations and remote professional analysis report; display the curve diagram of the same monitored component of every wind turbine and real-time comparison; display the query of historic records; the function of query and export of fault history database; the remote control function of the data acquisition and analysis instrument; remote parameters and threshold setting function of data acquisition and analysis instrument; and user's right setting function. The following figure shows the interface of wind turbine alarm of the system. For more interface icons, please refer to the conclusion of Shen (2014). (Fig. 27.5).

27.6.4 Remote Monitoring System

The remote monitoring system designed in this paper realizes the field monitoring of wind turbine by being connected with the field monitoring system via internet, through remote access to the database of the field monitoring system. The main function of the remote monitoring center is that when receiving the fault alarm information of a wind turbine on a wind farm or field staff asking for help, remote experts and technical personnel use professional analysis method to analyze the fault of wind turbine, locate the fault and predict the development trend of the fault, and give professional analysis report and suggestion which will be sent back to the field monitoring system to assist the field staff to complete the actual work.

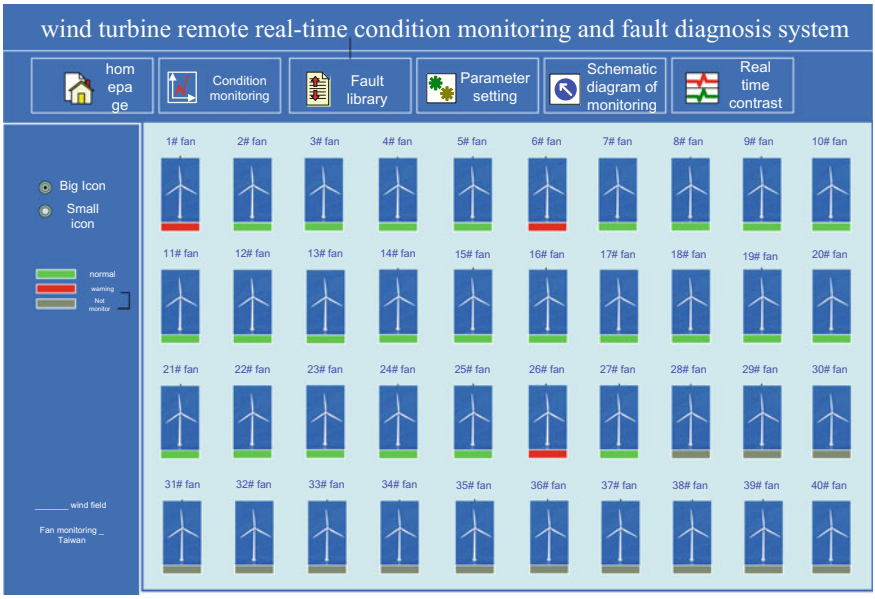

Fig. 27.5 Wind turbine alarm

The system has fully demonstrated its non-negative validity and reliability in the field test. For the specific field installation test procedure and test result analysis of the system, please refer to the content of Shen (2014).

27.7 Summary of This Chapter

Wind power load, wind instability and a large number of interfering signals generated in the operation of wind turbine cause early fault signal of wind power to be weak time-varying non-stationary signal; because of the complexity of the unit structure, it is difficult to establish a linear relationship between fault and symptom, seriously affecting the effectiveness and accuracy of the traditional analysis method based on the vibration signal, which causes the actual application effect of the condition monitoring and fault diagnosis method of wind turbine is not good; therefore, it is urgent to improve the existing fault diagnosis methods or introduce a new analysis method. Starting from the basic contents of the condition monitoring and fault diagnosis of wind turbine, step by step, after elaborating the necessity of condition monitoring and fault diagnosis of wind turbine, selection of monitoring points and failure diagnosis methods, this chapter completely introduces the design method of the whole structure and function of the condition monitoring and fault diagnosis system of the wind turbine by taking the condition monitoring and fault diagnosis system of large-scale wind turbine developed by the team led by the

author as a case. The main content of this chapter is the latest research results of our team in this field in recent years. But, considering the condition monitoring and fault diagnosis technology of wind turbine covers a wide range of content and the complexity of early fault of wind power, any method has its limitations and imperfections, the condition monitoring and fault diagnosis method and system of wind turbine introduced in this chapter has no exception. Therefore, if there is anything wrong with this chapter, welcome your comments and corrections.

References

Guo, D. (2012). *Research on condition monitoring and fault intelligent diagnosis system of wind turbine*.

Junjun, L., Zhengqiu, W., Xunqiong, T., & Bo, C. (2011). Review of wind power generation and its technical development [J]. *Power Construction,* 64–72.

Kang, D., & Liqi, J. (2000). Study on the limitations of application of demodulation analysis in mechanical vibration analysis [J]. *Mechanical Science and Technology*, 722—725.

Meng, E., Guo, D., & Wang, L. (2012). Design and implementation of condition monitoring and intelligent fault diagnosis system of wind turbine. *East China Electric Power, 40*(3), 0507.

Raghu, P.P., & Yegnanarayana, B. (1998). Supervised texture classification using a probabilistic neural network and constraint satisfaction model [J]. *IEEE* Transactions on *Neural Networks*, *9* (3), 516-522.

Shen, J. (2014). *Research on condition monitoring and intelligent fault diagnosis system of large scale wind turbine*.

Shen, J., Wang, L., & Meng, E. (2014). Study on fault diagnosis of wind turbine of improved wavelet packet combined with PNN. *Renewable Energy, 32*(4), 412.

Spech, D. F. (1990). Probabistic neural networks [J]. *Neural Net works, 3*(2), 109–118.

Wang, L., et al. (2013c). Data acquisition system of wind turbine. In *State Intellectual Property Office of the People's Republic of China (patent).*

Wang, L., et al. (2014a). System and method for remote real-time state monitoring and intelligent fault diagnosis of wind turbine. In *State Intellectual Property Office of the People's Republic of China (patent).*

Wang, L., et al. (2014b). Design of wireless monitoring system of wind turbine. *Transducer and Microsystem Technologies, 33*(5), 0084.

