Mariano Anderle *Editor*

Innovations in Land, Water and Energy for Vietnam's Sustainable Development

Springer

UNIPA Springer Series

The **UNIPA Springer Series** publishes single and co-authored thematic collected volumes, monographs, handbooks and advanced textbooks, conference proceedings, professional books, SpringerBriefs, journals on specific issues of particular relevance in six core scientific areas. The issues may be interdisciplinary or within one specific area of interest. Manuscripts are invited for publication in the following fields of study:

1- Clinical Medicine;
2- Biomedical and Life Sciences;
3- Engineering and Physical Sciences;
4- Mathematics, Statistics and Computer Science;
5- Business, Economics and Law;
6- Human, Behavioral and Social Sciences.

Manuscripts submitted to the series are peer reviewed for scientific rigor followed by the usual Springer standards of editing, production, marketing and distribution. The series will allow authors to showcase their research within the context of a dynamic multidisciplinary platform. The series is open to academics from the University of Palermo but also from other universities around the world. Both scientific and teaching contributions are welcome in this series. The editorial products are addressed to researchers and students and will be published in the English language.

More information about this series at http://www.springer.com/series/13175

Mariano Anderle
Editor

Innovations in Land, Water and Energy for Vietnam's Sustainable Development

UNIVERSITÀ
DEGLI STUDI
DI PALERMO

Springer

Editor
Mariano Anderle
Embassy of Italy to Vietnam
Hanoi, Vietnam

Preface

The recent steep economic growth of South East Asia, and particularly of Vietnam, poses challenges to land, water and energy management as well as to urban environment and nature protection. In the last years, the Italy–Vietnam research cooperation supported by the governments of the two countries has explored these hot topics underlining the existing close linkages between the various sectors.

The aim of this manuscript is thus to show what the most innovative trends in land, water and energy management are, using also advanced information and communication technologies developed by Italian and Vietnamese researchers.

Vietnam is a fast-growing country with an average gross domestic product (GDP) increase each year of around 7%. At the same time, investments in higher education, research and industrial innovation have been increased.

Fundings in research and development are around 0.40% of the GDP. In respect to all other national public spending, the amount dedicated to science and technology is around 1.6% with a workforce involved in research of 700 full-time researchers per million inhabitants.

Alongside research, Vietnam also invests in the higher education sector. Indeed, estimates by the Ministry of Education and Training in August 2018 show a total of 0.5% of the GDP for that sector.

Italy enjoys Vietnamese's respect and recognition. This recognition, where properly oriented, allows the research system to be focused on areas advantageous for the development of economic and commercial partnership between the two countries. Particularly relevant in this sense are the applied researches in the agri-food, medicine and health, biotechnologies and genomics, nanotechnologies, energy and environment, physics and space.

Moreover, the mature competences of the Italian research system, accustomed to international visibility, have backed up Vietnamese researchers and allowed them to enter in scientific organizations promoting research dissemination.

Bilateral cooperation in science and technology between Italy and Vietnam is regulated by an Agreement on Scientific and Technological Cooperation, signed in Hanoi on January 5, 1992. Since then, executive protocols for scientific and technological cooperation have supported, on a regular deadline schedule, the

implementation of bilateral projects. The last one, valid for the period 2017–2019, was signed in Rome in November 2016 between the Ministry of Science and Technology of Vietnam and the Ministry of Foreign Affairs and International Cooperation of Italy. In the framework of this executive protocol, a call for bilateral projects is typically launched and a few projects are selected and financed. During this last triennium—with the seal of the "1st Italy–Vietnam conference on science and technology-bilateral research experiences and projects: results and perspectives" held in Hanoi on November 17, 2017, and organized by the Embassy of Italy in Vietnam with the collaboration of Vietnamese Ministry of Science and Technology—all of the funded projects have shown a very intense activity in terms of scientific results achieved, efficacy of the bilateral collaborations and potential impact. It is also evident how the bilateral scientific activities between the two countries have indeed acquired a great visibility, increasing the creation of international networks, and, ultimately, eased the access to funds.

In this scenario, there are more than 100 signed and operative Memorandum of Understanding between Italian and Vietnamese scientific and higher education institutions, while the dissemination activity on the research topics ranges from local events to international conferences.

This book, comprising the most important scientific projects and results of the bilateral cooperation in the last three years, has been structured in three main sections: Environment, Climate Change and Land Management in Vietnam; Energy for Vietnam; and Cities and utilities in Vietnam.

In the first part, Environment, Climate Change and Land Management in Vietnam, a few chapters concentrate on water systems in Vietnam, including rivers and seacoasts. An efficient flood forecasting model shows the weakness of the delicate ecosystem of the rivers in China, Laos and Vietnam. The model allows flood control downstream, to prevent catastrophes in the heavily populated Northern Delta of the Red River and close to the Ca river, the Vinh area, both experiencing rapid development and increasing urbanization.

The worldwide famous Ha Long Bay is also a weak marine ecosystem, severely jeopardized by intensified tourism and climate change effects. The bilateral research puts into evidence the unique richness of this area witnessed by sedimentological and microfossil content of sediment samples.

Other studies in the area of agri-food explore new ways to fight against rice pathogens. Italy is indeed the largest Japonica rice producer in Europe, and Vietnamese rice belongs to the same variety. Both countries are under biotic and abiotic stresses related to drought, climate changes and pests.

Finally, the land and environment management with a Global Navigation Satellite System (GNSS) created within a framework of bilateral scientific cooperation is presented. The chapter describes the GNSS technology that easily allows precise determination of position, velocity and time (PVT) enabling hundreds of different applications in almost any field of modern life, from transportation to logistics, from surveying to disaster management, from natural resources monitoring to services for citizens.

Second part, Energy for Vietnam, is focused on energy and water. The steep growth of the country is a great challenge for the Vietnamese government that has to face the reinforcement of the electrical infrastructures. An overview about the measures the Vietnamese government is planning to take in the medium-short term is provided together with their impact on the Vietnamese power system. In this scenario, high-voltage direct current (HVDC) transmission is considered to reinforce the connections with China and other neighbouring countries.

New architectures for renewable energy installation, such as microgrids, are also considered. These being analyzed both as urban energy districts and as islanded systems.

In the same part, the reader is given a brief overview of the conventional biological wastewater treatment (WWT) plants and processing steps, with a specific focus on their solid residues such as sewage sludge. The latter could be valorized by anaerobic digestion (AD), converting it into valuable commercial products such as biogas and fertilizers, applying hence the important concept at the base of circular economy.

Finally, third part, Cities and Utilities in Vietnam, is focused on cities. Urbanization is also a consequence of the steep economic growth of the country, and a study on the urban morphology of the Vietnamese metropolis, with special attention to Hanoi and Ho Chi Minh City, is provided. The same technology proposed for environment and land management is in this section considered for road applications, enabling consistent improvements in traffic management and monitoring. According to the proposal, in Vietnam, public transport vehicles and lories should be equipped with black boxes to enable traffic management and control. Other contributions focus on the telecommunication infrastructures. In countries like Vietnam, where wired communication systems are not sufficient for satisfying the continuously increase of user demand, the development of a highly efficient wireless communication network is mandatory. A good solution to this problem is represented by the 5th-generation mobile network or simply 5G in which the design of antennas is a strategic element.

Hanoi, Vietnam   Mariano Anderle

Contents

Environment, Climate Change and Land Management in Vietnam

A Hydrometeorological Flood Forecasting Chain for the Red and Ca rivers (China, Laos and Vietnam) Part I—Investigated Areas and Model Setup

Roberto Ranzi, Lê An Ngô, Thanh Tùng Hoàng, Hoàng Sơn Nguyễn, Stefano Barontini, Giovanna Grossi, Baldassare Bacchi, Andrea Buzzi, Silvio Davolio, Oxana Drofa, Piero Malguzzi, Lê Thuy Đỗ, Van Hoa Võ, and Minh Cát Vũ

Abstract As a result of two research projects conducted over the years from 2005 to 2011, this paper describes the configuration of a coupled hydrometeorological flood forecasting system set up for the Red river basin (169,000 km^2 at the outlet) and the Ca river basin (23,830 km^2 at the outlet), across China and Vietnam, and Laos and Vietnam, respectively. The research objectives were to setup a hydrometeorological chain to forecast the discharge flowing into reservoirs and at some control sections for flood control in downstream areas of the Red river and Ca river. These areas include the heavily populated Northern Delta of Red river and, for the Ca river, the Vinh area. Both these areas are in fact experiencing a rapid population and economic development. The distributed hydrological model DIMOSHONG, which implements a modified Green and Ampt or an SCS—type routing scheme of the soil–water content, is forced by Quantitative Precipitation Forecasts provided by the BOLAM and the MOLOCH meteorological models or by raingauge observations. The forecasting chain has an Extended Kalman Filtering updating module based on hydrometric measurements now available in real-time for two pilot stations installed in Ghenh Ga section and Nam Dan section, in the Red and Ca rivers, respectively. Several floods in the basins, including the August 1971 and October 2010 catastrophic ones which flooded the Red river and the Ca river regions, were simulated providing useful results in view of improving flood alert systems.

R. Ranzi (✉) · S. Barontini · G. Grossi · B. Bacchi
Università degli Studi di Brescia, Brescia, Italy
e-mail: roberto.ranzi@unibs.it

L. A. Ngô · T. T. Hoàng · H. S. Nguyễn · M. C. Vũ
Thuyloi University, Hanoi, Vietnam

A. Buzzi · S. Davolio · O. Drofa · P. Malguzzi
CNR-ISAC, Bologna, Italy

L. T. Đỗ · V. H. Võ
National Center for Hydrometeorological Forecasting, Hanoi, Vietnam

1   Introduction

Flood hazard mitigation measures have been traditionally mainly focused on structural defense systems, including river embankment, diversion, by-pass and reservoirs construction. Starting with the XIX century, non-structural measures based on flood alert systems, land planning, flood risk mapping and communication have been implemented in some areas and today are becoming an important alternative for enhancing awareness of people about water-related hazards and for reducing flood risk. The effectiveness of flood alert systems is improving in the last decades as meteorological and hydrological forecast models become more accurate, and raingauge and hydrometric observing systems, which provide real-time data, are installed in several basins worldwide. This is the case also of Vietnam, a country which is prone to water-related hazards triggered by monsoon and tropical typhoons rainfall.

In this paper, we give an overview of the set up of a hydro-meteorological system aiming at extending the lead time and improving the effectiveness of forecasting for a better control of reservoirs in the Red river and in the Ca river, draining from China and Laos to Vietnam (Fig. 1). The system is based on three main components:

1. A quantitative precipitation forecast system provided by the BOLAM—MOLOCH Mesoscale Meteorological Model chain, nested into the ECMWF (for Red river) and NOAA-GFS (Ca river) Global Circulation Model initial and boundary conditions;
2. A hydrological flood forecast with the DIMOSHONG (DIstributed hydrological MOdel for the Shong HONG) hydrological model;
3. A real time observation system of rainfall and river level, providing runoff data for the hydrological model updating with an Extended Kalman Filtering (EKF).

The investigated area and the first two components of the model setup are described in the following.

2   The Red and the Ca Rivers in Vietnam

2.1   *The Red River Basin*

The Red river (Sông H`ông in Vietnamese) is an international basin located in China, Laos and Vietnam. With the area of about 169,000 km^2, it is the largest basin in Vietnam. Its source originates in China, flows to the southeast entering into Vietnam at Lao Cai, where the drained area is 41,000 km^2, and discharges into the Tonkin Gulf in the East Sea. Three main tributaries in the Red river system in Vietnam's territory
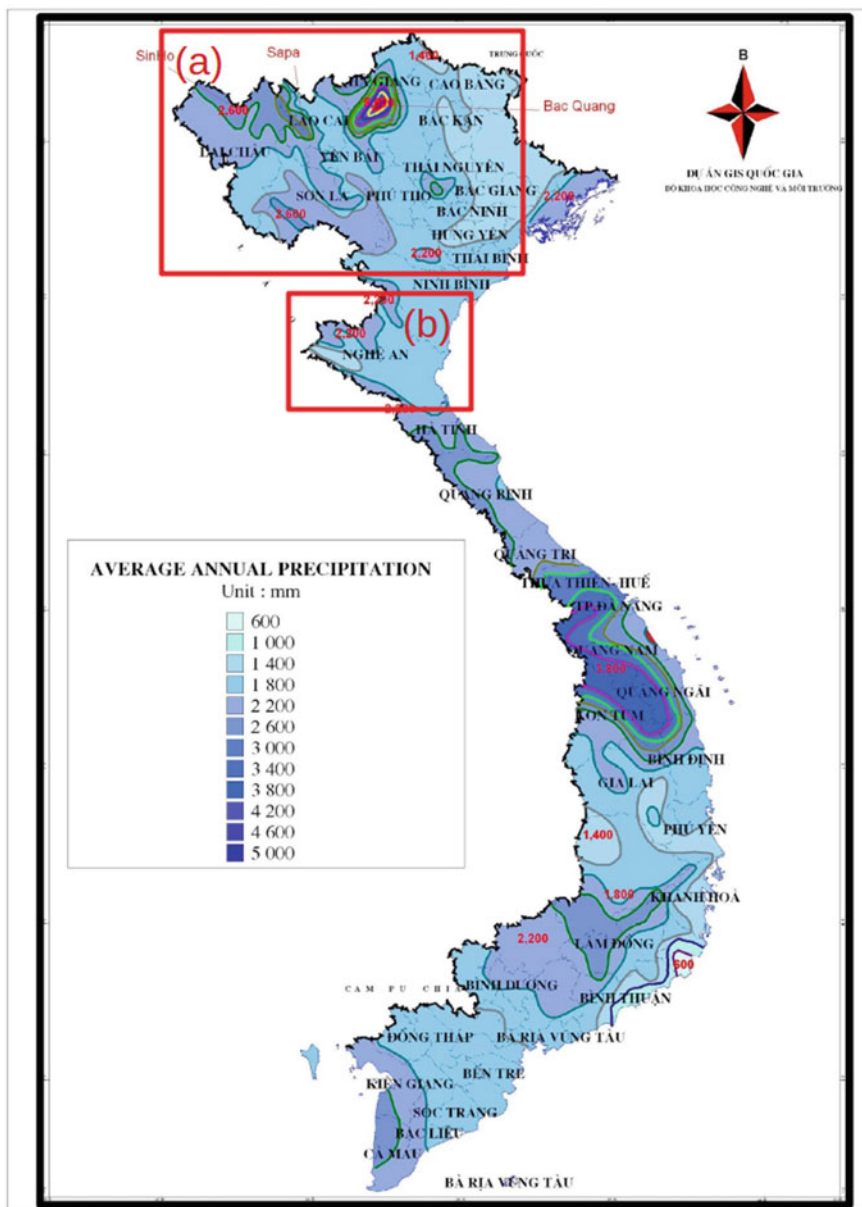
Fig. 1 Location of the Vietnamese part of the Red river (**a**) and of the Ca river (**b**) basin superimposed on the annual precipitation map of Vietnam

are called Lo (where rainfall may reach 4000 mm/year), Thao, and Da rivers (Fig. 2).
These three main tributaries join to form the Red river, upstream Hanoi, in Viet Tri.
The Red river Delta is a very important area for the society, education, culture and
also economics of Vietnam. As a consequence, both the Red river basin and its Delta
are thoroughly investigated to assess their proneness to water related issues (e.g. Luu
et al. 2010; Simons et al. 2016).

The climate of the basin ranges from tropical to subtropical. It is dominated by
the northeast monsoon in winter and the southwest monsoon in summer. The basin is
located next to the sub-humid tropical region of South East Asia. Since the topography
is divided into individual mountain ranges and climate regime is affected by many
synoptic weather systems, rainfall in the basin varies strongly in space and time.

A forecasting system in the Red river basin has been developed and implemented
by the Vietnamese National Flood Control Bureau. The warning time is the time lag
of the basin in the Vietnamese territory. This time lag can be extended by means
of Quantitative Precipitation Forecasts, issued some days in advance and obtained
from the meteorological model chain mentioned above and described in the second
section.
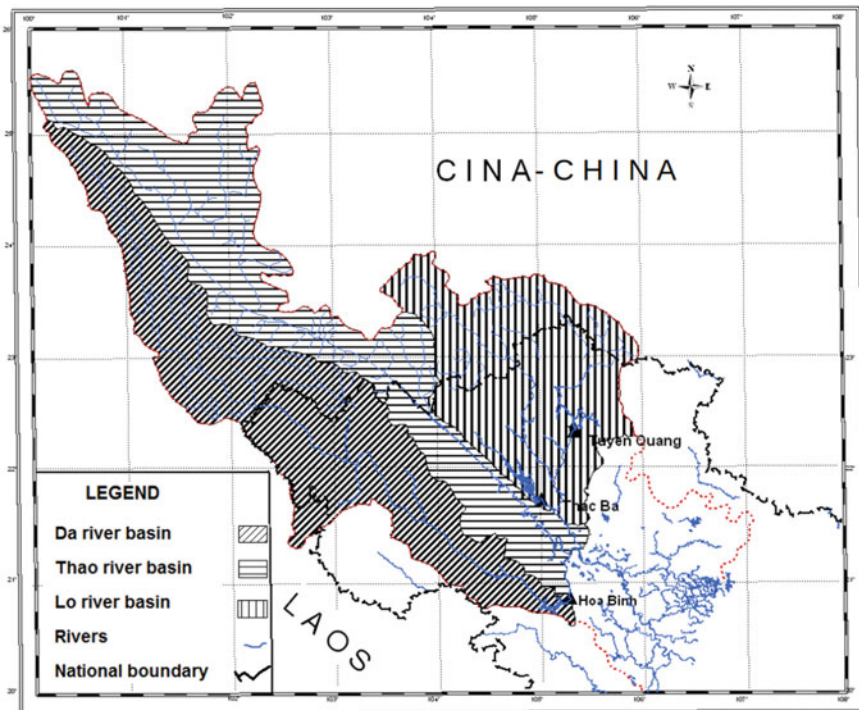
Fig. 2 The Red river basin at the Viet Tri junction, and its three major tributaries, together with
the Tuyen Quang, Thac Ba and Hoa Binh reservoirs

Rainfall Data

Annual precipitation amount in the Vietnam's part of the Red river ranges between 1600 and 2800 mm with mean annual 24 h precipitation ranging from 140 mm in Sa Pa and Sin Ho to 300 mm in Bac Quang (see Fig. 1). Raingauge data are both from automatic and mechanical stations, with hourly records, and from climatological stations, with 6 or 12 h manual recording time. Readings are conducted at 1, 7, 13, and 19 local time, corresponding to UTC + 7 h. In the Da river up to 23 raingauge data were available, 9 in the Thao river and 15 in the Lo river. In China data for only 7 raingauge stations were available for this study and this poses limitations on the accuracy of both the hydrological model simulations and the verification of the simulated rainfall fields.

Runoff Data

Runoff data with 6 h time steps were collected at 20 runoff stations for 5 major flood events occurred from 13 to 20 August 1971, from 13 to 20 August 1992, from 18 to 25 July 2000, from 10 to 19 August 2002 and from 19 to 25 July 2007. The most significant data are those of Vu Quang (draining 37,000 km^2) in the Lo river, Yen Bai (48,000 km^2) in the Thao river, Ta Bu (45,900 km^2) in the Da river and Son Tay (144,000 km^2) in the Red river downstream the junction of the three main tributaries, where the mean annual flow is 3500 m^3/s and the mean annual peak flow over the 1960–2004 period is 16,450 m^3/s.

Reservoirs

In the Red river system, there are 3 main reservoirs which were operative at the time when the research was conducted, others being built or planned in the Da and Lo rivers. The Hoa Binh reservoir (with 9.6 billion m^3 storage, in operation since 1989) is in the Da river, the Thac Ba (with 2.49 billion m^3 storage, in operation since 1971) and Tuyen Quang reservoirs (with 2.245 billion m^3 storage, in operation since 2005) in the Lo river. The missions of these reservoirs are: to protect the Ha Noi capital and the Red river Delta against all floods with 250 years return period at Son Tay (upstream of Hanoi), to keep the water level at Hanoi less than 13.4 m (above the hydrometric reference level), to generate hydropower and to provide water for irrigation (Ngo et al. 2007; Castelletti et al. 2012). They are modifying the runoff and sediment transport regime (Ranzi et al. 2012) downstream, but have already been able to mitigate the intensity of major floods, after the catastrophic one of August 1971 which reached 34,300 m^3/s in Son Tay. Because of their multiple purposes and their large storage capacity their effective operation would benefit from a warning of inflow some days in advance.

2.2 The Ca River Basin

Floods threaten frequently also the Ca river. Focusing in the recent past, in 2007 a series of five floods occurred in only one-month period causing huge damage on both people and properties in the central districts of Nghe An province in the Ca river basin; also in October 2010 about 50 casualties and widespread inundation occurred in the same area. Ca river basin is in the northern part of Central Vietnam, with latitude ranging from 18°15′ to 20°10′30″ North and longitude ranging from 103°45′20″ to 105°15′20″ East (Fig. 3). The main river originates from Muong Khut and Muong Lap in Laos at elevation of 1800–2000 m, flows down from the northwestern to southeast direction, enters into Nghe An province at Keng Du, flows still to the Eastern Sea at the Hoi estuary. This branch is properly named Ca river. The branch downstream of Ca river from Cho Trang to Hoi estuary, about 35 km long, should be named as Lam river rather than Ca river, and it is formed by Ca and La (originated from Ha Tinh) rivers. The natural area of the Ca river basin is 27,200 km^2 in total up to Hoi river mouth, of which 17,730 km^2 are in Vietnam and 9470 km^2 in Laos. In some documents additional 2650 km^2 are added to the Vietnamese part of the basin to account for wetlands, lagoons and little rivers in its downstream part which flow directly to the sea or through secondary branches of Ca river. With this addition the Vietnamese part of the basin is 20,460 km^2 and the total basin area is 29,850 km^2. The Ca river bed is very narrow, with steep slope in the upstream area, widens up in
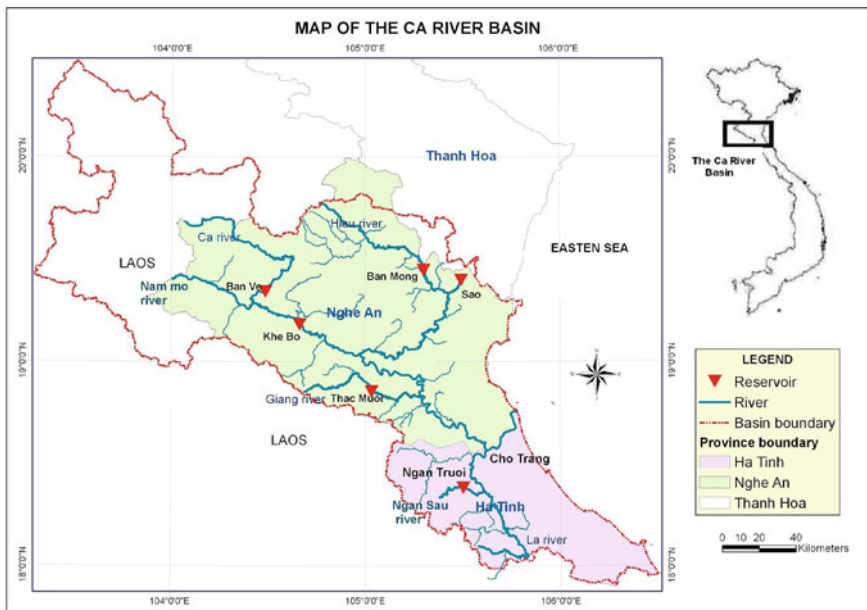
Fig. 3 Map of the Ca river, in Vietnam, with the main river branches, reservoirs and administrative boundaries

the middle basin (from Con Cuong to Anh Son), then joins the Hieu river on its left side. In the downstream area, the Ca river flows through the plain and finally joins the La river on the right side. The length of the main Ca river is 531 km in total, of which 361 km is in Vietnam and its mean elevation is 294 m a.s.l.

Hieu river is the largest tributary on the left side of the Ca river. The catchment area of the Hieu river is 5340 km^2, its length is 228 km, originates from Pu Hoat Range in the Laos-Vietnam boundary, flows into the Ca river at Anh Son. La river is the main tributary on the right side of the Ca river: it originates from the Giai mountainous area and enters into Ca river at Cho Trang. The catchment area of the La river is 3210 km^2.

Rainfall and Landcover Data

Mean annual precipitation in the basin ranges from 1100 mm in the upstream mountain area to 2400 mm close to the junction with the La river. Rainfall is concentrated in one rainy season from May to October when falls 80% to 85% of annual rainfall. Heaviest rainfall occurs in September when mean of maximum daily rainfall is up to 250 mm in the coastal area, and in 2010 a maximum value up to 800 mm was recorded in 24 h. Land is covered mainly by forests (66%) and by agricultural crops (16%). In flat land, Gleysols, characteristics of lagoons, wetlands or areas with shallow groundwater tables, are the least prone ones to infiltration due to fine soil texture and being frequently waterlogged. In the mountain range, upstream Acrisols, very common in tropical and subtropical areas, were considered less prone to infiltration than downstream ones, due to higher slopes. In the Laotian part of the basin also a small area of Leptosol is present, with small infiltration capacity due to their small thickness.

Reservoirs

In the Ca river basin, there are major reservoirs being operated or under construction such as Sao reservoir on the Sao river, Ban Mong reservoir on the Hieu river, Ban Ve reservoir and Khe Bo reservoir on the Ca river, and Thac Muoi reservoir on the Giang river. The operation of these reservoirs, which have a total storage capacity of 2.8 billion m^3, an installed power generation capacity of 485 MW (excluding Song Sao) has to meet flood control with 411 million m^3 storage available, hydropower generation, water supply for irrigating 33,487 ha, industrial, domestic and environmental demand (Vietnam Institute of Water Planning 2004). Operation rules are set for single reservoirs, mainly for dam safety reasons, and no regulation practices are set for the overall reservoir system. For mitigation of damages caused by floods, as well as to take into account socio-economic and environmental effects of reservoirs operation, it is necessary to extend research on both medium range rainfall and runoff forecast and to set a comprehensive scientific basis for operation of the reservoir system (Hoang et al. 2012).

3 The Meteorological BOLAM—MOLOCH Model

3.1 The BOLAM Hydrostatic Mesoscale Meteorological Model

BOLAM is a Mesoscale Meteorological Model (MMM) based on hydrostatic primitive equations and terrain—following coordinates. It was developed starting in the early 1990's at ISAC in Bologna. A detailed description of the dynamics and numerical schemes of the model can be found in Buzzi et al. (1994) and in subsequent papers, to which we refer for the sake of brevity. The model was continuously developed and updated over the last years and applied in Europe (Malguzzi et al. 2006), the Mediterranean, Africa, the Pacific Ocean and, within this project, in South-East Asia, in an area comprising China, Vietnam and Laos. A more recent version of the model (GLOBO) was developed to run as Meteorological Global Model for medium range meteorological forecasts (Malguzzi et al. 2011).

The initial and lateral boundary conditions from a Global Circulation Model (GCM in Fig. 4) for the Red river study have been obtained from the ECMWF archives (see e.g. Persson, 2001), by retrieving the 6—hourly analyses available at $0.5° \times 0.5°$ resolution. Snow cover, sea surface temperature, and soil temperature and wetness at four soil layers have been derived also from the ECMWF analyses. For the Ca river, the NOAA—GFS initial and boundary conditions were adopted. In order to
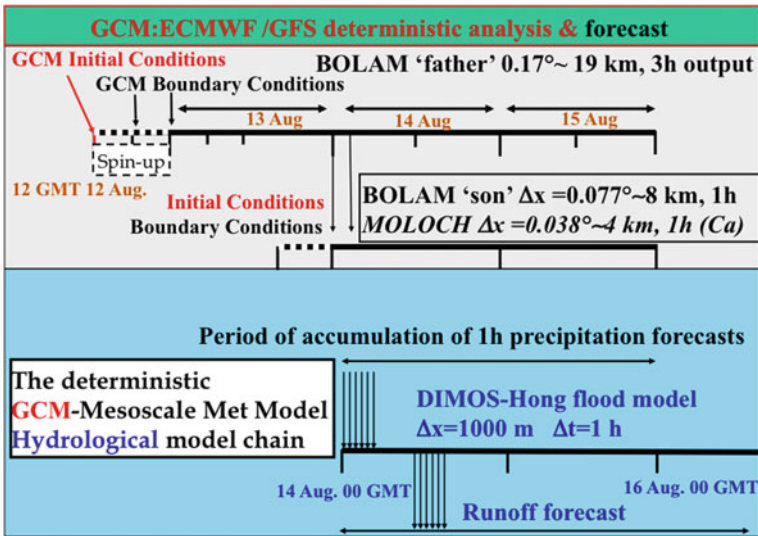
Fig. 4 Scheme of the GCM (Global Circulation Model)—BOLAM/MOLOCH Mesoscale Meteorological Model—DIMOSHONG Hydrological model forecasting chain setup to forecast floods in the Vietnamese part of the Red and Ca rivers. Precipitation forcing to the hydrological model can be provided also by raingauge observations

provide the high resolution simulations with more suitable boundary data than those available from global models, a dynamical downscaling chain has been organized: global analysis data, after undergoing a pre—processing procedure, have been used as initial and boundary conditions for simulations at a relatively coarse spatial resolution (about 19 km), intermediate between the global analyses resolution and the inner grid resolution. The coarse model integrations are conventionally named "BOLAM—father" runs. Output data obtained from the BOLAM—father simulations have been used as initial and boundary conditions for simulations at high spatial resolution. Such model simulations are conventionally indicated as the "BOLAM-son" runs (see the MMM in Fig. 4).

In the BOLAM—father configuration, the following parameters for the numerical solution have been used: rotated geographical system with the coordinate origin set at 21.3°N and 102.5°E; horizontal grid distance of 0.17° (about 19 km in both latitudinal and longitudinal directions); 38 hybrid levels in the vertical; basic time step of 150 s; output data storage every 3 h; duration of each continuous run of 72 h. In the BOLAM—son configuration, the following parameters have been used: same rotated geographical system as for BOLAM-father; horizontal grid distance of 0.077° (about 8 km); 40 hybrid levels; time step of 72 s; output data storage every 1 h; duration of each continuous run of 48 h.

3.2 The MOLOCH Non-hydrostatic Mesoscale Meteorological Model

For the Ca river meteorological forecasting the high resolution meteorological forecasting was implemented using both the BOLAM—son configuration and the MOLOCH non—hydrostatic meteorological model at 4 km resolution (see Fig. 4), also developed at CNR—ISAC. MOLOCH is a nonhydrostatic, convection—resolving model: this means that atmospheric convection equations are explicitly resolved by the numerical scheme instead of being parameterised, on the basis of already solved meteorological variables, as in BOLAM. It integrates the fully compressible set of equations with prognostic variables (pressure, temperature, specific humidity, horizontal and vertical velocity components, and five water species) represented on the lat–long rotated, Arakawa C-grid (Malguzzi et al. 2006). Terrain-following coordinates, relaxing smoothly to horizontal surfaces away from the earth surface, are employed. Model dynamics are integrated in time with an implicit scheme for the vertical propagation of sound waves, while explicit, time-split schemes are used for the remaining terms. Three-dimensional advection is computed using the eulerian WAF scheme. Horizontal second-order diffusion and divergence damping is included to prevent energy accumulation on the shorter space scales. The physical scheme consists of radiation (as for BOLAM), subgrid turbulence, microphysics, and soil water and energy balance. The turbulence scheme is based on an E-l closure, where the Turbulent Kinetic Energy equation (including advection) is

evaluated. Surface turbulent fluxes of momentum, specific humidity, and temperature are computed by the classical Monin—Obukhov theory with Businger/Holtslag functions in the unstable/stable case. The microphysical scheme is based on the parameterization proposed by Drofa and Malguzzi (2004) to take into account fast and slow processes of transformation of specific humidity and cloud quantities. The soil model is a four-layer model including vegetation.