Wang, L., & Shen, J. (2014). Research on fault diagnosis method of wind turbine transmission system based on improved wavelet packet and envelope spectrum. *Acta Energiae Solaris Sinica, 35*(9), 1771.

Wang, L., et al. (2012). Fault diagnosis of wind turbine based on improved wavelet combined with BP network. *Proceedings of the CSU-EPSA, 24*(2), 0053.

Wang, L., et al. (2013a). Research on data communication of wind power remote monitoring and diagnosis system. *East China Electric Power, 41*(1), 0148.

Wang, L., et al. (2013b). Research on grid-connected cooperative control of wind power based on multi-agent. *Renewable Energy, 31*(3), 0036.

Wang, L., et al. (2015a). Prediction of vibration characteristics of wind turbine of improved grey Elman neural network. *Renewable Energy, 33*(6), 0876.

Wang, L., Jiao, J., & Yin, S. (2015). Wireless condition monitoring system of yaw device of wind turbine. In *State Intellectual Property Office of the People's Republic of China (patent).*

Yang, J. (2005). *Wavelet analysis and its application in engineering [M]* (pp. 63–97). Beijing: Machinery Industry Press.

Yin, S., Jiao, J., & Wang, L., et al. (2013). Wireless monitoring system of wind farm. In *State Intellectual Property Office of the People's Republic of China (patent).*

Zhang, D. (2011). *MATLAB neural network programming [M]* (pp. 179–180). Beijing: Chemical Industry Press.

Chapter 28
Reducing the Life Cycle Environmental Impact of Buildings Following a Simulation-Optimization Approach

Joan Carreras, Dieter Boer, Luisa F. Cabeza, Marc Medrano, Laureano Jiménez and Gonzalo Guillén-Gosálbez

Abstract Identifying building designs with minimum cost and environmental impact is a fundamental topic in the transition towards a more sustainable residential sector. Energy efficiency strategies can reduce energy consumption in buildings, thereby decreasing their environmental impact without compromising comfort. Among these energy efficiency strategies, building insulation is particularly appealing due to its low cost and high potential energy savings. This study presents a systematic methodology for determining the optimal insulation thickness of the external surfaces of a building. Our approach, based on multi-objective optimization, minimizes simultaneously the cost and environmental impact associated with both the construction materials and the energy consumed over the operational phase of the building. The thermal loads are calculated by EnergyPlus, a building energy simulation program widely used by architects and engineers. The environmental impact is quantified following the life cycle assessment methodology and explicitly incorporated into the multi-objective model as an additional objective

J. Carreras · D. Boer (✉)
Departament d'Enginyeria Mecanica, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain
e-mail: dieter.boer@urv.cat

L. Jiménez · G. Guillén-Gosálbez (✉)
Departament d'Enginyeria Química, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain
e-mail: g.guillen05@ic.ac.uk; g.guillen05@imperial.ac.uk

G. Guillén-Gosálbez
Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College, South Kensington Campus, London SW7 2AZ, UK

L.F. Cabeza · M. Medrano
GREA Innovació Concurrent, Edifici CREA, Universitat de Lleida, Pere de Cabrera s/n, 25001 Lleida, Spain

M. Medrano
CESDA (European University College of Aviation), University Rovira i Virgili, Carretera del Aeropuerto s/n, Apartado de Correos 481, 43206 Reus, Spain

G.M. Kopanos et al. (eds.), *Advances in Energy Systems Engineering*,
DOI 10.1007/978-3-319-42803-1_28

to be optimized along with the cost. We applied our approach to a case study of a house-like cubicle, for which four potential European locations with different climate conditions and electricity prices were considered. Solutions that reduce around 40 % both, the cost and environmental impact, with respect to a cubicle without insulation were identified. Our systematic method is intended to assist decision-makers in the design of more sustainable buildings.

Keywords Multi-objective optimization · Life cycle assessment (LCA) · Modelling · Buildings · Insulation

Nomenclature

Abbreviations

IEA	International Energy Agency
MOO	Multi-objective optimization
LCA	Life cycle assessment
PDE	Partial differential equations
NSGA-II	Non-dominated sorting genetic algorithm-II
HVAC	Heating, ventilating and air conditioning
COP	Coefficient of performance
PU	Polyurethane
MW	Mineral wool
EI99	Eco-indicator 99
GLO	Average global impact
OF	Objective function
ACH	Air changes per hour

Variables

COST	Cost [€]
UCOST	Unitary Cost [€/kg]
M	Quantity [kg]
CONS	Consumption [kWh]
IMP	Impact [EI99 points]
UIMP	Unitary Impact [EII99 points/kg]

Indices

TOT	Total
MAT	Materials
EN	Energy

Sets

K Set of construction materials indexed by k
N Set of years indexed by n

Symbols

ir Electricity inflation rate (%)
x Decision variables
X Space of feasible solutions

28.1 Introduction

Nowadays the building sector represents 40 % of the total annual energy consumption worldwide (IEA 2008). Given its importance, many countries in the OECDE have dictated measures to reduce energy consumption in buildings. In March 2007, the European Parliament approved a binding legislation with several goals: to achieve a 20 % reduction in EU greenhouse gas emissions from 1990 levels; to increase the share of EU energy consumption produced from renewable resources to 20 %; and to improve the EU's energy efficiency by 20 % (Anon 2009). To achieve these key targets, several energy strategies can be adopted. Among them, building insulation appears as a promising one, since it decreases the cooling and heating demand without compromising comfort and it is applicable in both, new and refurbished buildings.

Today the current trend is to implement thicker insulation materials in order to reduce energy consumption. This strategy, however, may lead to sub-optimal solutions when one seeks to optimize the economic and environmental performance simultaneously. Indeed, thicker insulation does not necessarily imply less impact (neither less total cost), mainly because the impact embodied in the insulation material (as well as its cost) may be quite large (Blengini and Di Carlo 2010; Stephan et al. 2013).

Environmental and economic objectives tend to be conflicting goals in many engineering problems. Multi-objective optimization (MOO) is the prevalent method to solve problems with multiple objectives (Gebreslassie et al. 2009; Brunet et al. 2012; Antipova et al. 2014; Sabio et al. 2014; Grossmann and Guillén-Gosálbez 2010; Guillén-Gosálbez and Grossmann 2009; Carreras et al. 2014). Typically, the final result of a MOO problem consists of a set of Pareto optimal solutions, each achieving a unique combination of objective function values. From this set, decision-makers are asked to identify the best alternative according to their preferences.

Previous studies in building design focused on seeking optimal thermal insulation strategies for buildings considering (only) either the economic performance (Daouas 2011; Pan et al. 2012; Zhu et al. 2011) or the environmental impact (Audenaert et al. 2012; Asif et al. 2007) as separate unique objectives. Ozel (Ozel 2013b) showed that the optimal insulation thickness does not depend on the insulation layer location in the wall. Al-Sanea et al. (Al-Sanea et al. 2005) proposed a method to evaluate the optimum insulation thickness from economic data

(i.e. electricity price, cost of insulation material) and using as well the coefficient of performance of the HVAC equipment. In these studies, the optimal insulation thicknesses vary from 5 to 16 cm depending on the case study.

This chapter describes a method for the multi-objective optimization of buildings. The final goal is to find the optimal thermal insulation solutions that minimize simultaneously the environmental impact and the economic cost. The environmental impact is evaluated via life cycle assessment (LCA) principles, which quantify the damage caused in the manufacturing, operational and dismantling phases of the building under study (Menoufi et al. 2013; Singh et al. 2011; Ortiz et al. 2009; Cabeza et al. 2014). We show through numerical examples how our multi-objective optimization (MOO) approach can identify solutions with better economic and environmental performance, thereby assisting decision makers in the transition towards a more sustainable building sector.

28.2 Problem Statement

We are given a cubicle type building (details about the cubicle can be found in Sects. 28.4.1 and 28.4.2) for which a set of different insulation materials and different thicknesses are available. The main objective is to find the insulation configuration that simultaneously minimizes the economic cost and the environmental impact of the building. In the analysis, we need to consider the cost and environmental impact of the construction materials, along with the cost and environmental impact of the energy consumed for heating and cooling.

28.3 Methodology

Our methodology relies on a simulation-based optimization framework that combines a model that predicts the building's performance with an optimization algorithm that seeks the optimal building design. The next section introduces the simulation software very briefly, while the following one describes the objective functions to be minimized.

28.3.1 Mathematical Model

Following the simulation-optimization approach mentioned above, the energy loads are calculated using EnergyPlus v.8 (EnergyPlus 2015), a commercial simulator that models energy and water use in buildings. EnergyPlus contains a system of partial differential equations (PDE) that enables the calculation of the energy consumed by the building for given values of the input parameters, including

climatic conditions, thermal properties of the materials and layout of the building. Details on this software can be found elsewhere (EnergyPlus 2015).

28.3.2 Objective Functions

In the ensuing sections, we present the equations used to determine the environmental and economic performance of each thermal insulation alternative.

28.3.2.1 Economic Indicators

The economic performance accounts for the cost of the construction materials and the cost of the electricity consumed for heating and cooling during the operational phase of the building. The final goal is then to reduce the total cost of the building ($COST^{TOT}$) (Ozel 2013b; Kaynakli 2012; Ozel 2013a; Yu et al. 2009), which is calculated by Eq. (28.1).

$$COST^{TOT} = COST^{MAT} + COST^{EN} \tag{28.1}$$

where, $COST^{MAT}$ accounts for the cost of the construction materials, whereas $COST^{EN}$ denotes the cost of the electricity consumed during the operational phase of the building.

The materials required for the construction of the building are presented in Table 28.1, which displays as well standard values of mass and cost for a standard cubicle with one cm of insulation in all of its surfaces. The cost and thermo-physical properties of the insulation materials are presented in Table 28.2. Data were retrieved from the LIDER (LIDER 2009) and ITeC (BEDEC 2011) databases. The total materials cost associated with the construction of the cubicle is quantified as shown in Eq. (28.2).

Table 28.1 Inventory list of the cubicle construction materials and their corresponding cost

Component	Used mass (kg)	Cost (€)
Brick	5,456	287
Base plaster	518	43
Cement mortar	608	30
Steel bars	262	157
Concrete	1,240	44
In-floor bricks	1,770	62
Asphalt	153	317
PU (1 cm)	20	79
MW (1 cm)	18	55

Table 28.2 Properties of the insulation materials

Insulation material	Density (kg/m^3)	Thermal conductivity (W/(m K))	Specific heat (J/(kg K))	Cost (€/m^3)
Polyurethane	45	0.027	1,000	175
Mineral wool	40	0.04	1,000	122

$$COST^{MAT} = \sum_{k \in K} UCOST_k^{MAT} \cdot M_k \tag{28.2}$$

where $UCOST_k^{MAT}$ is the unitary cost per kilogram of material k, and M_k is the mass expressed in kilograms of material k used in the building (i.e., kg of mortar).

The energy consumed to fulfill the cooling and heating requirements over the operational lifetime of the building is obtained via Eq. (28.3):

$$COST^{EN} = \sum_{n \in N} CONS_N \cdot UCOST^{EN} \cdot (1 + ir)^n \tag{28.3}$$

where $CONS_N$ is the amount of energy consumed (expressed in kWh) for heating and cooling in year n (note that it is assumed that this annual energy consumption remains constants over the years), $UCOST^{EN}$ is the unitary cost of one kWh of electricity, and ir is the yearly increase in the energy cost.