4   The Hydrological DIMOSHONG Model

The output of meteorological simulations or raingauge-observed precipitation is then used as meteorological forcing to the distributed hydrological model implemented in this study called DIMOSHONG, standing for DIstributed hydrological MOdel for the Shong HONG. The DIMOSHONG model is developed, for the Red river basin, from the DIMOSOP model (Ranzi et al. 2003). The topography was represented by the Digital Elevation Model resampled at 1 km resolution from the Global DEM GTOPO30, provided by the USGS—EROS, in a raster format, lat—long coordinate system. The river network is then determined through the Hortonian order. The other physical parameters are estimated from the landuse, geological maps took from USGS, and Vietnam Atlas.
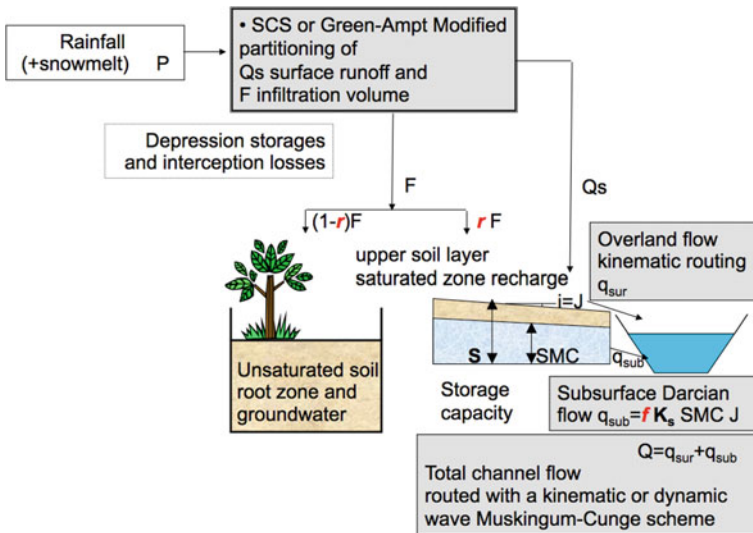
Fig. 5 Scheme of the DIMOSHONG model. In evidence the four main components of the model: **a** the rainfall partitioning scheme of rainfall, **b** the overland flow, **c** the subsurface Darcian interflow and **d** the channel runoff routing

The model is basically built by four components (Ngo et al. 2010, see Fig. 5). The first is the soil model that aims at describing in a relatively simple mathematical scheme the complex physical processes that include the interception of rainfall by vegetation, the storage in surface ponds and the partitioning of the resulting rainfall volume into infiltration and runoff. This component of the model can be either of the 'infiltration excess' or of the 'saturation excess' type. For the 'infiltration excess' scheme a Green and Ampt Modified scheme was adopted, assuming that the wetting front propagates downward with a rate corresponding to an equivalent saturated conductivity decreasing with depth. Lateral subsurface interflow is assumed, in a Darcian framework, proportional to the lateral soil transmissivity and local topographic gradient (Ngo 2008). The saturation excess model is developed according to the Soil Conservation Service—Curve Number method, but the soil columns are interconnected allowing for subsurface lateral flow propagation with a Darcian scheme, which is the second component of the model.

Infiltration excess or saturation excess surface runoff is then routed as overland flow with a linear reservoir model with time constant estimated from a kinematic wave equation.

The sum of overland flow and subsurface lateral interflow contribute to channel flow, which is propagated along the river network according to a Muskingum—Cunge scheme with celerity and diffusivity which can be estimated with a constant event-dependent value or changing dynamically during the flood development.

5   Conclusions

In this first paper the meteorological and hydrological model's component of a flood forecasting systems for two basins in Vietnam was presented. The BOLAM—MOLOCH—DIMOSHONG model chain was designed with a spatial resolution of the meteorological model of 8 km and 4 km, respectively, for the Red and Ca rivers, considering their respective size. The hourly time resolution and 1 km spatial resolution of the hydrological model output is suitable for most applications, including warnings for reservoir operations and alert systems for the inhabited areas along the rivers.

References

Buzzi A, Fantini M, Malguzzi P, Nerozzi F (1994) Validation of a limited area model in cases of Mediterranean cyclogenesis: surface fields and precipitation scores. Meteorol Atmos Phys 53:137–153

Castelletti A, Pianosi F, Quach X, Soncini-Sessa R (2012) Assessing water reservoirs management and development in Northern Vietnam. Hydrol Earth Syst Sci 16:189–199

Drofa OV, Malguzzi P (2004) Parameterisation of microphysical processes in a non-hydrostatic prediction model. In: Proceedings of the 14th international conference on clouds and precipitation, World Meteorological Organization, Bologna, Italy, 19–23 July 2004

Hoang TT, Ranzi R, Barontini S, Cat VM (2012) Medium range rainfall and flood forecasting for reservoir system operation in the Ca river basin (Vietnam). IAHR Asian-Pacific Division Congress, Korea, p 13

Luu TNM, Garnier J, Billen G, Orange D, Némery J, Le TPQ, Tran HT, Le LA (2010) Hydrological regime and water budget of the Red River Delta (Northern Vietnam). J Asia Earth Stud 37:219–228

Malguzzi P, Buzzi A. Drofa O (2011) The meteorological global model GLOBO at the ISAC—CNR of Italy assessment of 1.5 yr of experimental use for medium-range weather forecasts. Weathern Forecasting 26:1045–1055

Malguzzi P, Grossi G, Buzzi A, Ranzi R, Buizza R (2006) The 1966 'century' flood in Italy: a meteorological and hydrological revisitation. J Geophys Res 111:D24106. https://doi.org/10.1029/2006JD007111

Ngo LL, Madsen H, Rosbjerg D (2007) Simulation and optimization modelling approach for operation of HoaBinh reservoir in Vietnam. J Hydrol, 269–281

Ngo LA (2008) Flood forecast in the Red river basin (China-Vietnam) using a hydro-meteorological modeling system. Ph.D Thesis, Politecnico di Milano, 170pp

Ngo LA, Vu MC, Hoang TT, Buzzi A, Drofa O, Do LT, Barontini S, Ranzi R (2010) A hydrometeorological flood forecasting system for the reservoir control in the Red River. In: Proceedings of the ICOLD-international commission on large dams international symposium "Dams & Sustainable Water Resources Development" (on CD), Hanoi 23–25 May 2010, 11pp

Persson A. User guide to ECMWF forecast products, ECMWF Meteorological Bulletin M3.2, 2001, 105 pp.

Simons G, Bastiaanssen W, Ngô LA, Hain CH, Anderson M, Senay G (2016) Integrating global satellite-derived data products as a pre-analysis for hydrological modelling studies: a case study for the Red River Basin. Remote Sens 8:269

Ranzi R, Bacchi B, Grossi G (2003) Runoff measurements and hydrological modelling for the estimation of rainfall volumes in an Alpine basin. Quart J R Meteorol Soc 129 Part B, n 588:653–672

Ranzi R, Le TH, Rulli MC (2012) A RUSLE approach to model suspended sediment load in the Lo river (Vietnam): effects of reservoirs and land use changes. J Hydrol 422–423: 17–29. https://doi.org/10.1016/j.jhydrol.2011.12.009

Vietnam Institute of Water Planning (2004) The Ca River basin master plan. Ministry of Agriculture and Rural Development

A Hydrometeorological Flood Forecasting Chain for the Red and Ca Rivers (China, Laos and Vietnam) Part II—Applications and Results

Lê An Ngô, Thanh Tùng Hoàng, Hoàng Sơn Nguyễn, Minh Cát Vũ, Lê Thuy Đỗ, Van Hoa Võ, Silvio Davolio, Oxana Drofa, Stefano Barontini, and Roberto Ranzi

Abstract Applications and results of the hydrometeorological flood forecasting chain which has been setup for the Red river and the Ca river are briefly presented here. Five major floods occurred in the Red river in the 1971–2007 period and several floods occurred in the Ca river from 2006 to 2011, including the October 2010 catastrophic flood, were simulated with the BOLAM—MOLOCH—DIMOSHONG forecasting chain. The use of runoff observations at fourteen hydrometric stations in the Lo river, a major branch of the Red river, enabled the distributed model updating with an Extended Kalman Filtering, showing an improvement of flood forecasts compared to the results obtained with the rainfall-runoff model without hydrometric observations. Perspectives for improving the forecasting chain and for an operational implementation including surface observations and meteorological corrections using weather-types classification are discussed.

Keywords Red river · Ca river · Transnational river basins · Hydrometeorological flood forecast · Floodplain protection · Extended Kalman Filter

L. A. Ngô · T. T. Hoàng · H. S. Nguyễn · M. C. Vũ
Thuyloi University, Hanoi, Vietnam

L. T. Đỗ · V. H. Võ
National Center for Hydrometeorological Forecasting, Hanoi, Vietnam

S. Davolio · O. Drofa
CNR-ISAC, Bologna, Italy

S. Barontini (✉) · R. Ranzi
Università degli Studi di Brescia, Brescia, Italy
e-mail: stefano.barontini@unibs.it

1 Introduction

In this paper, which follows the description of the investigated areas and the setup of the hydrometeorological flood forecasting chain (Ranzi et al. 2020), a quantitative evaluation of results obtained in flood forecasting on the Red river, flowing from China to North Vietnam, using the BOLAM meteorological and the DIMOSHONG hydrological model are presented in the second section. In the third section results on the Ca river, flowing from Laos to the provinces of Nghe An and Ha Tinh in Vietnam, for which the non-hydrostatic MOLOCH model is nested into BOLAM, are presented together with hydrological simulations of the October 2010 flood. This episode was characterized by 800 mm of rainfall accumulated in a single day in Chu Le, which caused about 50 casualties, and about 100,000 houses inundated.

2 Meteorological and Hydrological Forecasts in the Red River

2.1 *Verification of Precipitation Forecasts*

Precipitation forecasts for five major floods occurred in August 1971, August 1996, July 2000, August 2002 and July 2007 in the Red river were simulated with a BOLAM—son model, a 8 km resolution model grid nested into a BOLAM—father, with 19 km resolution gridchain. Initial and boundary conditions were provided by the ECMWF model, as described in Ranzi et al. (2020). Precipitation forecasts were compared at the event scale (Fig. 1) and at the daily scale (Fig. 2) with raingauge observations, interpolated on a 1 km grid with a nearest—neighbour algorithm. The comparison was conducted by spatial averaging rainfall on the Vietnamese part only, where the available raingauge data density was close to $1/(1000 \text{ km}^2)$, of Lo river at Vu Quang, Thao river at Phu Tho and Da river at Hoa Binh basins, and the Red river at Son Tay.

As shown in Fig. 1, the coefficient of determination of simulated versus observed rainfall at the event scale, in each of the investigated areas, ranges from 0.06 only, in the Lo river, to 0.87 in the Da river. However, the BOLAM model at high resolution tends to overestimate observed precipitation by 42%, on average. Figure 2 shows that the time pattern of modeled daily precipitation reproduces quite satisfactorily the observations, especially for the 1971 flood, both at low and high model resolution, and for the 1996 and 2007 events. The overestimation of observed precipitation by the model in 2002, instead, is quite significant. Although the BOLAM model was already verified to be capable of reproducing tropical cyclones quite satisfactorily, in comparison with other mesoscale models (Nagata et al. 2001), reasons of the bias of the BOLAM model are possibly due to the earlier scheme adopted for microphysics and convection. In fact, the new version of the model applied to the Ca river, with an upgraded microphysics and convection scheme, better performed.
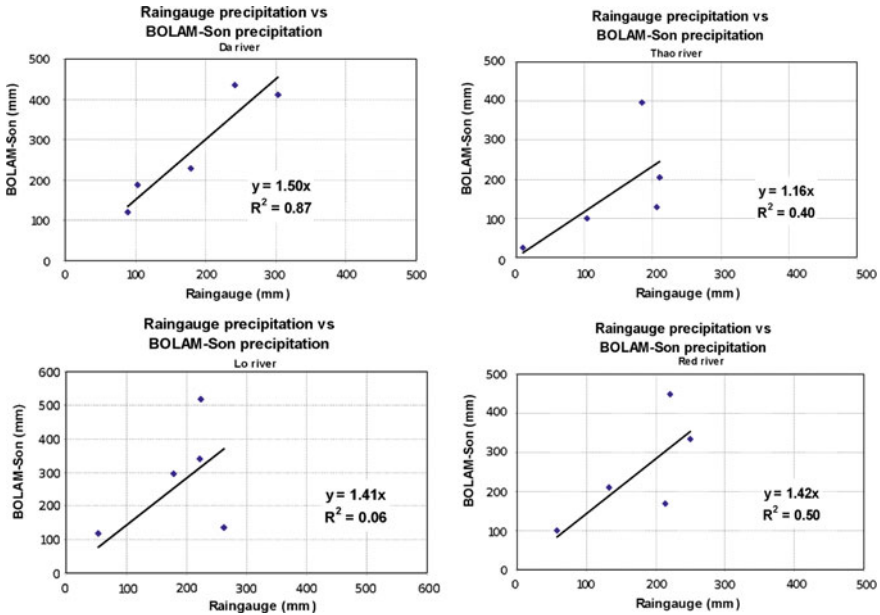
Fig. 1 Comparison of the observed versus simulated total precipitation averaged over the three major sub-basins and the Red river for the five flood events selected (1971, 1996, 2000, 2002 and 2007). The results of BOLAM with high resolution are adopted

2.2 Runoff Observations and Forecasts with an Extended Kalman Filtering

Hydrological simulations with any model are affected by several sources of errors, mainly due to inadequate rainfall forcing, processes conceptualization and parameters adopted. For this reason, runoff observations are fundamental to correct the model's state or parameter variables on the basis of simulated and observed output. For the model's adaptation to be effective runoff data must be easily accessible with a time delay much shorter that the response time of the basin. If this happens we can say that the data are available in "real-time".

In the perspective of the hydrological model's adaptation to runoff observations, two surveys were conducted by staff members of the Thuyloi University and the University of Brescia in December 2007 in the Lo basin and in October 2009 in the Ca basin to identify existing hydrometric stations suitable for the installation of automatic monitoring instruments with data transmission systems. The installation of two ultrasonic hydrometers manufactured by CAE was completed in 2008 and 2012, in Ghenh Ga and Nam Dan, respectively in the Lo and Da basins (see Fig. 3). Because of the availability of hydrometric data in real-time in Ghenh Ga since 2008, an adaptive hydrological forecasting system was firstly setup for the Lo basin. The theory for adapting hydrological models to observations has been developed starting
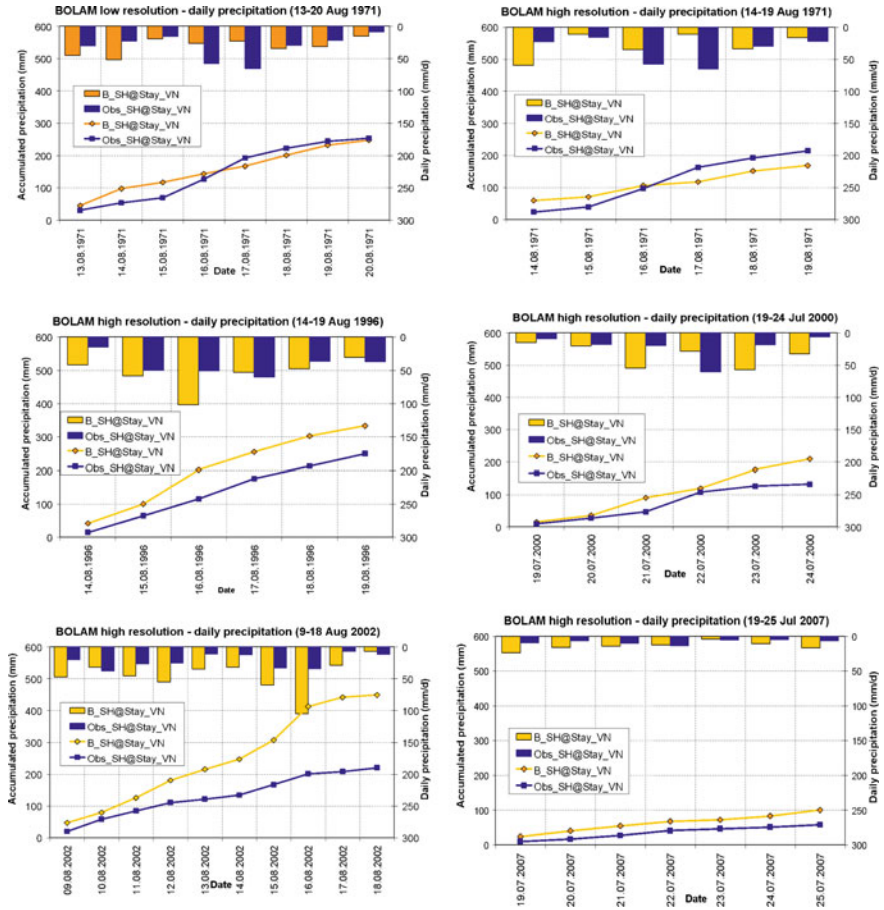
Fig. 2 Comparison of the observed (Obs_SH@Stay_VN) versus BOLAM (B_SH@Stay_VN) simulated total daily and accumulated precipitation averaged over the Red river (SôngH`ông) at Son Tay for the five flood events selected (1971, 1996, 2000, 2002 and 2007). The BOLAM—high resolution results are shown, with the exception of the top left panel where the BOLAM—low resolution results are shown

in the 70s of the last century (Todini et al. 1976; O'Connell and Clarke 1981). It is based, in principle, on either modeling forecast errors in a stochastic and statistical approach or in updating model's state and parameter values. Montanari and Grossi (2008), among others, provided recent advances on the error modelling framework. Moradkhani et al. (2005) discuss and provide novel techniques for the second class of model updating, based on Ensemble Kalman Filtering (EnKF). Early theory of model's updating was mainly based on the assumption of linearity of hydrological systems and often use "classical" Kalman filtering method. Because of the non-linearity of most hydrological processes, including infiltration, runoff formation and

Fig. 3 Surveys conducted by the University of Brescia (Roberto Ranzi) and Thuyloi University staff (Vu Minh Cat, Hoang Thanh Tung, Le Van Chin, Nguyen Hoang Sơn and Tran Kim Chau) for the installation of a tipping-bucket raingauge and ultrasonic hydrometer installed by the CAE company (Bologna, Italy) in a well (left) in the Ghenh Ga station on the Lo river. On the right image the ultrasonic hydrometer ready for the installation on the Nam Dan bridge, in the Ca river basin. On the right Mr. Nguyen Xuan Tien the Vice-Director of the Central Northern Provinces Flood Forecasting Centre, responsible for flood forecasting on the Ca river

propagation, the Extended Kalman Filtering (EKF) method should be, in principle, more suitable for hydrological models.