28.3.2.2 Environmental Indicators

To quantify the environmental impact of the building (due to both, the generation of the energy consumed during its operational phase and the manufacture of the building materials), we follow the LCA methodology Eco-indicator 99 (Eco-Indicator 99 2000). Data from the EcoInvent database (Ecoinvent 2015) is used in the calculations. This indicator quantifies the impact in 10 specific categories, which are grouped into three damages (ecosystem quality, human health and resources depletion) that are further aggregated into a single score.

Hence, the total environmental impact (variable IMP^{TOT}) accounts for the impact of the construction materials (variable IMP^{MAT}), and the impact of the energy consumed for heating and cooling over the operational phase of the building (variable IMP^{EN}):

$$IMP^{TOT} = IMP^{MAT} + IMP^{EN} \tag{28.4}$$

The total impact of the materials for the construction of the cubicle (IMP^{MAT}) is calculated as follows:

$$IMP^{MAT} = \sum_{k \in K} IMP_k^{MAT} \cdot M_K \tag{28.5}$$

where IMP_k^{MAT} is the impact per kilogram of material k (note that the value of this parameter can be retrieved from the EcoInvent database for a wide range of materials (Ecocostsvalue 2014), and M_K is the corresponding quantity (expressed in kilograms) of raw material k.

To translate the energy consumed into environmental impact, we use data on the electricity production mix of the country where the building will be installed. It is important to clarify that the impact of energy generation differs across countries, while the impact of materials is assumed to be the same regardless of the location. The total impact of the energy consumed (IMP^{EN}) is calculated as follows:

$$IMP^{EN} = \sum_{n \in N} CONS_N \cdot UIMP^{EN} \tag{28.6}$$

where $UIMP^{EN}$ is the environmental impact per kWh of energy in each country and $CONS_N$ is the yearly consumed energy in period n (which is provided by the building design model).

Table 28.3 shows a list of the materials employed in the construction of the cubicle and their corresponding environmental impact. For illustrative purposes, Table 28.3 shows as well the environmental impact of a cubicle with one cm of insulation in all of its surfaces.

28.3.3 Solution Procedure

In mathematical terms, the EnergyPlus simulation model can be expressed as an explicit function $f^{MOD}(x_1, \ldots, x_j)$ of the decisions variables x_1 to x_j. This function contains two components (i.e. cost and environmental impact) whose values are determined after solving the simulation model for fixed values of the decision variables (i.e. insulation thicknesses). Hence, the multi-objective model can be expressed in compact form as follows:

$$\min_{x \in X}\left(COST^{TOT}, IMP^{TOT}\right) = \min_{x \in X}\left(f_1^{MOD}(x_1, \ldots, x_j), f_2^{MOD}(x_1, \ldots, x_j)\right)$$

where X is the space of possible solutions and x_1 to x_j are the decision variables. Note that the set X is defined by explicit constraints imposed externally as well as implicit constraints solved internally by the building simulator. Within the former group, we define inequality constraints imposing lower and upper bounds on the values of the decision variables:

Table 28.3 Environmental impact punctuation (EI99 point) of the materials implemented for the construction of the cubicle

Component	Used mass (kg)	EI 99 (Points/kg)	Total EI99 (Points)
Brick	5,456	0.0196	106.714
Base plaster	518	0.0126	6.552
Cement mortar	608	0.0147	8.939
Steel bars	262	0.0135	3.531
Concrete (m3)	0.577	18.8780	10.888
In-floor bricks	1,770	0.0160	28.237
Asphalt	153	0.0284	4.342
Disposal bricks	5,456	0.0028	15.078
Disposal plaster	518	0.0057	2.976
Disposal mortar	608	0.0062	3.798
Disposal concrete + steel bars	1,492	0.0042	6.203
Disposal in-floor bricks	1,770	0.0028	5.029
Disposal asphalt	153	0.0020	0.307
PU	20	0.3973	8.046
MW	18	0.1024	1.842
Disposal PU	20	0.0743	1.504
Disposal MW	18	0.0073	0.132

$$\underline{x_i} \leq x_i \leq \overline{x_i} \forall i = 1, \ldots, j$$

where $\underline{x_i}$ and $\overline{x_i}$ are the lower and upper bounds, respectively, on variable x_i.

There are many MOO methods available to solve multi-objective optimization problems (Das and Dennis 1998; Ehrgott 2005; Messac et al. 2003; Copado-Méndez et al. 2014). The solution of these models is given by a set of Pareto points that represent the optimal trade-off between the conflicting objectives considered in the analysis (Brunet et al. 2012; Gebreslassie et al. 2009). These Pareto optimal solutions cannot be improved simultaneously in all of their objectives without necessarily worsening at least one of them.

In mathematical terms, $x \in X$ is a Pareto optimal solution if there does not exist any $x' \in X$ such that $f_k(x') \leq f_k(x)$ for all k, and at $f_l(x') \leq f_l(x)$ for some l. If x' is Pareto optimal, then $z' = f(x')$ is called non-dominated point or efficient point.

In this study, without loss of generality, we solve the multi-objective model using multi-objective genetic algorithms. Hence, the simulation software is coupled with an optimization algorithm implemented in JEPlus+EA (2015) that minimizes the multi-dimensional objective function by modifying the values of the decision variables (see Fig. 28.1). This optimization software makes use of the multi-objective genetic algorithm NSGA-II (see details in JEPlus+EA 2015). Note that due to the nonconvex nature of the model, this approach cannot guarantee convergence to the globally optimal Pareto set. Our framework, however, is general