Extended Kalman Filter relies on linearization using first order approximation of Taylor series of the state-space model (1), which can be non-linear, and can be used when the non-linearity of the system is weak. In the state-space non-linear model it is written:

$$x_{t+1} = F_t(x_t) + G_t(x_t)w_t \tag{1}$$

$$y_t = H_t(x_t) + v_t \tag{2}$$

Above (1) is the state-update equation and (2) is the observation equation. In this system x_t is an n—sized state vector (e.g. water volume in the surface streams and channel), w_t the error vector, y_t is the measurement vector (e.g. runoff at some hydrometric station) and v_t is the observation error vector. $G_t(x_t)$ is a non-constant, state-dependent weighting matrix. Both w_t and v_t are assumed to be independent zero-mean white Gaussian noise processes.

In summary, the filtering algorithm is applied in the following way. Given the error covariance matrix P, it is projected with Eq. (3). Then the Kalman gain K is computed from Eq. (4). The measurements enable to estimate v_{t+1} to update the state with Eq. (5) and the error covariance with (6):

$$P_{t+1:t} = F_t' P_{t:t} F_t'^T + G_t' Q_t G_t'^T \tag{3}$$

$$K_{t+1} = P_{t+1:t} H_t'^T \left[H_t' P_{t+1:t} H_t'^T + R_t \right]^{-1} \tag{4}$$

$$x_{t+1:t+1} = x_{t:t+1} + K_{t+1}v_{t+1} \tag{5}$$

$$P_{t+1:t+1} = \left(I - K_{t+1}H_t'\right)P_{t+1:t} \tag{6}$$

As we observed a power-law type of relationship between simulated storage in surface water and observed runoff at hydrometric stations, Nguyen (2011) explored the potential of EKF for flood forecasting in the Lo river using the DIMOSHONG distributed model. While the theory of Kalman filtering, with linear or non-linear models, is well defined for state-space models with a limited number of state variables, more difficult is model updating for distributed hydrological models as the state variables, including surface runoff and subsurface soil moisture, can easily become several thousands, considering the variability in each computational grid cell. In principle each observed variable, as runoff at one gauging station, provides information useful to update one state variable.

In our application to the Lo river runoff observed at each of 14 streamgauges was used to update the surface water storage spatially averaged in the river network upstream the measurement stations. The updated storage is then distributed over each cell with a multiplicative correction. The water stored in the Tuyen Quang and Thac Ba reservoirs, which influence runoff simulations in Chiem Hoa, Tyuen Quang and Vu Quang, was accounted by computing the volume stored during their operation in post processing mode. As shown in Table 1 and, for example, in Fig. 4 the application of the EKF is capable of improving, on average, the model's performances quantified in terms of Nash–Sutcliffe efficiency, RMSE, peak error timing and percentage error. However, as the EKF can lead to unstable results when the nonlinearity in the system is strong, in some cases the model's performances are better without the application of EKF.

3 Meteorological and Hydrological Forecasts in the Ca River

3.1 Verification of Precipitation Forecasts

The accuracy of the BOLAM—MOLOCH precipitation forecasting chain was evaluated by simulating, at the National Center of Hydro-Meteorological Forecasting (NCHMF) based in Ha Noi, 31 precipitation events occurred from 2008 to the end of 2010. Some additional events occurred in 2011, including the September 2011 flood, have been simulated using BOLAM only. The versions of the BOLAM and MOLOCH models delivered by ISAC—CNR in November 2010 were used for the simulations. In computing performance statistics, interpolated forecast precipitation is compared to point measurements at seven raingauge stations in the Ca river, so the representativeness of point values influences the statistics. In an overall

Table 1 DIMOSHONG model performance in simulating daily runoff at four stations in the Lo river basin in standard calibration mode (CAL), using observed rainfall as meteorological forcing, and after the application of the Extended Kalman Filtering (EKF). In *italics,* the best results in the CAL versus EKF comparison

	Chiem Hoa (16,500 km^2)		Ham Yen (11,900 km^2)		Tuyen Quang (29,600 km^2)		Vu Quang (37,000 km^2)	
2006 Event	*CAL*	*EKF*	*CAL*	*EKF*	*CAL*	*EKF*	*CAL*	*EKF*
Nash Efficiency	0.77	*0.87*	*0.74*	0.65	0.90	*0.92*	*0.82*	0.80
Mean flow (m^3/s)	1356	1295	777	700	2354	2180	2544	2440
RMSE (m^3/s)	342	*237*	*287*	305	500	*434*	*7433*	7464
Time error (days)	1.00	*0.00*	1.00	1.00	0.00	0.00	0.00	0.00
Peak error (%)	12.28	*7.33*	*1.07*	9.63	21.05	*4.13*	26.46	*11.93*
2008 Event	*CAL*	*EKF*	*CAL*	*EKF*	*CAL*	*EKF*	*CAL*	*EKF*
Nash Efficiency	0.91	*0.92*	0.51	0.51	0.72	*0.85*	0.57	*0.73*
Mean flow (m^3/s)	1073	1085	1165	1355	2325	2595	2709	2924
RMSE (m^3/s)	186	*176*	*662*	769	714	*565*	965	*759*
Time error (days)	1.00	1.00	5.00	*0.00*	5.00	*2.00*	4.00	*1.00*
Peak error (%)	7.46	*6.57*	*8.27*	29.58	17.19	*1.82*	11.18	*9.63*
2009 Event	*CAL*	*EKF*	*CAL*	*EKF*	*CAL*	*EKF*	*CAL*	*EKF*
Nash Efficiency	0.88	*0.89*	*0.58*	0.54	*0.86*	0.85	0.72	*0.74*
Mean flow (m^3/s)	1041	1043	949	1056	2092	2102	2394	2356
RMSE (m^3/s)	198	*190*	*533*	555	*479*	431	743	*664*
Time error (days)	0.00	0.00	0.00	0.00	1.00	*0.00*	1.00	1.00
Peak error (%)	7.99	*7.70*	*50.35*	63.91	6.54	*0.83*	33.98	*29.77*
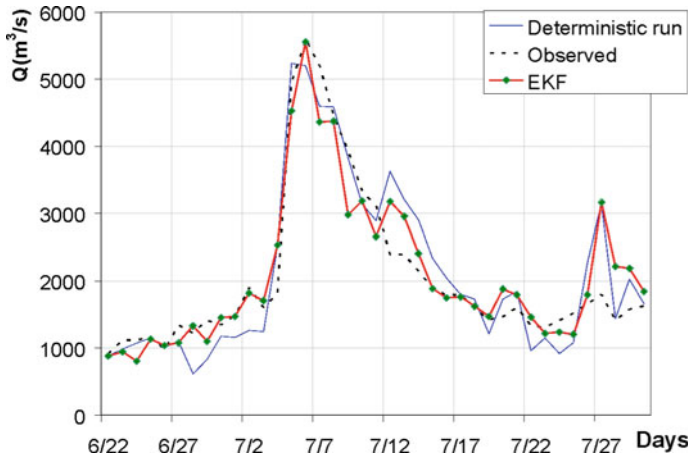
Fig. 4 Comparison of deterministic and Extended Kalman Filtered forecasts with DIMOSHONG versus observations at the Tuyen Quang station during the 2009 flood

assessment of the 31 events, the mean error and correlation coefficient obtained by comparing forecast and observed precipitation accumulated in 24—hours indicate that the BOLAM—MOLOCH chain (in practice, the MOLOCH model) tends to underestimate observed rainfall by a factor ranging from −60% to 0, and the linear correlation coefficient is on average 0.55 ranging from 0.2 to 0.8, depending on the site.

Focusing on the heaviest flood event occurred from 14 to 18 October 2010, meteorological models better perform and the error statistics improve. Between 16 and 17 October, both model forecasts agree very well with the observations, in terms of intensity and location: precipitation exceeding 200 mm in 24 h is centered around Vinh, in the lower part of the Ca-Lam river basin and in the La river basin branch (see Fig. 5, top panels).

The comparison among the fields of precipitation, accumulated over the 120 h period between 12 UTC of 13 October to 12 UTC of 18 October 2010 (shown in Fig. 5, bottom panels), allows to draw a general qualitative evaluation of the models performance in forecasting this severe weather episode. Extremely high amounts of rainfall insisted over a large portion of the Viet Nam territory, both along the coast and over the mountain chain, extending from north of the Ca river region to the city of Hue. Precipitation in excess of 500 mm affected a wide area of Central Vietnam.

BOLAM provides a satisfactory prediction in the region extending south of the Ca riverbasin, slightly underestimating the total rainfall amount, although forecasting more than 400 mm of precipitation. However, forecast heavy rainfall affects only marginally the Ca river, being characterized by a rapid decrease with latitude in proximity of Vinh. Thus, the flood magnitude in the target basin is slightly underestimated.
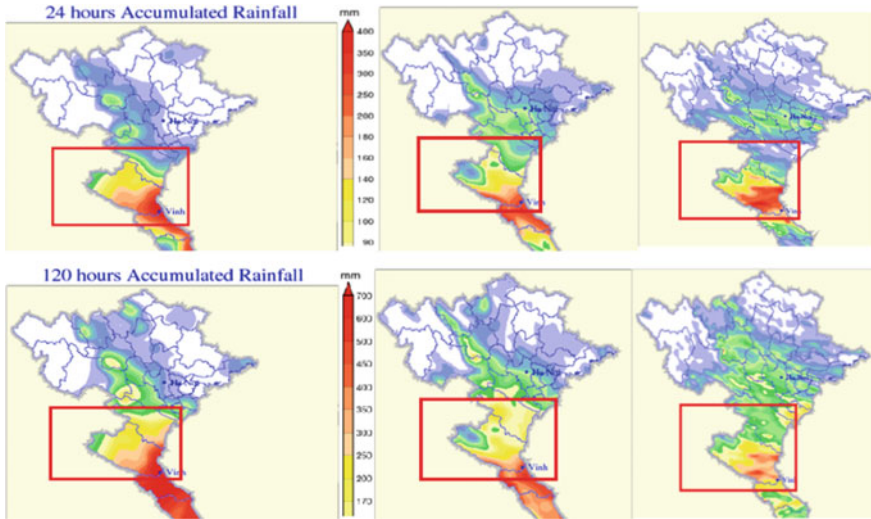
Fig. 5 Comparison of the raingauge-observed (top left panel) versus simulated total 24-h accumulated precipitation at 12 UTC (19 Local Time) of October 17 2010 with the BOLAM model (top central panel) and MOLOCH model (top right panel) during the period of heaviest precipitation in the October 2010 flood on the Ca river (in the red box). In the bottom panels observed 120-h accumulated precipitation until 18 October at 12 UTC, BOLAM-predicted (center) and MOLOCH (right)

The MOLOCH forecast precipitation fields is affected by a general underestimation in the wide scale region, but the prediction is reliable in the Ca riverbasin, both in intensity and location, with an underestimation, however in the La river area, south of Vinh. Overall, the spatial distribution of model precipitation appears in good agreement with the observations and the most intense rainfall is correctly predicted over the mountains. Over the Ca river basin, MOLOCH is apparently able to provide a much better forecast with respect to BOLAM, displaying some maxima along the coast north of Vinh and a wide area of precipitation, whose accumulated amount is around 200 mm.

As in the central provinces rainfall can occur because of a complex combination of weather types, including typhoons and tropical low pressure, Inter-Tropical Convergency Zone, high pressure in the Pacific Ocean, a lookup table of weather conformations producing medium and heavy rains in the Ca river Basin has been proposed by Hoang et al. (2012) to correct, when appropriate, model's predictions.

3.2 Runoff Observations and Forecasts

The October 2010 and September 2011 floods were simulated with a complete BOLAM–MOLOCH–DIMOSHONG hydrometeorological flood forecasting chain (Ngo et al. 2010, Ranzi et al. 2020). The highest water level at Nam Dan, located just upstream the junction with the La river, was 7.44 m on 19 October 2010, 7 a.m. between the level II and level III warning which is set to 7.9 m. But in the La river water level at Linh Cam station reached 7.28 m, over the level III warning set at 6.5 m: it was the fourth highest level recorded since 1975.

The flood forecasts by the DIMOSHONG model with forecast lead time, Δt_f, of 48 h from the forecast time, t_f, were tested in two case: (1) without forecast rainfall and (2) with rainfall forecast from the BOLAM model. In the first case at the time of the forecast only observed rainfall is adopted for the meteorological forcing of the hydrological model. In the second case, at the time of the forecast t_f, the observed rainfall is used until t_f, and for the successive 48 h the rainfall forecast from BOLAM is used as input to the DIMOSHONG model.

The experience in forecasting with DIMOSHONG for some floods such as the October 2010 and the September 2011, with case (1) and case (2) was performed in this research. The result of the forecasting experience indicates that, in case (2), the calculated runoff hydrographs better match with observed runoff than in case (1). In the upstream stations as Nghia Khanh (Fig. 6) and Muong Xen (near the Viet Nam–Laos border), DIMOSHONG using rainfall from BOLAM model shows much more suitable results, also because observed rainfall does not cover the catchment area in Laos. For downstream stations the forecast experiment is not very representative yet because the effect of reservoirs is not taken into account in the present stage of the research project. It is planned to be implemented in the future.
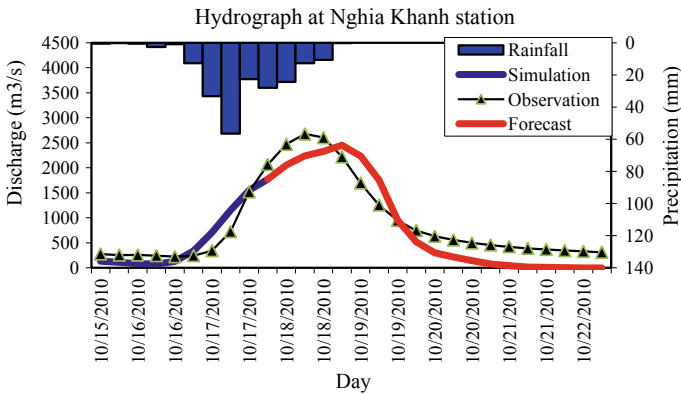
Fig. 6 Flood forecast for the Ca river basin at the Nghĩa Khánh station (area 4020 km^2 in the Hieu river) with the observed precipitation until the forecast time t_f at 01 h on 18th October 2010 h and BOLAM-predicted rainfall thereafter for the forecast lead time Δt_f of 48 h. The hydrograph is simulated with the DIMOSHONG model

4 Conclusions

A hydrometeorological flood forecasting system, complete with two mesoscale meteorological models BOLAM and MOLOCH coupled with the distributed hydrological model DIMOSHONG, was tested for both the Red river and Ca river in Vietnam. The BOLAM hydrostatic meteorological model simulations of five floods occurred in the Red river show quite a satisfactory capability of forecasting rainfall up to three days in advance, with good timing but some overestimation of rainfall amount. Improvements have been achieved with the new version of the model and with the non-hydrostatic MOLOCH model applied to the Ca river. In this case some underestimation of precipitation was observed, confirming that precipitation is a difficult variable to be predicted in tropical areas, because of the combination of different weather types and of the widespread presence of convective instability. The results show, however, the usefulness of meteorological modelling for precipitation estimation in the upstream part of both basins, in China and Laos, in case a lack of raingauge information occurs. Flood hydrographs simulated by the DIMOSHONG model, especially when surface observations are used for model adaptation with the Extended Kalman Filtering, are fairly well reproduced when observed precipitation is used and when the mesoscale model is accurate enough in the rainfall prediction as for the 1971 or 2010 flood. The tests conducted so far encourage the joint effort for a further improvement of the flood forecasting system, considering the regulation of reservoirs and using real-time hydrometric data. The results achieved indicate that the BOLAM–MOLOCH–DIMOSHONG model chain can provide information useful to assist medium term flood forecasting for the Red and Ca rivers, thus enabling flood hazard mitigation measures, as the operation of reservoirs, or the issue of alerts for towns along the rivers.