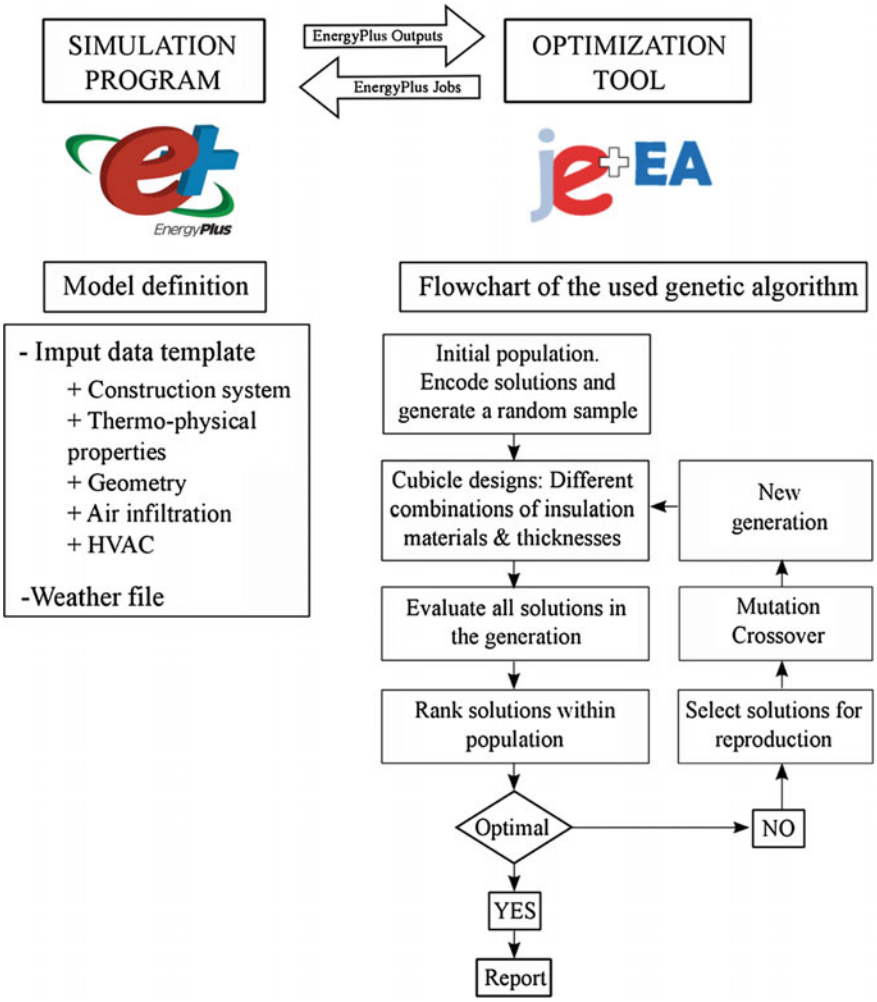

Fig. 28.1 *JEPlus+EA* optimization process coupled with *EnergyPlus*

enough to accommodate other optimization algorithms, in a way similar as we previously did in the past in other engineering problems (Brunet et al. 2012).

28.4 Case Study

28.4.1 Cubicle Description

To demonstrate the capabilities of the approach presented, let us consider a house-like cubicle like the ones installed in the experimental setting of the GREA

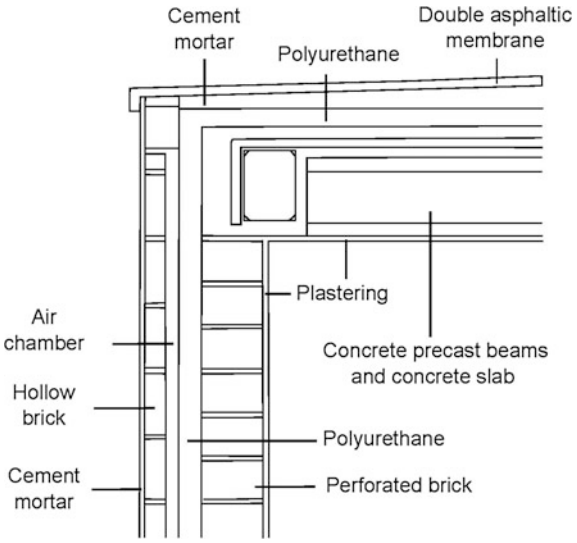

Fig. 28.2 Construction profile of the experimental cubicles

group located in Puigverd (Lleida, Spain). The cubicles have five plane walls of 2.4 × 2.4 × 0.15 m, but differ in the construction strategy followed. In the location of the experimental installation, there is a continental Mediterranean climate characterized by dry hot summers, moderate cold winters, and important daily temperature oscillations between day and night (Rivas-Martinez, S., Rivas Sáenz, S., Penas Merino n.d.).

Our cubicles behave as a conventional Mediterranean construction system (Fig. 28.2). The structure of the cubicle consists of four mortar pillars with reinforcing bars. The walls have six layers, an exterior cement mortar cover, a structure of hollow bricks, a 5 cm air chamber, the insulation layer, a perforated bricks structure, and an interior cover of a plaster plastering layer. In our analysis, we will consider cubicles that differ only in the insulation layer. The cubicle floor consists of a layer of concrete (3 × 3 m) with reinforcing bars, while the roof is composed of an exterior double asphaltic membrane with a cement mortar base that covers the insulation layer. The structural configuration of the roof is made of concrete precast beams and 5 cm of concrete slab, while the internal finish consists of a plaster plastering layer. A reference cubicle with no insulation is considered (Menoufi et al. 2012; Cabeza et al. 2010) for comparison purposes.

28.4.2 Model Specifications

For the HVAC, we consider a heat pump with a COP of 3. The internal set point temperature is fixed to 24 °C (Cabeza et al. 2010; Castell et al. 2013). No mechanical or natural ventilation is considered, and a fixed infiltration rate of 0.12

air changes per hour is assumed (DOE 2013). The cubicle is assumed to be empty (i.e. no internal mass neither human occupancy). The building lifespan is 20 years (Ozel 2014; Fokaides and Papadopoulos 2014). The construction materials are paid the first year of the time horizon. The electricity cost is increased by 5 % every year.

The insulation thicknesses fall in the range 1–30 cm (EURIMA 2014). Two insulation materials, PU and MW, are analyzed in detail. Different insulation thicknesses for the roof and the walls can be established (but the four walls should always show the same insulation thickness).

28.4.3 Considered Locations

Four European locations were considered according to the Köppen–Geiger Climate Classification (Kottek et al. 2006) (Table 28.4). This classification characterizes the climate conditions using three characters. The first character defines the main climate: A: equatorial, B: arid, C: warm temperate, D: snow and E: polar. The second character defines the level of precipitation: W: desert, S: steppe, f: fully humid, s: summer dry, w: winter dry, m: monsoonal. Finally, the third character describes the temperature: h: hot arid, k: cold arid, a: hot summer, b: warm summer, c: cool summer, d: extremely continental, F: polar frost, T: polar tundra. Data on the electricity impact was obtained from (EUROSTAT 2013), while the electricity cost was retrieved from (Ecoinvent 2015).

The computational framework took around 1900 to 2000 CPU seconds to generate around a hundred of Pareto solutions in each experiment (PU, MW) on a computer HP Compaq Pro 6300 SFF with an Intel Core Processor 3.30 GHz and 3.88 GB of RAM. Each experiment was limited to 200 generations with an initial population size of 10. Each calculation was repeated 10 times in an attempt to avoid locally optimal solutions.

Table 28.4 Climate type, electricity cost and electricity impact for the four locations

Locations	Climate	Electricity cost (€/kWh)	Electricity impact (EI99 points/kWh)
Lleida (Spain)	BSk	0.223	0.034
Dublin (Ireland)	Cfb	0.203	0.043
Athens (Greece)	Csa	0.156	0.089
Berlin (Germany)	Dfb	0.292	0.030

28.5 Results and Discussions

The optimal solutions for the different locations are presented in Fig. 28.3. Each point in the figures represents a cubicle with a specific insulation configuration. In all of the scenarios, we obtain the extreme economic and environmental optimal solutions and a set of intermediate Pareto solutions in between.

The best economic solution is achieved using PU regardless of the location. PU is more expensive than MW, but its thermal conductivity is lower, so its energy savings compensate for the extra cost. On the other hand, the environmental impact is minimized using MW, as the impact embodied in this material is significantly lower compared to PU. In all of the scenarios, the roof presents an insulation thickness from 1 to 3 cm thicker than the one required in the walls. Hence, to move from the extreme optimal economic solution to solutions with less environmental impact, the model decides to change the insulation material (from PU to MW) and to increase the insulation thickness. In all of the locations, with a marginal economic effort we can attain significant improvements from the environmental viewpoint (e.g. in Lleida, increasing the cost by 0.5 % leads to a reduction of 11 % in environmental impact).

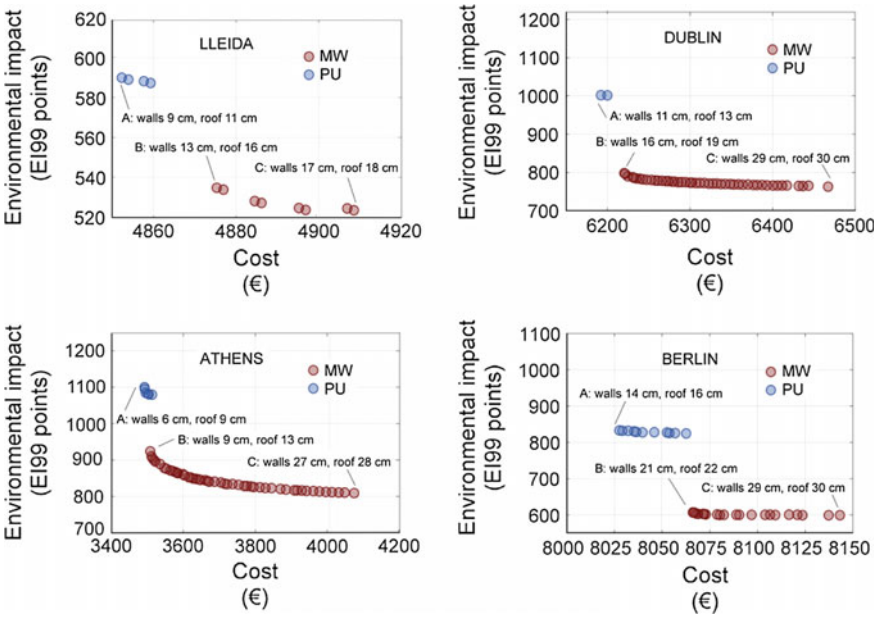

Fig. 28.3 Pareto set cost versus environmental impact (Eco-indicator 99) for the four locations. In the figures, A is the best economic solution, B is an intermediate optimal solution and C is the solution with minimum impact

Athens is the location with the lowest cost, mainly because its electricity price is the cheapest and its climatic conditions are less severe than in the other locations. Berlin leads to the most expensive solutions due to its high electricity cost and harsh climate conditions. Lleida shows the lowest environmental impact, as both the impact embodied in electricity and the climate conditions are moderate. Note that Athens, despite showing mild weather conditions (less energy required by the HVAC), has the highest environmental impact due the high impact embodied in its electricity mix.

Table 28.5 displays the improvements attained by points A (extreme economic optimal solution), B (intermediate optimal solution) and C (extreme environmental optimal solution) with respect to the base case (cubicle without insulation) in the four locations. Economic savings range from 22 to 50 % depending on the scenario. The minimum cost solution in Dublin shows the best improvements with respect to the base case. Berlin shows the best environmental impact improvement (45 % reduction), while Athens leads to the smallest (15 %). All intermediate optimal solutions (B) show significantly better environmental performance than the minimum cost solution.