References

Hoang TT, Ranzi R, Barontini S, Cat VM (2012) Medium range rainfall and flood forecasting for reservoir system operation in the Ca riverbasin (Vietnam). IAHR Asian-Pacific Division Congress, Korea, p 13

Montanari A, Grossi G (2008) Estimating the uncertainty of hydrological forecasts: a statistical approach. Water Resour Res 44:W00B08. https://doi.org/10.1029/2008WR006897

Moradkhani H, Hsu K, Gupta HV, Sooroshian S (2005) Uncertainty assessment of hydrologic model states and parameters: sequential data assimilation using particle filter. Water Resour Res 41:W05012. https://doi.org/10.1029/2004WR003604

Nagata M, Leslie L, Kamahori H, Nomura R, Mino H, Kurihara Y, Rogers E, Elsberry RL, Basu BK, Buzzi A, Calvo J, Desgagne M, D'Isidoro M, Hong S-Y, Katzfey J, Majewski D, Malguzzi P, McGregor J, Murata A, Nachamkin J, Roch M, Wilson C (2001) A mesoscale model inter-comparison: a case of explosive development of a tropical cyclone (COMPARE III). J Meteorol Soc Jpn 79:999–1033

Ngo LA, Vu MC, Hoang TT, Buzzi A, Drofa O, Do LT, Barontini S, Ranzi R (2010) A hydrome-teorological flood forecasting system for the reservoir control in the Red River. In: Proceedings of the of ICOLD symposium Hanoi, 23–25 May 2010, 11pp

Nguyen HS (2011) Real time flood forecasting in the Lo river with surface observations and extended Kalman Filtering. Ph.D. Thesis , Politecnico di Milano, 157pp

O'Connell PE, Clarke RT (1981) Adaptive hydrological forecasting—a review. Hydrol Sci J 26:179–205

Ranzi R, Ngô LA, Hoàng TT, Nguyễn HS, Barontini S, Grossi G, Bacchi B, Buzzi A, Davolio S, Drofa O, Malguzzi P, Đỗ LT, Võ VH, Vũ MC (2020) A hydrometeorological flood fore-casting chain for the Red and Ca rivers (China, Laos and Vietnam). Part I—investigated areas and model setup. In: Anderle M (ed) Innovations in land, water and energy for Vietnam's sustainable development. Springer, Berlin

Todini E, Szollosi-Nagy A, Wood E (1976) Adaptive state-parameter estimation algorithm for real time hydrologic forecasting; a case study. In: IISA/WMO workshop on the recent developments in real time forecasting/control of water resources systems, Laxemburg

Ha Long Bay (Viet Nam) Sediment and Microfaunal Background Assessment for Future Monitoring Actions

Alessandra Negri, Cong Do Thung, Caterina Morigi, Anna Sabbatini, Simona Giunta, Antonietta D'Agnessa, Vu Tran Ngoc Cam, and Massimo Sarti

Abstract In order to assess the quality of the Ha Long bay environment we performed sampling campaign in April 2003, January 2004 and more recently in August 2018. Sedimentological, microfauna and calcareous nannofossil analyses have been conducted on samples collected in different location of the bay. This allowed obtaining preliminary description of the sedimentary environment and a census of the benthic microfauna to assess the biodiversity and to start to create a database in order to monitor the change in time of the organism distribution. In fact, the identification of indicators that, based on the knowledge of past events, can reveal in advance the changes to come in such a way as to proactively prepare the instruments for their mitigation. These are bio-geological indicators whose use in environmental analysis is expected to be very interesting. In particular, a group of protozoa with a calcareous shell (Foraminifera) seems very promising. The fact that these organisms leave an imprint in the ancient sediments (but also simply "old") allows to reconstruct the "pre-impacts", and to compare them with the "impacted" modern conditions therefore to recognize a baseline condition to obtain useful information about possible actions suitable for the protection of the environment. This chapter illustrates the result of our analyses and is a base for future, more detailed studies.

Keywords Ha Long bay · Heritage sites · Impact · Human activities · Ecosystems

A. Negri · A. Sabbatini · S. Giunta · A. D'Agnessa · V. T. N. Cam · M. Sarti
DISVA, Università Politecnica Delle Marche, Ancona, Italy

C. Do Thung (✉)
Institute of Marine Environment and Resources, Vietnam Academic of Science and Technology, Hai Phong City, Vietnam
e-mail: thungdocong@gmail.com

C. Morigi
Dipartimento Di Scienze Della Terra, Università Di Pisa, Pisa, Italy

1   Introduction

Ha Long bay is a UNESCO world heritage sites made up by spectacular limestone karsts pillars and isles (in the north of Vietnam, Quang Ninh Province. Due to its beautiful seascape, it is a popular travel destination although less is known about its biological biodiversity which integrate the scenic landscape. In recent time, it is also developing and amplifying new infrastructures (e.g.: harbour) that represent, together with indiscriminate tourism, form of anthropogenic impact on the biodiversity of the Vietnamese limestone islands. So, this habitat contains highly complex and dynamic ecosystems, strongly impacted by human activities. The complex interplay between the controlling natural and anthropogenic factors is still poorly understood, due to the absence of reliable bio-indicator methods for the quality of these transitional ecosystems.

The high biodiversity of Ha Long bay area is under treat: habitat destruction (deforestation, dredging, dumping, coastal urbanization), oil pollution, eutrophication (aquaculture, agriculture, tourism, industry), industrial continuous pollution, lack of integrated management, no enforcement of management, natural hazards and depletion of natural resources (overfishing).

Population increase induces most of the processes leading to these biodiversity loss causes. Other primary driving forces to loss of biodiversity are the lack of knowledge of local people (fishermen), natural hazards, climate change and the lack of a management plan integrated at all governmental levels and law enforcement. The degradation of the Ha Long bay ecosystem has a direct impact on the local communities. The clearest example is the effect of the depletion of natural resources on the livelihoods of local fishermen and tourism industry (expected if the current trend is maintained). Biodiversity loss also changes the structure of the food web, which leads to a decreased ecosystem health. The resulting impoverished ecosystem will have a negative impact on the appeal of the site as world natural heritage, therefore biodiversity loss will negatively impact the livelihoods of local people depending on tourism. Further analysis revealed some self-enforcing mechanisms in biodiversity loss event chain. As an example, tourism is used to compensate for the unsustainable fishery sector, which leads to a higher food demand by the extra tourists which leads to more aquaculture so more eutrophication, therefore, less healthy ecosystem and less fishery sustainability.

A recent United Nations project (https://www.biodiv.be/cebios2/partnerships/institutional/vietnam) focuses on understanding the processes involved in the first direct cause of biodiversity loss in Halong bay: habitat destruction. Infrastructure developments increase dredging and dumping activities of the area, also the development of more port activities demands more dredging. These activities in combination with coastal erosion due to mangrove cutting and dumping of coal mine waste into the bay damage the sea grass beds and the coral reefs. Aim of the project is the modeling of the sediment flows of the bay to study the effect on coral bleaching, this allowing to discriminate if corals are under treat because of human activities related to turbidity or is climate change with higher temperatures the more endangering factor.

In this context, a joint approach based on biotic indicators like Foraminifera (marine protozoa) and abiotic like sedimentological analyses provide a snapshot of the actual environmental state. In addition, Foraminifera short life cycles respond and register them allowing to assess the healthy status of the environment, impacted by both natural and human activities, by studying the presence/absence of sensitive species and their ecology.

Currently, Foraminifera represent a tool as bio-indicators because they occur in areas where other organisms traditionally used for bio-monitoring are not present and also because their ability to fossilize allows reconstructing pre-anthropic conditions in sites more or less anthropized nowadays.

Through the sedimentological and microfaunal analyses of the collected samples, we aim, at exploring the almost completely unknown biodiversity of Vietnamese Foraminifera and use these organisms to build a new bio-monitoring tool fitted for Vietnam. This approach joint in the future to an effective monitoring system based on environmental and electrical measurements, advanced instrumentation and data acquisition systems, intelligent distributed systems and signal, and image processing, will contribute to better monitor the dynamics of habitats in shallow ecosystems with endangered coral reefs such as Halong Bay, and hence to make the most ecologically sensitive decisions for management, taking into account the ecosystem services for the local communities.

2 Materials and Methods

2.1 Samples Collection

During the oceanographic cruise of April 2003 in the Tonchino Gulf, in the Halong Bay, 7 sites were selected along a transect from 6 to 25 m water depth (Fig. 1 and Table 1); In the framework of the project "Study on the biodiversity of Vietnamese limestone islands, Proposed solutions and models for use, conservation and sustainable development" (IMER) from 20 to 28 August 2018 the Ha Long Bay and four marine karstic lakes, respectively: Hò ca Hong, Hang Sang, But Xam and Me Cung, have been sampled (Table 2) which have peculiar morphologic and physical characteristic as well as marine fauna, different from the one out in the bay, from which they are separated, or in partially connection. Sampling was performed by grab and the sediment was stored in vials with 10% formaline buffered with sodium borate in order to study both modern coccolithophorid and Foraminifera present in the sample.

2.2 Laboratory Preparation

In the laboratory, we subsampled 20 ml of sediment for each sample and we stained them with Rose Bengal solution (1 g l^{-1}) for 24 h (Sabbatini et al. 2012).

Fig. 1 Location of the sampling sites 2004

Table 1 Sample sites of oceanographic cruises April 2003 and January 2004

Sample	Water depth (m)	Latitude (°N)	Longitude (°E)	Sediment data
1	15	20° 46′ 39″	107° 10′ 5″	Silty sand
2	15	10° 47′ 57″	107° 9′ 29″	Clayey sand
3	25	20° 49′ 11″	107° 8′ 10″	Sand
4	18	20° 51′ 10″	107° 8′ 34″	Silty sand
5	12	20° 52′ 14″	107° 6′ 4″	Silty sand
6	8	20° 54′ 50″	107° 4′ 16″	Silty sand
7	6	20° 56′ 55″	107° 4′ 47″	Silty sand

Later on, we washed through 63 microns sieved each sub-sample and sorted in water under a binocular microscope for Rose Bengal stained benthic Foraminifera.

The residues were stored in 10% formaline buffered with sodium borate until the analysis. For the analysis, we counted only the species that presented coloured (pink or red) protoplasm. Taxonomy follows Loeblich and Tappan (1994). Most

Table 2 Sample sites of oceanographic cruise August 2018

Site	Depth (m)	Replicates	Latitude (N)	Longitude (E)
Hò Ca Hong	1.0	1A	20°52'46"	107°12'9"
Hò Ca Hong	1.0	1B	20°52'46"	107°12'9"
Hò Ca Hong	1.0	1C	20°52'46"	107°12'9"
Hò Ca Hong	4.5	1A	20°52'46"	107°12'9"
Hò Ca Hong	4.5	1B	20°52'46"	107°12'9"
Hò Ca Hong	4.5	1C	20°52'46"	107°12'9"
Hang Sang	1.0	2A	20°47'33.0"	107°08'00.5"
Hang Sang	1.0	2B	20°47'33.0"	107°08'00.5"
Hang Sang	1.0	2C	20°47'33.0"	107°08'00.5"
Hang Sang	8.0	2A	20°47'33.0"	107°08'00.5"
Hang Sang	8.0	2B	20°47'33.0"	107°08'00.5"
Hang Sang	8.0	2C	20°47'33.0"	107°08'00.5"
Bui Xam	1.0	3A	20°50'05"	107°07'07"
Bui Xam	1.0	3B	20°50'05"	107°07'07"
Bui Xam	1.0	3C	20°50'05"	107°07'07"
Bui Xam	4.0	3A	20°50'05"	107°07'07"
Bui Xam	4.0	3B	20°50'05"	107°07'07"
Bui Xam	4.0	3C	20°50'05"	107°07'07"
Me Cung	1.0	4A	20°50'01"	107°05'03"
Me Cung	1.0	4B	20°50'01"	107°05'03"
Me Cung	1.0	4C	20°50'01"	107°05'03"
Me Cung	4.0	4A	20°50'01"	107°05'03"
Me Cung	4.0	4B	20°50'01"	107°05'03"
Me Cung	4.0	4C	20°50'01"	107°05'03"

of the well-stained calcareous and agglutinated specimens were mounted dry on micropaleontological slides to provide taxonomic reference material. Sedimentological analyses were conducted on smear slide after observation and count of 300 randomly detected grains under the optical microscope.

Finally, samples preparation, for calcareous nannofossil study, followed standard smear slide technique, with the use of Norland Optical Adhesive. No centrifugation was applied to concentrate the biogenic fraction in order to retain the original composition of the nannofossil assemblage. Qualitative analyses were performed with a light microscope at 1250× magnification.

2.3 Results

From the sediment compositional point of view, after a first speditive analysis under the microscope, three main components were detected:

- 85% Quartz.
- 10%Heavy minerals (Pyrite and hematite).
- 5% Others (siliceous spicules, organic matter, spores, volcanic glass *).

As for the matrix, we distinguish between:

- 90% clay.
- 10% limestone (coccolithophores).

Preliminary results from grain size analyses performed at different depth in the karstic lakes at Ho Ca Hong evidence the occurrence of silt as the modal class, whereas the coarser grains are referred to mollusk shells. Interestingly, at Ho Ca Hong site coal debris appear to drape the bottom surface probably as a consequence of a disastrous flood in the region of Quan Ninh in 2015, where extensive Coal Mines are located.

As for the foraminiferal content, first we want to highlight the abundance of *Ammonia tepida* in the samples 1A-1 m and 4.5 m (Ho Ca Hong) samples (Table 2). Some of these specimens show theratological deformation indicative of stressed condition.

The analysis of the samples collected in 2004 shows significant differences in the foraminiferal assemblages. The density along the transect decreases. We identified 46 taxa of benthic Foraminifera. In the stations 1, 2 and 3 we detected the lower number of taxa. In station 4 to 7 the number strongly increases (Fig. 2). In station 1, the offshore site, the assemblage mainly consists of agglutinated taxa (96%), Lituolids and *Trochammina* spp., whereas Miliolids contribute for the 4% of the assemblage. In the benthic foraminiferal community of station 2 the percentage of Lituolids decreases (33%), whereas *Trochammina* spp. shows the same percentage (47%). Other agglutinated species are present but in lower percentages (*Lagenammina* spp., *Eggerella* spp. and *Siphotextularia* spp.). *Calcareous* taxa (4%) are only represented by Miliolids. The assemblage in site 3 is similar to site 2 but the number of species slightly increases and the number of calcareous taxa decreases.

Fig. 2 Diversity of benthic Foraminifera across the transect collected in January 2004
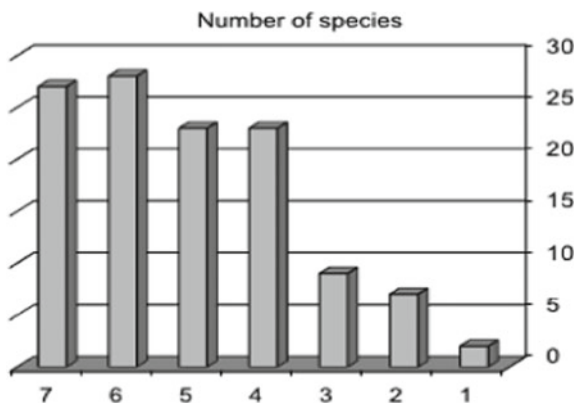
Fig. 3 Density of benthic Foraminifera across the studied transect collected in 2004
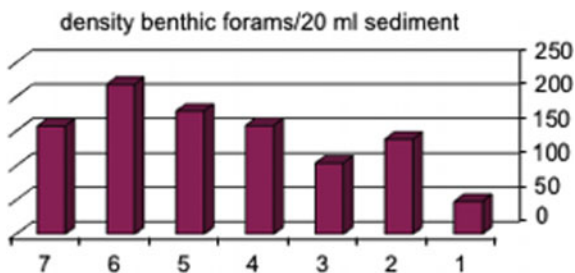
The assemblage in station 4 consists of agglutinated and calcareous taxa in the same percentage. Important species are *Ammonia* spp., *Lagenammina* spp., and *Trochammina* spp. (Fig. 3).

Ammonia spp., occurs for the first time with the 27% of the total abundance. Station 5 is characterised by the lower percentage of agglutinated taxa. In station 6 the density has high value (about 220 number of specimens per 20 ml of sediment) and the community is well diversified. Finally, the assemblage of station 7 shows abundant *Lagenammina* spp., *Ammonia* spp., *Bolivina* spp. and *Cancris auriculus*,

observed few specimens of calcareous nannofossils and a low number of species, although well preserved. *Gephyrocapsa oceanica* (normal size and small size); *Reticulofenestra* spp., *Emiliania huxleyi* (small size) dominate the assemblage whereas some reworked specimens from older sediments (i.e. *Watznaueria ovata*) are always present. Samples 2, 4 and 6 are less rich in coccolithophorids, while samples 6 and 7 show a higher number of species, with the presence of *Helicosphaera carteri* and *Syracosphaera pulchra*, although rare.

3   Discussion and Conclusion

Ha Long bay is characterized by fine quartz with subordinated calcareous biogenic (coccolithophorid) sediment. It is interesting to evidence the occurrence of a coal debris layer draping the Ho ca Hong karstic lake, consequent to the disastrous flood event occurred in 2015 (Quan Ninh). The foraminiferal assemblages occurring at the seven stations along the transect offshore of the Ha Long Bay and at the karstic lake of Ho Ca Hong show remarkable differences. The total density of benthic Foraminifera shows a decrease along the transect from the coast to offshore. The high number of benthic Foraminifera could be a result of a relatively high amount of organic matter input due to the presence of the numerous rivers in the Ha Long Bay. Ideally the high density is a good indicator of the input of the organic matter to the sea floor. Several papers report the link between this physical parameter and the biological response, in fact bottoms with excess food availability can harbour rich benthic communities which suggests that the benthic fauna in many areas may be considered food-limited (Gooday 1986; Jorissen et al. 1992; Rosenberg 1995; Moodley et al. 1998). However,

it is worth mentioning the strong negative correlation between the organic flux to the sea floor and the oxygen consumption. In fact, the benthic foraminiferal faunas are mainly controlled by both parameters: organic flux and bottom water oxygenation. A high food supply could lead to long-term severe hypoxic or anoxic conditions. Although the benthic Foraminifera seem to survive to this stressed situation, the community could be strongly affected and resulting into a low density, low diversity composition and the occurrence of species with an opportunistic behaviour (Jorissen et al. 1995).

The analysis reveals the dominance of agglutinated taxa in the offshore stations (from station 3 to 1, Fig. 4). In these sites, the sediment is composed by coarse sand and silty sand. Taxa with agglutinated test preferably live in coarse sediment (Murray 1991). In deeper sites of the Atlantic Ocean, agglutinated association dominate in sandy, organic poor sediments influenced by strong bottom current (Harloff and Mackensen 1997). Stations 1 to 3 are located along a channel in a morphological depression. In this channel bottom current is usually strong, as testified also by higher content of coarse sand in the sediment. This allows a good oxygenation of the sea bottom that is important for the agglutinated benthic Foraminifera. In the four stations, from 4 to 7 near the coast the benthic association is composed by more that 50% by calcareous taxa with a higher diversity. We may suppose that the environmental condition is optimal to sustain the benthic community and the oxygen concentration to the sea bottom and the organic flux have a good balance that favouring a development of the microfauna. Among calcareous taxa dominate morphotypes of the genus *Ammonia*, a genus highly resistant to environmental variations especially to organic enrichment (Jorissen 1988).

The Ha Long bay contain highly complex and dynamic ecosystems (Do Cong Thung 2000; Do Cong Thung and Sarti 2004; Pham Dinh Trong and Do Cong Thung 1999) that could be strongly impacted by human activities. The low biodiversity and density found in the outer part of the bay could be also explained by an accumulation of natural stress factors, which are today intensified by a strongly increased anthropogenic pressure. The complex interplay between the controlling natural and anthropogenic factors is still poorly understood, explaining the absence
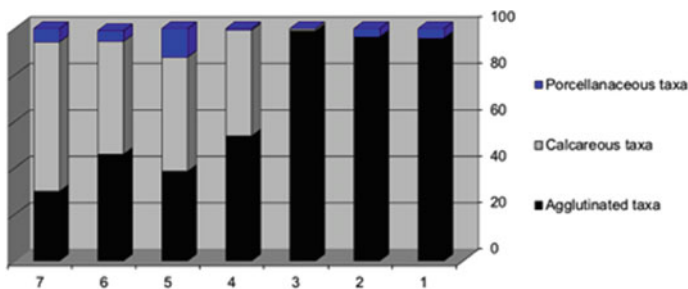
Fig. 4 Variability of foraminiferal test composition across the transect

of reliable bio-indicator methods for the quality of shallow water transitional ecosystems. In this context, we studied the foraminiferal distribution in a transect along the fragile ecosystem of the Ha Long bay and we interpret this distribution in function of the natural environmental factors (physical/chemical factors and sediment). This first census represents an ecological baseline and it will let us to compare data with future faunal distribution analysis to investigate whether increased undergoing anthropogenic impact and climate change are to blame as cause of major faunal changes in this ecosystem. This work has to be intended as an attempt to obtain the background knowledge based on biotic and abiotic indicators which if coupled to effective environmental measurements based on advanced instrumentation and data acquisition systems will allow to evaluate the quality status of marine ecosystems of these areas of Vietnam.

References

Jorissen FJ (1988) Benthic foraminifera from the Adriatic Sea: principles of phenotypic variation. Utrecht Micropaleontol Bull 37:7–139

Jorissen FJ, Barmawidjaja DM, Puskaric S, van der Zwaan GJ (1992) Vertical distribution of benthic Fora-minifera in the northern Adriatic Sea: The relation with high organic flux. Mar Micropaleontol 19:31–46

Jorissen FJ, De Stigter HC, Widmark JGV (1995) A conceptual model explaining benthic foraminiferal microhabitats. Marine Micropaleontol 26:3–15

Loeblich Jr AR, Tappan H (1994) Foraminifera of the Sahul shelf and Timor Sea. Edit by Cushman foundation special publication no. 31

Moodley L, Heip CHR, Middelburg JJ (1998) Benthic activity in sediments of the northwestern Adriatic Sea: sediment oxygen consumption, macro- and meiofauna dynamics. J Sea Res 40:263–280

Murray JW (1991) Ecology and palaeoecology of benthic foraminifera. Harlow Longman Editor

Rosenberg R (1995) Benthic marine fauna structured by hydrodynamic processes and food availability. Neth J Sea Res 34:303–317

Sabbatini A, Bonatto S, Bianchelli S, Pusceddu A, Danovaro R, Negri A (2012) Foraminiferal assemblages and trophic state in coastal sediments of the Adriatic Sea. J Mar Syst 105–108:163–174

Do Cong Thung (2000) Biodiversity of mollusks in Halong Bay and Catba Islands. In: Proceedings of the 11th congress & workshop: tropical marine Mollusk programme, vol II. Phuket Marine Biological Center Special Publication, pp 369–372

Do Cong Thung, Sarti M (2004) Biodiversity conservation in the Coastal zone of Vietnam. Ha Noi National University Publication, p 253

Marine Biodiversity in Ha Long Bay and Cat Ba Archipelago (VietNam)

Do Cong Thung, Nguyen Dang Ngai, Dau Van Thao, Nguyen Van Sinh, Dao Minh Dong, Barbara Calcinai, and Carlo Cerrano

Abstract Cat Ba islands, located in the North East of Vietnam, in the area of Ha Long City (Quang Ninh Province) and Cat Hai Island District (Hai Phong City), is the largest limestone archipelago in Vietnam. Typical limestone islands including tropical rainforests, caves, mangroves, tidal ecosystems, coral reefs, and soft-bottom ecosystems characterized the area. Cat Ba hosts high level of diversity with 4622 recorded species of marine and terrestrial plants and animals. Among animals, 130 species, mainly corals, are threated and are listed in the Vietnam Red Book and in the IUCN Red list; the marine mammal *Sousa chinensis* and the sea turtle *Eretmochelys imbricata* are considered in danger and in an extremely critical state. Saltwater lakes are known to host several endemic and rare species. In Cat Ba these peculiar habitats are numerous and their biotic community includes 165 marine vegetal and animal species; some of them are endemic or rare like several sponge species and the jellyfish *Mastigias* sp. Rare and precious species of global conservation value, the unique naturalistic characteristics of the area, hosting various and rich ecosystems, support Ha Long Bay, Cat Ba Archipelago as a world heritage site according to the criteria of biodiversity (IX and X criteria).

D. C. Thung · N. D. Ngai · D. Van Thao
Institute of Marine Environment and Resources (IMER), Vietnam Academy of Science and Technology (VAST), Hanoi, Vietnam
e-mail: thungdc@imer.vast.vn

N. Van Sinh
Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology (VAST), Hanoi, Vietnam

D. M. Dong
Bach Long Vi Marine Protected Area Management Board, Hai Phong, Vietnam

B. Calcinai · C. Cerrano (✉)
Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Via Brecce Bianche, 60131 Ancona, Italy
e-mail: c.cerrano@staff.univpm.it

1   Introduction

In the North-Eastern Vietnam, Cat Ba Archipelago in Ha Long Bay covers an area over 1000 km^2, and it belongs to Ha Long city (Quang Ninh province) and Cat Hai district (Hai Phong city) (Fig. 1). Cat Ba Archipelago, is characterized by thousands of karsts islands with evident erosive processes generating in tropical moist condition. The impressive naturalistic value of the area is due to its majestic limestone towers, arches and caves caused by coastal erosion (https://whc.unesco.org/en/).

In the area of Bai Tu Long Bay, Ha Long Bay and Cat Ba Archipelago, 119 marine lakes are recorded, covering a total area of 1475.6 ha. The largest lake (28.8 ha) is Ang Vem and the smallest (0.7 ha) is Ang Tre Moi. By using Google Earth, Vermeulen and Anker (2017) counted 138 sea lakes only for Ha Long, Cat Ba (Fig. 1), and in total they estimated about 400 saltwater lakes all over the world. Thus, Halong, Cat Ba holds about 1/3 of the global number of saltwater lakes.

Ha Long Bay, Cat Ba Archipelago is characterized by 7 important ecosystems of limestone islands typical on the tropical seas, including: (i) the tropical rainforest, (ii) the terrestrial cave; (iii) the mangrove, (iv) the tidal, (v) the salt-lakes, (vi) the coral reef, (vii) the soft bottom (Sea Ecosystems). Marine caves are likely present but still unexplored.

The level of diversity and of endemic species in the area is high (21) both for plant and animals such as *Trachypithecus poliocephalus* (Cat Ba Langur), *Goniurosaurus catbaensis* (Cat Ba Tiger Gecko), *Tiwaripotamon edostilus* (Cat Ba cave crab), *Ficus superba* var. *alongensis* (Halong ficus), *Jasminum alongensis* (Halong Jasmin).
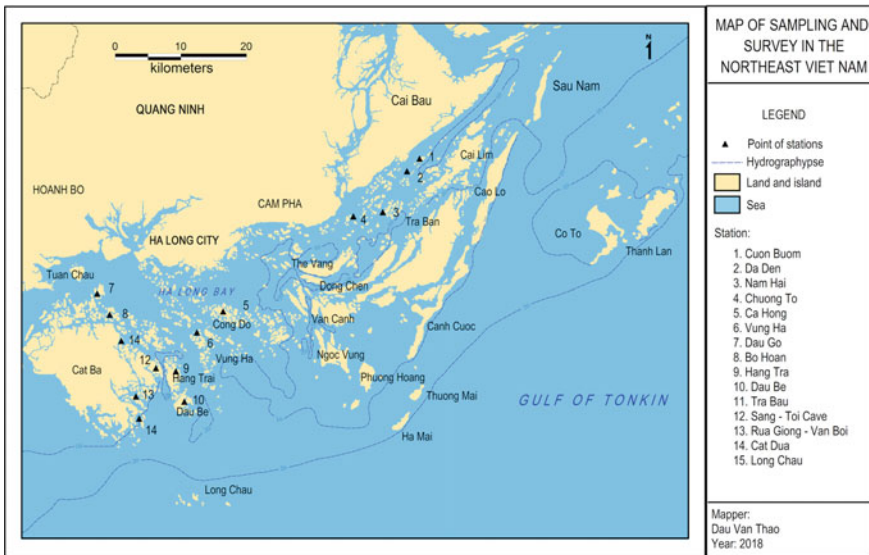
Fig. 1   Ha Long Bay and Cat Ba archipelago and the sampling locations

The present chapter provides an exhaustive report about the marine biodiversity of Ha Long Bay, Cat Ba Archipelago deriving from numerous years of field researches in the area. The peculiar environmental characteristics of the area and its high level of diversity support the area as a World Heritage Site according to the criterion of biodiversity (criterion ix and x https://whc.unesco.org/en/criteria/).

2 Materials and Methods

2.1 Location

Sampling areas include 3 island groups from Ha Long to Long Chau (Fig. 1). In Ha Long, we have monitored the islands of Cong Do (Ca Hong lake), Van Ha, Thien Cung, Dau Go, Thien Long, Bo Nau, Luon Cave, Tien Cave, Hang Trai, Dau Be.

Cat Ba Island represents the limestone island of the UNESCO biosphere reserve, where the influence of fresh water from the continental shelf is particularly evident; here six key areas of study were included: Cat Dua, Van Boi, Van Ha, Hang Toi cave, Hang Sang cave, Tra Bau.

The Long Chau archipelago, is located outside the coastal island system; here the uninhabited location of the Cay Bang Bay area (Fig. 1) was selected.

During three expeditions in April 2003, September 2003 and April 2004 the staff of IMER (VAST) and the Italian staff explored eight saltwater lakes in Ha Long, Cat Ba region. They comprised Me Cung, Bu Xam, Ang Du I, Ang Du II, Ba Ham lake, Ang Vem, Ang Tham, Ang Luon. Among them, particular attention was devoted mainly to Ang Du I and Me Cung being entirely enclosed and not adjoining outside. The other lakes presented different level of connection with the surrounding sea.

2.2 Collection and Analysis of Marine Organisms

Basically, we followed specific regulations of the Provisional Process of the Integrated Marine Survey issued by the State Committee for Scientific and Technological Research (1981), and the process of investigation of marine resources and environment (IMER 2014). In addition, the methodologies of several international organizations (i.e. Reef Check (https://www.reefcheck.org/) and Sea grassNet (https://www.seagrassnet.org/)) were followed too. Assessment of local diversity by quantitative sampling methods was carried out as described by English et al. (1997). Several methods of processing and analysis have been adopted in relation with different groups of organisms (see for example Calcinai et al. 2006; Azzini et al. 2007; Thin 2007; Thung 2014).

3   Results

3.1   Species Diversity

Including both marine and terrestrial species of flora and fauna, up to date 4622 species have been recorded in the area nominated as the World Heritage Site.

Among these, 2464 species, 53.3% of the total, belong to the terrestrial biota; aquatic biota comprises 2158 species (46.7%) and among these, 11 species are fresh water fishes.

Table 1 shows that the major number of species belongs to zoobenthos (734 species, 34.2% of total), followed by phytoplankton (450 species, 21%), marine fish (361 species, 16.8%), corals (247 species 11.5%) and seagrasses with only 4 species (0.29% of the total) (Table 1).

The biotic community in saltwater lakes includes 165 species considering both the vegetal and animal realm. Among them, 21 species are sea weeds, 46 species belong to sponges, 41 species are corals, 37 species are molluscs (19 species belong to the class of Gastropoda and 18 species are Bivalves), 8 species of crustaceans, 6 species of echinoderms, and 6 species are osteichthyes. Common genera of corals are *Acropora*, *Porites*, and *Favia*. Typical molluscs species are *Brachidontes* spp., *Anomalodiscus squamosus*, *Protapes gallus*, *Anadara kagoshimensis*, *Isognomon legumen*, *Pinctada martensii*. Jellyfish, belonging to genus *Mastigias*, appear in large number in 6 investigated lakes.

Considering the zoobenthos, the phylum Porifera has been intensively studied in these past years. In the area, a total of 63 species of demosponges were identified. Some are present in the whole area, while others have been exclusively detected in the marine lakes (Calcinai et al. 2006; Azzini et al. 2007). Unfortunately, according to the last surveys, the sponge diversity seems drastically decreased during the last 15 years (Cerrano et al., present volume).

Table 1   Species diversity of aquatic organisms living in Ha Long Bay, Cat Ba Archipelago

Groups of organisms	Families	Genera	Species	%
Mangroves	24	29	31	1.4
Sea weeds	35	56	160	7.5
Seagrasses			4	0.2
Phytoplankton	38	105	450	21.0
Zooplankton	47	60	154	7.2
Zoobenthos	146	345	734	34.2
Corals	18	55	247	11.5
Marine fish	71	136	361	16.8
Marine mammals and turtles			6	0.3

Table 2 Rare and precious marine species

Taxa	Vietnam Red Book 2007	IUCN Red List 2017
Zoobenthos	6	0
Corals	1	110
Marine turtles	4	3
Marine mammals (*Sousa chinensis*)	1	1
Total number of species	12	114

3.2 The Rare and Precious Species Listed in Vietnam Red Book and in the Red List of IUCN

Among the species recorded in the studied area, 12 are listed in the Vietnam Red Book (2007) and 114 species in the IUCN Red list (2017). The highest species number belongs to the coral group (110 species in the IUCN Red List); other threatened species belong to the zoobenthos, and some species of turtles (3 species for the IUCN and 4 species for Vietnam Red Book). One species of marine mammal (*Sousa chinensis*) is considered in danger (Table 2). The marine turtle, *Eretmochelys imbricata*, is now considered in a particularly and extremely critical state, CR level of IUCN 2012.