Table 28.5 Comparison of the base case results and the best economic and environmental results for the four locations

	Cubicle model	Economic cost (€)	EI99 (eco-points)	Improvement (calculated with respect to the base case and expressed in %)	
				Economic	EI99
Lleida	Base case	8,602	916	0	0
	A	4,852	588	44	36
	B	4,876	530	43	42
	C	4,910	523	43	43
Dublin	Base case	12,317	1,253	0	0
	A	6,190	1,004	50	20
	B	6,221	798	49	36
	C	6,469	763	47	39
Athens	Base case	5,201	1,294	0	0
	A	3,493	1,100	33	15
	B	3,511	924	32	29
	C	4,076	809	22	37
Berlin	Base case	15,235	1,092	0	0
	A	8028	835	47	24
	B	8066	606	47	45
	C	8142	599	47	45

28.6 Conclusions

The population is becoming more aware of the implications of anthropogenic impacts on the environment and consequently starting to demand cost-effective products with lower environmental impact.

Following this general trend, this chapter has presented a systematic approach integrating building modeling, multi-objective optimization and life cycle assessment that identifies optimal building designs according to economic and environmental criteria simultaneously. To illustrate the capabilities of this methodology, a house-like cubicle was designed using our approach to select the optimal type and thickness of insulation material in four different European locations.

Numerical results show how significant economic and environmental improvements (win-win scenarios) can be obtained with respect to a base case (cubicle with no insulation) in all of the locations. The economic improvements range from 22 to 50 %, while the environmental impact is reduced between 15 and 45 %. Polyurethane shows better economic performance, while mineral wool has less environmental impact. Intermediate optimal solutions of mineral wool are particularly appealing, as they attain significant reductions in environmental impact at a marginal increase in cost.

Our approach aims to promote cost-effective building designs and in doing so guide decision-makers during the development of more effective environmental regulations in the building sector.

References

Al-Sanea, S. A., Zedan, M. F., & Al-Ajlan, S. A. (2005). Effect of electricity tariff on the optimum insulation-thickness in building walls as determined by a dynamic heat-transfer model. *Applied Energy*, *82*(4), 313–330. Retrieved May 2014. http://www.sciencedirect.com/science/article/pii/S0306261904001965.

Anon. (2009). European Parliament and Council of the European Union Decision No 406/2009/EC of the European Parliament and of the council of 23 April 2009 on the efforts of member states to reduce their greenhouse gas emissions to meet the community's emission reductio. *Official Journal of the European Union*, *L140*, 136–148.

Antipova, E., et al. (2014). Multi-objective optimization coupled with life cycle assessment for retrofitting buildings. *Energy and Buildings*, *82*, 92–99.

Asif, M., Muneer, T., & Kelley, R. (2007). Life cycle assessment: A case study of a dwelling home in Scotland. *Building and Environment, 42*(3), 1391–1394.

Audenaert, A., De Cleyn, S. H., & Buyle, M. (2012). LCA of low-energy flats using the Eco-indicator 99 method: Impact of insulation materials. *Energy and Buildings, 47*, 68–73.

BEDEC. (2011). BEDEC Database. Retrieved March 2015 from http://www.itec.es/nouBedec.e/bedec.aspx.

Blengini, G. A., & Di Carlo, T. (2010). The changing role of life cycle phases, subsystems and materials in the LCA of low energy buildings. *Energy and Buildings*, *42*(6), 869–880. Retrieved September 9, 2014 from http://www.sciencedirect.com/science/article/pii/S0378778810000022.

Brunet, R., Reyes-Labarta, J. A., et al. (2012). Combined simulation-optimization methodology for the design of environmental conscious absorption systems. *Computers & Chemical Engineering, 46*, 205–216.

Brunet, R., Cortés, D., et al. (2012). Minimization of the LCA impact of thermodynamic cycles using a combined simulation-optimization approach. *Applied Thermal Engineering*, *48*, 367–377. http://www.sciencedirect.com/science/article/pii/S135943111200275X.

Cabeza, L. F., et al. (2010). Experimental study on the performance of insulation materials in Mediterranean construction. *Energy and Buildings*, *42*(5), 630–636. Retrieved January 2, 2014 from http://www.sciencedirect.com/science/article/pii/S0378778809002825.

Cabeza, L. F., et al. (2014). Life cycle assessment (LCA) and life cycle energy analysis (LCEA) of buildings and the building sector: A review. *Renewable and Sustainable Energy Reviews*, *29*, 394–416. Retrieved January 2, 2014 from http://www.sciencedirect.com/science/article/pii/S1364032113005777.

Carreras, J., et al. (2014). Multi-objective optimization of thermal modelled cubicles considering the total cost and life cycle environmental impact. *Energy and Buildings*, *88*, 335–346. Retrieved December 29, 2014 from http://www.sciencedirect.com/science/article/pii/S0378778814010585.

Castell, A., et al., 2013. Life Cycle Assessment of alveolar brick construction system incorporating phase change materials (PCMs). *Applied Energy*, *101*, 600–608. Retrieved January 2, 2014 from http://www.sciencedirect.com/science/article/pii/S0306261912005119.

Copado-Méndez, P. J., Guillén-Gosálbez, G., & Jiménez, L. (2014). MILP-based decomposition algorithm for dimensionality reduction in multi-objective optimization: Application to environmental and systems biology problems. *Computers & Chemical Engineering*, *67*, 137–147. http://www.sciencedirect.com/science/article/pii/S0098135414001112.

Daouas, N. (2011). A study on optimum insulation thickness in walls and energy savings in Tunisian buildings based on analytical calculation of cooling and heating transmission loads. *Applied Energy*, *88*(1), 156–164. http://www.sciencedirect.com/science/article/pii/S0306261910002990.

Das, I., & Dennis, J. E. (1998). Normal-boundary intersection: A new method for generating Pareto optimal points in multicriteria optimization problems. *SIAM Journal on Optimization, 8* (3), 631–657.

DOE. (2013). Residential Prototype Building Models. U.S. Department of Energy. Retrieved May 2015 from http://www.energycodes.gov/development/commercial/90.1_models.

Ecocostsvalue. (2014). Ecocostsvalue. Retrieved May 2015 from www.ecocostsvalue.com.