3.3 Environmental Factors of Marine Lakes

Environmental conditions of saltwater lakes are different in comparison with those present outside, in the open sea. Especially, temperature of water inside lakes is often high, (29–30 °C), salinity is between 9 and 24‰ (changing according to season); the content of Nitrate and Phosphate is high (Table 3).

4 Conclusions and Discussion

The records of 2147 species of marine organisms in the area clearly evidence that the region of Ha Long Cat Ba is characterized by unique ecosystems, which have high valuable conservation.

Up to now, 4622 species of plants and animals, on both land and underwater, have been recorded. The high level of biodiversity in Ha Long Bay, Cat Ba Archipelago, with over hundreds of rare species and global conservation values, supports the area as a World Heritage Site according to the criterion of biodiversity (criterion ix and x).

Table 3 Environmental factors of lakes in Ha Long Bay, Cat Ba Archipelago (from Ngai et al. 2015 with modifications)

Parameters	Ang Du	Ang Dau Be	Ang Qua bang	Allowable limits
Temperature (°C)	29.0	32.10	30.9	30
DO	8.64	9.03	7.63	≥ 5
pH	7.62	7.80	7.88	6.5–8.5
Salinity (‰)	9	24	23	–
Nitrite (NO_2^-)	5.68	5.35	4.94	<10
Nitrate (NO_3^-)	76.50	108.60	105.6	60
Ammonium (NH_4^+)	32.72	36.57	49.61	70
Phosphate (PO^{3-}_4)	15.84	17.92	18.00	15
COD (mg/l)	1.96	2.16	1.83	3

Unfortunately, owing to the increasing human pressures on the area, numerous species are threatened even if already included in the Vietnam Red Book and in the IUCN Red List.

The numerous marine lakes are precious habitats characterized by peculiar environmental conditions and diversity. Temperature and salinity, for example, strongly varies with season, provoking sharp thermal and osmotic stresses on local populations, forcing them to adapt to these extremely variable environmental conditions.

The saltwater lake ecosystem is a natural laboratory for studying evolutionary adaptation of organisms with the selection of the most adapted isolated species (Swift et al. 2016). Very likely, the data that have been collected up to date, do not manifest all the species diversity of saltwater lakes ecosystem; for example, the Echinoderm and Osteichthyes have been little investigated. Lakes, being ecosystems more or less separated from surrounding sea, are strongly prone to extreme weather condition, climate change and pollution impacts. The recent discovery of a new species of sponge in the Hò Cà Hong lake confirms this peculiarity and highlights the importance to protect and valorise the diversity of the area.

From the above characteristics, saltwater lakes ecosystem and the surrounding sea have created a unique chance to study the relation between the fluctuation of physical climate factors and the evolution of ecology and diversity in marine populations. In this way, they can be considered as uncommon and unrepeatable natural laboratories where short-term effects of climate change may be studied and monitored.

References

Azzini F, Calcinai B, Cerrano C, Bavestrello G, Pansini M (2007) Sponges of the marine Karst lakes and of the coast of the islands of Ha Long Bay (North Vietnam). In: Custodia MR, Lobo-Hajdu G, Hajdu E, Muricy G (eds) Porifera research: biodiversity, innovation and sustainability. Museo Nacional, Rio de Janeiro, pp 157–164

Calcinai B, Azzini F, Bavestrello G, Cerrano C, Pansini M, Thung DC (2006) Boring sponges from the Ha Long Bay (Tonkin Gulf, Vietnam). Zool Stud 45(2):201–212

Cerrano C, Bavestrello G, Bertolino M, Thung DC, Pansini M, Sarti M, Núñez Pons L, Calcinai B. The Ha Long Bay marine ecosystem. An unprecedented opportunity for evolutionary studies on marine taxa, present volume

English S, Wilkinson C, Baker V (1997) Survey manual for tropical marine resources, 2nd edn. Australian Institute of Marine Science, Townsville, 390 p

Institute of Marine Resources and Environment (2014) Process of investigation and survey of marine resources and environment. Science and Technology Publishing House, 291 p

Ngai ND, Van Thao D, Thung DC, Thuy LT, Tien DD, Van Quan N, Van Chien P (2015) Biological community in submerged caves and marine lakes in Ha Long-Cat Ba area, Vietnam. J Life Sci 9:541–548. https://doi.org/10.17265/1934-7391/2015.11.006

State Committee for Science and Technology (1981) Provisional proceedings of the integrated marine survey. Scientific and Technical Publishing House, 205 p

Swift HF, Gomez Daglio L, Dawson MN (2016) Three routes to crypsis: stasis, convergence, and parallelism in the Mastigias species complex (Scyphozoa, Rhizostomeae). Mol Phylogenet Evol 99:103–398115

Thin N (2007) Methods of plant research. Publishing House Vietnam National University, 171 p

Thung DC (2014) Conservation of coastal biodiversity in Vietnam. Scientific and Technical Publishing House, 420 p

Vermeulen JJ, Anker K (2017) Outstanding global values in geology and environment in Cat Ba archipelago and Halong Bay. In: Proceeding of the biodiversity conservation solutions in Ha Long bay and Cat Ba archipelago. IUCN Vietnam, pp 16–23

The Ha Long Bay Marine Ecosystem. An Unprecedented Opportunity for Evolutionary Studies on Marine Taxa

Carlo Cerrano, Giorgio Bavestrello, Marco Bertolino, Maurizio Pansini, Laura Núñez-Pons, Massimo Sarti, Do Cong Thung, and Barbara Calcinai

Abstract The Ha Long Bay (Vietnam) is a shallow area located in the northern part of the Tonkin Gulf, in the South China Sea. It includes more than 3000 islands of variable sizes. A continuous action of karstic processes, initiating around 280 million of years ago, shaped the formation of extraordinary caves and shallow salt-water lakes. These impressive structures recall the marine lakes found in Indonesia and Palau, creating one of the most important UNESCO's World Heritage Sites. The protected area covers a surface of 434 km^2 and comprises 775 islands, forming a triangle with the Dau Go Island (Driftwood Island) to the west, the Ba Ham Lake (Three Shelter Lake) to the south, and the Cong Tay Island to the east. During the Pleistocene, multiple sea level rise episodes that span for 10,000 – 15,000 years contributed to the geographic isolation of marine taxa. Along the coasts of Ha Long Bay islands, several coral species formed very shallow reefs that host peculiar sponge assemblages adapted to low depth and murky waters. In the marine lakes instead, corals can be quite rare while sponges are common. Their diversity though, is only known for a few lakes. Extreme variations in environmental conditions occur yearly in the lakes, due to heavy monsoon rains that cause stratification of the water column and drastic alterations in the communities. The aim of this study is to review the sponge diversity inside marine lakes and along the coastal areas of this peculiar

C. Cerrano (✉) · M. Sarti · B. Calcinai
Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, 60131 Ancona, Italy
e-mail: c.cerrano@staff.univpm.it

G. Bavestrello · M. Bertolino · M. Pansini
Dipartimento di Scienze della Terra dell'Ambiente e della Vita, Università di Genova, Corso Europa 26, 16132 Genova, Italy

L. Núñez-Pons
Department of Integrated Marine Ecology (EMI), Stazione Zoologica Anton Dohrn (SZN), Napoli, Italy

D. C. Thung
Institute of Marine Environment and Resources (IMER), Vietnam Academy of Science and Technology, VAST, Hanoi, Vietnam
e-mail: thungdocong@gmail.com

region, comparing data collected more than fifteen years ago with recent data. During three expeditions between 2003 and 2004 we recorded 63 demosponges, of which 46 were living in and outside marine lakes, 17 were characteristic outside the lakes, and 23 were exclusive inside the lakes. After 15 years, a survey in August 2018 was performed to combine morphological data with DNA barcoding, to reveal and enhance the value of the biodiversity of the area. The diversity of sponges displayed drastic recession, showing the decline of more than half of the species typically present. In lieu, the exploration of a new lake (with very low salinity and high-water temperature) unveiled a seemingly new sponge species able to endure these conditions. Our findings suggest the urgency to develop measures of protection for these extraordinary but endangered environments, where biodiversity is still poorly explored. Indeed, these characteristic systems subjected to such intense geophysical isolating mechanisms represent optimal natural laboratories for the study of evolution and speciation of marine biota developing from segregated peripheral populations.

1 Introduction

The Ha Long Bay is located in the northern part of the Tonkin Gulf, in a shallow area of the South China Sea harboring more than 3000 islands of diverse sizes. The main characteristic of the bay is the presence of karsts peaks that are to be found also elsewhere in the tropics, in a vast geographic area including Thailand, Southern China and Philippine Islands. The Ha Long Bay main calcareous islands are as high as 400 m and the maximum water depth of canals in between can be 25 m. These karst towers arose owing to the uplift of oceanic sediments after the subsidence of the massive Pacific plate (Fenart et al. 1999). When the area got flooded due to raising sea-level (about 18,000 years ago, after the last glacial maximum) the Ha Long Bay towers became islands.

Karstic processes, enhanced by the tropical conditions, carved out of the limestone many depressions, leading to the formation of shallow salt-water lakes. Lakes are characteristic marine habitats, differing one another, according to the degree of isolation from the surrounding sea. Although they are no more in pristine conditions, having been exploited for mollusc harvesting and aquaculture by local fishermen, they may give new insights into evolutionary processes, clarifying the degrees of isolation and genetic connectivity of their fauna. In a few lakes, the communication with the surrounding sea is detectable by the flow of tidal streams. Sometimes artificial canals have been built to enhance this water exchange. However, most of the lakes are apparently closed, being connected to the sea only through the cavities of the karsts system. They host a peculiar fauna of stingless jellyfish (Cerrano et al. 2006) and may be compared, for their environmental and biological conditions, with the meromictic marine lakes of Palau (Hamner and Hamner 1998; Dawson and Hamner 2005) and with the anchialine lakes of East Kalimantan (Indonesia) (Becking et al. 2011).

In the surrounding sea, the shallow water bay environment is characterized by low water movement due to the presence of islands, variable salinity around 32‰, and varying water temperatures between a minimum of 23 °C at the end of winter and 29°–30 °C in the late summer (Tang 2001), and high sedimentation rates making water generally turbid.

The anchialine lakes may be regarded as small marine habitats surrounded by land, as islands are surrounded by sea. Even if they are communicating with the sea, there is always an environmental hindrance that marine organism larvae have to overcome to enter the lake, and vice versa. Such a situation reminds of—in the marine environment—the concept of insularity (McArthur and Wilson 1967; Johnson et al. 2000; Hubbell 2001). In such "islands of sea" the biodiversity is an equilibrium between immigration and extinction processes strongly affected by sharp temperature and salinity variations, which may enhance endemism, probably through selection mechanisms.

Porifera are among the most represented phyla and their diversity in the area has not been intensely studied. The first papers were those from Lindgren (1898), Dawydoff (1952) and Lévi (1961), and more recently new records were added by Calcinai et al. (2006) and Azzini et al. (2007). In 2013, Thai reviewed the literature of the area and updated the sponge fauna of Vietnam counting 201 species in total.

The aim of this chapter is to review the sponge diversity inside marine lakes and along the adjacent coasts of this unique region, adding new data obtained by a recent survey conducted in the area to suggest adequate measures of conservation.

2 Materials and Methods

Past surveys were conducted in April 2003, September 2003, and April 2004, within a cooperation project among the Polytechnic University of Marche (UNIVPM), University of Genoa (UNIGE) and the Institute of Marine Environment and Resources (IMER) of Hai Phong for the study of biodiversity and conservation in a coastal area of Vietnam. The lakes studied in the past surveys were (i) open lakes, Hang Du II, Hang Tham, Hang Luong, (ii) semi-enclosed lakes, Dau Be, Cat Ba, Me Cung and (iii) enclosed lakes, Hang Du I, Bui Xam (see Cerrano et al. 2006 and Azzini et al. 2007 for details). Fifteen years later, in August 2018, a survey was organized in the lakes Hang Du I, Bui Xam, Me Cung and in Ho Ca Hong, this last one surveyed for the first time. Sponge collection was qualitative, and samples were collected by SCUBA diving, by visually oriented transects. Inside the lakes, samples were collected down to 3 m of depth, and outside the islands, on the shallow reefs, down to 7–8 m depth. Regarding the old samplings, the general physical-chemical parameters of the lakes are reported in Cerrano et al. (2006).

After collection on board, samples were divided into~ 3 cm^3 pieces and fixed in ethanol 70% for morphological observations, andin 4% formaldehyde solution. Smaller fragments were preserved in absolute ethanol for DNA barcoding

approaches. Voucher specimens in 70% and formaldehyde are deposited at UNIVPM, while 100% ethanol fragments are at the EMI Dept. of SZN.

Spicule and skeletal preparations were made according to the techniques described by Rützler (1978) for light and electron microscopy (SEM). Mean length and width values were calculated after measuring at least 25 spicules per type.

3   Results

During the past surveys 182 specimens were collected leading to the identification of 63 demosponges. In the lakes, 46 sponge taxa were found, while 40 were the sponges identified in the bay. The sponges exclusive from the lakes were 23, while 17 were detected only outside the lakes. The most common species in all the studied area was *Dysidea cinerea* Keller, 1889 (in the 60% of stations), followed by *Tethya seychellensis* (Wright 1881) (in the 53%). Other 23 species were found both in the lakes and in the bay. The highest number of sponges (29) was recorded in Bui Xam; in the lake of Hang Du I, intensely studied, only 10 sponges were identified. *Suberites diversicolor* (in Azzini et al. 2007 listed as *Suberites* sp. 1) was found in Hang Du I and in Dau Be only, communicating with the sea by small conduits. Other two sponges such as *Pione carpenteri* and *Spirastrella decumbens* were exclusively present in the enclosed lakes of Hang Du I and Bui Xam.

During the last survey (2018) we have detected 29 sponge taxa (this includes 9 new records), which increase the total number of sponge species in Ha Long bay to 72. Comparing the records reported by lake across previous documented surveys, here we recorded only 5 sponge taxa out of 10 previously found in Hang Du I lake (*C. aurivilli*, *Pseudosuberites* sp., *Suberites diversicolor*, *Suberites* sp. 1 and *Tethya seychellensis*) and 14 species out of 29 in Bui Xam lake. *Amorphinopsis fenestrata*, a new record for the area, and was exclusively found in this lake though. In Me Cung lake we found 8 species out of 10 previously recorded. The new species, *Cladocroce* sp., already found in the past surveys in five lakes and in two coast sites (Azzini et al. 2007), was collected again in Bui Xam and Me Cung lakes, as well as outside Bui Xam and Hang Du I lakes. The lake Ho Ca Hong, characterized by extreme environmental conditions, and surveyed only in the last expedition, hosted a single species of sponge; it was identified as *Spongilla* sp. that is a species new to science (under description) and the first record of a fresh water sponge for Vietnam. The species *Suberites diversicolor*, in the previous surveys identified as *Suberites* sp. 1, and found in the last survey (2018) exclusively in Hang Du I, is a new record for Vietnam. *Hyrtios erectus*, present outside Hang Du I karst lake, is a new record for Ha Long Bay, while *C. burapha* and *Pseudoceratina verrucosa* (Table 1) are two new records for Vietnam.