Eco-Indicator 99. (2000). PRé Consultants. The Eco-indicator 99A damage oriented method for life cycleimpact assessment. Methodology report and manual for designers. *Technical report, PRé Consultants,* Amersfoort, The Netherlands.

Ecoinvent. (2015). The Ecoinvent Center. A competence centre of ETH; PSI; Empa & ART. Retrieved April 2015 from http://www.ecoinvent.ch/.

Ehrgott, M. (2005). *Multicriteria optimization*. http://www.ncbi.nlm.nih.gov/pubmed/24694125.

EnergyPlus. (2015). EnergyPlus, energy simulation software. Retrieved March 2015 from http://apps1.eere.energy.gov/buildings/energyplus/.

EURIMA. (2014). Retrieved April 2015 from http://www.eurima.org/.

EUROSTAT. (2013). Eurostat database. http://epp.eurostat.cec.eu.int.

Fokaides, P. A., & Papadopoulos, A. M. (2014). Cost-optimal insulation thickness in dry and mesothermal climates: Existing models and their improvement. *Energy and Buildings*, *Part A*, *68*, 203–212. Retrieved November 25, 2013 from http://www.sciencedirect.com/science/article/pii/S0378778813005689.

Gebreslassie, B. H., et al. (2009). Design of environmentally conscious absorption cooling systems via multi-objective optimization and life cycle assessment. *Applied Energy*, *86*(9), 1712–1722. Retrieved May 7, 2014 from http://www.sciencedirect.com/science/article/pii/S0306261908003048.

Grossmann, I. E., & Guillén-Gosálbez, G. (2010). Scope for the application of mathematical programming techniques in the synthesis and planning of sustainable processes. *Computers & Chemical Engineering, 34*, 1365–1376.

Guillén-Gosálbez, G., & Grossmann, I. E. (2009). Optimal design and planning of sustainable chemical supply chains under uncertainty. *AIChE Journal, 55*, 99–121.

IEA. (2008). Promoting energy efficiency investments. Case studies in the residential sector.

JEPlus+EA. (2015). JEPlus+EA, an EnergyPlus simulation manager for optimization studies. Retrieved May 2015 from http://www.jeplus.org/.

Kaynakli, O. (2012). A review of the economical and optimum thermal insulation thickness for building applications. *Renewable and Sustainable Energy Reviews*, *16*(1), 415–425. Retrieved May 13, 2015 from http://www.sciencedirect.com/science/article/pii/S1364032111004163.

Kottek, M., et al. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift, 15*, 259–263.

LIDER. (2009). Ministerio de Fomento, Government of Spain - LIDER, V. 1.0. Retrieved May 2015 from http://www.codigotecnico.org.

Menoufi, K., et al. (2012). Evaluation of the environmental impact of experimental cubicles using Life Cycle Assessment: A highlight on the manufacturing phase. *Applied Energy*, *92*, 534–544. Retrieved January 2, 2014 from http://www.sciencedirect.com/science/article/pii/S0306261911007173.

Menoufi, K., et al. (2013). Life cycle assessment of experimental cubicles including PCM manufactured from natural resources (esters): A theoretical study. *Renewable Energy*, *51*, 398–403. Retrieved January 2, 2014 from http://www.sciencedirect.com/science/article/pii/S0960148112006477.

Messac, A., Ismail-Yahaya, A., & Mattson, C. A. (2003). The normalized normal constraint method for generating the Pareto frontier. *Structural and Multidisciplinary Optimization, 25*, 86–98.

Ortiz, O., Castells, F., & Sonnemann, G. (2009). Sustainability in the construction industry: A review of recent developments based on LCA. *Construction and Building Materials, 23*(1), 28–39.

Ozel, M. (2013a). Determination of optimum insulation thickness based on cooling transmission load for building walls in a hot climate. *Energy Conversion and Management*, *66*, 106–114. http://www.sciencedirect.com/science/article/pii/S0196890412003846.

Ozel, M. (2014). Effect of insulation location on dynamic heat-transfer characteristics of building external walls and optimization of insulation thickness. *Energy and Buildings*. http://www.sciencedirect.com/science/article/pii/S037877881300697X.

Ozel, M. (2013b). Thermal, economical and environmental analysis of insulated building walls in a cold climate. *Energy Conversion and Management*, *76*, 674–684. Retrieved November 26, 2013 from http://www.sciencedirect.com/science/article/pii/S019689041300472X.

Pan, D., et al. (2012). The effects of external wall insulation thickness on annual cooling and heating energy uses under different climates. *Applied Energy*, *97*, 313–318. http://www.sciencedirect.com/science/article/pii/S0306261911008038.

Rivas-Martinez, S., Rivas Sáenz, S., & Penas Merino, A. (2002). *Worldwide bioclimatic classification system* (Vol. 1, pp. 1–638). Global Geobotany.

Sabio, N., et al. (2014). Multi-objective optimization under uncertainty of the economic and life cycle environmental performance of industrial processes. *AIChE Journal*, p.n/a–n/a. http://doi.wiley.com/10.1002/aic.14385.

Singh, A., et al. (2011). Review of life-cycle assessment applications in building construction. *Journal of Architectural Engineering, 17*(1), 15–23.

Stephan, A., Crawford, R. H., & de Myttenaere, K. (2013). A comprehensive assessment of the life cycle energy demand of passive houses. *Applied Energy*, *112*, 23–34. Retrieved September 7, 2014 from http://www.sciencedirect.com/science/article/pii/S0306261913004996.

Yu, J., et al. (2009). A study on optimum insulation thicknesses of external walls in hot summer and cold winter zone of China. *Applied Energy*, *86*(11), 2520–2529. http://www.sciencedirect.com/science/article/pii/S0306261909000841.

Zhu, P., Huckemann, V., & Fisch, M. N. (2011). The optimum thickness and energy saving potential of external wall insulation in different climate zones of China. *Procedia Engineering*, *21*, 608–616. http://www.sciencedirect.com/science/article/pii/S1877705811048909.