Table 1 List of the sponge species collected from lakes (in) and coastal areas (out) in Ha Long Bay, during past (see Azzini et al. 2007) and present surveys

	in				out			
	Hang Du I	Bui Xam	Me Cung	Hò Cà Hong	Hang Du I	Bui Xam	Me Cung	Hò Cà Hong
Acanthella hispida (Pulitzer-Finali, 1980)					X			
Amorphinopsis excavans (Carter, 1887)		X						
Amorphinopsis fenestrata (Ridley, 1884)		X						
Chondrilla australiensis (Carter, 1873)		X			X			
*Cladocroce burapha** (Putchakarn et al., 2004)		X			X	X		
Cladocroce sp.		X	X		X	X		
Cliona aff. *celata* (Grant, 1826)		X						
Cliona orientalis (Thiele, 1900)		X						
Cliothosa aurivilli (Lindgren, 1897)	X							
Dactylospongia sp.						X		
Dysidea sp.			X					
Echinodictyum asperum (Ridley and Dendy, 1886)			X					
Haliclona (*Gellius*) *cymaeformis* (Esper, 1794)		X			X	X		
Haliclona (*Haliclona*) sp. 1					X			
Hyrtios erectus (Keller, 1889)					X			
Mycale (*Zygomycale*)*parishii* (Bowerbank, 1875)			X					
Mycale (*Mycale*) *philippensis* (Dendy, 1896)		X	X					
Neopetrosia sp.								X
Protosuberites sp.		X						
Psammocinia sp.			X					
*Pseudoceratina verrucosa** (Bergquist, 1995)							X	
Pseudosuberites sp.	X							
Spongilla sp. *				X				
*Suberites diversicolor** (Becking and Lim, 2009)	X							

(continued)

Table 1 (continued)

	in				out			
	Hang Du I	Bui Xam	Me Cung	Hò Cà Hong	Hang Du I	Bui Xam	Me Cung	Hò Cà Hong
Suberites sp. 1	X							
Siphonodicyon mucosum (Bergquist, 1965)		X						
Spheciospongia solida (Ridley and Dendy, 1886)		X					X	
Spongia sp.		X	X			X		
Tethya seychellensis (Wright, 1881)	X	X	X					

New records for Vietnam are marked with *

4 Discussion

The total number of species recorded in the area during all the periods of sampling (72) (Table 1, Azzini et al. 2007) is not particularly high considering other similar studies in tropical areas (Calcinai et al. 2017). The number of species found in the nine surveyed lakes (52) is comparable with the number of sponge species (45) in four anchialine lakes of East Kalimantan (Voogd et al. 2006).

Salinity variations, causing rapid environmental changes, especially in the most sheltered marine lakes of the Ha Long Bay, could have had a selective action on the sponge fauna. In Ho Ca Hong lake only one species (*Spongilla* sp.) tolerates the extreme high temperature and low salinity conditions recorded during the summer.

The differences in the number of sponge species recorded in two enclosed lakes of comparable size and characteristics, Bui Xam with 29 species and Hang Du I with only 10, were striking; the new survey seems to confirm these differences (14 vs. 5 respectively) and unfortunately puts in evidence an equally striking loss of diversity.

The differences in diversity between the two lakes may be due mainly to a different level of water exchange with the open sea, causing in Hang Du I lake high degree of stratification of the water column leading to intense physical disturbances (Cerrano et al. 2006). Very few cases of marine fauna mortality following intensive rains are documented (Goodbody 1961) and closed bays, coastal areas (Cerrano et al. 2001) and lagoons and marine lakes are obviously more exposed to these disturbances. Persistence of high water temperatures in late summer may stress benthic organisms trigging mass mortalities, as already recorded for sponges and gorgonians (Cerrano et al. 2000). Seasonal variations, such as those recorded in Hang Du I lake, are not so extreme and long lasting to cause mass mortalities, but remarkably they do affect the sponge fauna. A similar scenario seems to occur in Ho Ca Hong, where a single species of sponge was detected probably due to the peculiar environmental conditions found in the period of August (summer).

The diversity recession we have registered is diffused and cannot be imputed to seasonal cycles, because surveys were conducted in similar periods (end of August and September), and the decrease involves all the surveyed areas, also Me Cung lake, considered semi-enclosed.

High level of sedimentation rate is also a factor negatively affecting filter-feeding organisms such as sponges (Di Camillo et al. 2013; Bell and Smith 2004). The general deterioration of the area, correlated with an increase of sedimentation could explain the recession of some excavating sponges such as *Pione carpenteri*, *S. tentorioides*, and *C. hancocki*.

Distribution of species from Ha Long bay encopasses the Indian and Western Pacific Oceans, the Indo-Malayan region and Australia; another species as *S. diversicolor* was described in enclosed lakes similar to the Vietnamese lakes (Becking and Lim 2009) and has been recently recorded in Hawaii (Núñez-Pons et al. 2017); genetic analysis of connectivity could clarify the rate of connectivity among Vietnamese lakes, surrounding areas and the rest of biogeographic living in sympatry with the congeneric and widespread pacific sponge *Cladocroce burapha*.

Further work is needed to fully describe this particular area, but, overall, the number of recorded species may not be considered low, taking into account the uniformity of the surveyed environment. The last surveys put in evidence a general decrease of the diversity, confirming the urgency of developing effective management and protection actions of this peculiar and valuable marine region. Evaluating negative trends through historical data is highly relevant for understanding these habitats from an evolutionary point of view. We greatly encourage developing tailored measures of protection for these extraordinary but endangered environments, where many species might indeed disappear before being described and studied. Our concern is tightly related to the fast development of the area, in what regards a continuous tourism growth.

References

Azzini F, Calcinai B, Cerrano C, Bavestrello G, Pansini M (2007) Sponges of the marine karst lakes and of the coast of the islands of Ha Long Bay (North Vietnam). In: Custódio MR, Lobo-Hajdu G, Hajdu E, Muricy G (eds) Porifera research: biodiversity. Innovation and sustainability, Rio de Janeiro, pp 157–164

Becking LE, Lim SC (2009) A new *Suberites* (Demospongiae: Hadromerida: Suberitidae) from the tropical Indo-West Pacific. Zoologische Mededelingen 83:853–862

Becking LE, Renema W, Santodomingo NK, Hoeksema BW, Tuti Y, de Voogd NJ (2011) Recently discovered landlocked basins in Indonesia reveal high habitat diversity in anchialine systems. Hydrobiologia 677:89–105

Bell JJ, Smith D (2004) Ecology of sponge assemblages (Porifera) in the Wakatobi region, south-east Sulawesi, Indonesia: richness and abundance. J Mar Biol Assoc U K 84:581–591

Calcinai B, Azzini F, Bavestrello G, Cerrano C, Pansini M, Thung DC (2006) Boring sponges from the Ha Long Bay (Tonkin Gulf, Vietnam). Zool Stud 45(2):201–212

Calcinai B, Bastari A, Bavestrello G, Bertolino M, Horcajadas SB, Pansini M, Makapedua DM, Cerrano C (2017) Demosponge diversity from North Sulawesi, with the description of six new species. ZooKeys 680:105–150

Cerrano C, Bavestrello G, Bianchi CN, Cattaneo-Vietti R, Bava S, Morganti C, Morri C, Picco P, Sara G, Schiaparelli S, Siccardi A, Sponga F (2000) A catastrophic mass-mortality episode of gorgonians and other organisms in the Ligurian Sea (north-western Mediterranean), summer 1999. Ecol Lett 3:284–293

Cerrano C, Magnino G, Sarà A, Bavestrello G, Gaino E (2001) Necrosis in a population of *Petrosia ficiformis* (Porifera, Demospongiae) in relation with environmental stress. Ital J Zool 68(2):131–136

Cerrano C, Azzini F, Bavestrello G, Calcinai B, Pansini M, Sarti M, Thung DC (2006) Marine lakes of karst islands in Ha Long Bay (Vietnam). Chem Ecol 22(6):489–500

Dawson MN, Hamner WM (2005) Rapid evolutionary radiation of marine zooplankton in peripheral environments. Proc Natl Acad Sci U S A 102(26):9235–9240

Dawydoff C (1952) Inventaire des animaux benthique récoltés par moi dans le domaine maritime Indochinois. Poriferes, in: Contribution a l'étude des invertébrés de la faune marine benthique de l'Indochine. Suppléments au Bulletin Biologique de France et de Belgique, vol 37, pp 46–51

De Voogd N, Weerdt WH, van Soest RWM (2006) The sponge fauna of the Anchialine Lakes of Kakaban and Maratua (East, Kalimantan, Indonesia). In: Custodio MR, Lobo-Hajdu G, Hajdu R, Muricy G (eds) Biodiversity, innovation, sustainability: book of abstracts. 7th International sponge symposium, Armaçao de Buzios, Rio de Janeiro, Brazil, 7–13 May 2006, Museu Nacional, RJ

Di Camillo CG, Bartolucci I, Cerrano C, Bavestrello G (2013) Sponge disease in the Adriatic Sea. Mar Ecol 34(1):62–71

Fenart P, Cat NN, Drogue C, Van Canh D, Pistre S (1999) Influence of tectonics and neotectonics on the morphogenesis of the peak karst of Halong Bay, Vietnam. Geodinamica Acta 12:193–200

Goodbody I (1961) Mass mortality of a marine fauna following tropical rains. Ecology 42(1):150–155

Hamner WM, Hamner PP (1998) Stratified marine lakes of Palau (Western Caroline Island). Phys Geogr 19:175–220

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography, vol XIV. Princeton University Press, Princeton, New Jersey, 375 p

Johnson KP, Adler FR, Cherry JL (2000) Genetic and phylogenetic consequences of island biogeography. Evolution 54:387–396

Lévi C (1961) Eponges Intercotidales de Nha Trang (Viet Nam). Archives de Zoologie Expèrimentale et Génerale 100:127–150

Lindgren NG (1898) Beitrag zur Kenntniss der Spongienfauna des Malayischen Archipels und der Chinesischen Meere. Zoologische Jahrbücher Abteilung für Systematik, Geographie und Biologie der Thiere 11:283–378

McArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton University Press, Princeton, New Jersey, p 224

Núñez-Pons L, Calcinai B, Gates RD (2017) Who's there? – First morphological and DNA barcoding catalogue of the shallow Hawai'ian sponge fauna. Plos ONE. https://doi.org/10.1371/journal.pone.0189357

Rützler K (1978) Sponges in coral reefs. In: Stoddart DR, Johannes RE (eds), Coral reefs: research methods. Monographs on oceanographic methodology. Vol. 5. Unesco, Paris, pp. 299–313

Tang VT (2001) The Eastern Sea resources and environment. The Gioi Publishers, Hanoi

Thai MQ (2013) A review of the diversity of sponges (Porifera) in Vietnam. In: The 2nd international workshop on marine bio-resources of Viet Nam, pp 109–115

Pathobiome Studies as a Way to Identify Microbial Co-operators and/or Antagonists of the Incoming Plant Pathogen

Cristina Bez, Hang Dinh Thuy, Minh Nguyen Hong, Iris Bertani, and Vittorio Venturi

Abstract Why should there be a collaboration project involving rice, Italy and Vietnam? There are many reasons starting from rice itself. Rice cultivation is fundamental for the nutrition of the world population: it is grown in more than hundred countries, with a total harvested area of approximately 158 million hectares. The annual production is around 700 million tons, with 90% occurring in Asia (https://ricepedia.org/rice-as-a-crop/rice-productivity). Will this production be sufficient in the next decades? World population is estimated to reach 9.7 billion people by 2050 and 11.2 billion people by 2100 (UN in United Nations, Department of Economic and Social Affairs, Population Division World Population Prospects, 2017). How to feed this large population and how to do it in a sustainable way will be a major future challenge. The current methodologies in rice cultivation are no longer sustainable due to the strong environmental impact they bring with such as deforestation, water scarcities, soil depletion and high levels of greenhouse gas emissions. New ways that protect and enhance yields in a sustainable and integrated manner, are needed to decrease the use of chemicals (fertilisers and pesticides) and improve productivity which will also need to cope with the decrease of agricultural land, the increment of no-food crops, the rural depopulation (more than two-thirds of the world population in 2050 will live in urban areas—UN in United Nations, Department of Economic and Social Affairs, Population Division World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352), 2014), pests, diseases and climate changes. Unfortunately, the current rate of progress in agriculture is not sufficient to eradicate hunger even by 2050, so the development of sustainable technologies is a worldwide priority (FAO in The future of food and agriculture. Trends and challenges. FAO, Rome, 2017). Rice cultivation, due to its diffusion, has also strong impact on atmosphere (de Miranda et al. in Am J Plant Sci 6:2009–2018, 2015), on soils and on health of consumers and farmers due to the large use of fertilizers, pesticides and herbicides. In Vietnam, for instance, the value of pesticide market

C. Bez · I. Bertani · V. Venturi (✉)
Bacteriology Laboratory, ICGEB, Padriciano 99, 34149 Trieste, Italy
e-mail: Vittorio.Venturi@icgeb.org

H. D. Thuy · M. N. Hong
IMBT, Hanoi, Vietnam

was estimated to be more than 800 million USD in 2015 (Vietnam Pesticide Association: https://cacasiasummit.com/Uploads/Download/7-Vietnam%20Agricuture%20and%20Status%20of%20Pesticides%20Market.pdf) and is expected to reach 6.8% of CARG annually in 2016–2020 (Mordor Intelligence: https://www.mordorintelligence.com/industry-reports/vietnam-crop-protection-chemicals-market).
Currently Italy is one of the top ten pesticide consumers in the world (Pretty and Bharucha in Insects 6:152–182, 2015). Therefore, studies focused on the research of new avenues for a more sustainable rice agriculture are urgently needed. Moreover, rice represents, more than other cereals, a cultural link between different countries and people, and, as consequence of this, interest for this crop crosses boundaries. Italy is the first European rice producer with 216.019 hectares used for this crop. Italian rice belongs to the Japonica group and some varieties are cultivated just in Italy contributing to enhance and spread the Italian culture in the world. Rice is also Vietnam's most important crop (7 million hectares are dedicated to rice cultivation) and is the primary staple food of the Vietnamese diet. Vietnamese rice belongs to the Japonica variety too, and its cultivation, as the Italian one, is under biotic and abiotic stresses related to drought, climate changes and pests.

1 Rationale and Project Description

The microbial community that lives in association with the plant is called microbiome and is considered as a second plant genome (Berendsen et al. 2012). It plays a crucial role in plant health, nutrient uptake and stress tolerance. Plant microbiomes can be a significant ally for the plant controlling colonization/infection by plant pathogens. Alternatively part of the plant microbiome could be an ally to an incoming pathogen resulting in a multispecies cooperative microbial disease (Lamichane and Venturi 2015). The concept of monostrain/monospecies infections is now challenged and different studies indicate interactions between pathogens and the residential microbiota (Hosni et al. 2011; Kolenbrander et al. 2010; Guo et al. 2014). Novel methodologies allow the analysis of the total microbial community associated with a particular niche or host, thus opening new avenues to the studies on the role of microbiome in plant health and disease. Microbial members of the plant microbiome compete for nutrients, produce anti-microbial molecules or lytic enzymes (proteases, lipases, chitinases, etc.), consume pathogen stimulatory compounds (Doornbos et al. 2012; Lugtenberg and Kamilova 2009) or boost the immune system of the plant for a rapid defence response (Van der Ent et al. 2009), thus being important tools and ally to control colonization and/or infection by plant pathogens. Comparing rice microbiomes of healthy and infected plants from the same location could provide insights on the potential role of the microbiome in rice resistance towards incoming bacterial pathogens thus providing a way to recognize and use them as microbial biocontrol agents against pathogens.

It is currently unknown what is the role and potential effect on the microbiome of two important bacterial diseases of rice: bacterial blight, caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), endemic in Asia, and the foot rot, caused by *Dickeya zeae*, an emerging rice disease in Italy. Understanding the plant microbiome at the site of infection (this is called "pathobiome") of these two rice diseases could reveal potential commensal/resident bacteria that can cooperate with the pathogen. On the other hand, comparing the pathobiome with the plant microbiome of healthy plants harvested in the same field, could highlight microbial taxa which can keep away the pathogen from colonizing the plant.

This study is designed to compare the plant microbiome present in healthy rice plants harvested in Italy and in Vietnam with the pathobiome of rice infected by *D. zeae* (in Italy) and *X. oryzae* pv. *oryzae* (*Xoo*, in Vietnam). During several working meetings and visits between the two research partner groups, the experimental approach and methodologies were planned such as sampling methods, DNA purification procedures, 16S amplicon libraries preparation, DNA sequencing methodologies and bioinformatic analysis to be performed. The experimental design involved a total of 262 rice plants in 5 locations and 5 different rice varieties (two in Italy and three in Vietnam) in order to have a set of statistically significant data for answering the following questions, (i) which bacterial genera are responsible for the core microbiome of the rice plant independently from locations, cultivation methodologies and growing seasons, (ii) which are the overall differences in the microbial community composition of diseased and healthy plants and (iii) which are the specific bacterial taxa enriched in healthy and diseased microbiomes. Collected data will allow identifying the bacteria genera enriched in healthy plants, which could be responsible for a higher degree of resistance of the plant to the disease; these enriched taxa could be potential antagonists of the pathogens and be used as biocontrol agents. Similarly, bacterial taxa which are significantly enriched in diseased samples, might be involved in teaming up with the incoming pathogen exacerbating and accelerating the development of the disease thus being putative partners of the pathogen. The ultimate goal is to isolate microbial strains to develop as bioinoculants for the biocontrol of the two rice diseases. In addition, it is of interest to identify pathogen associated partners in the microbiome in order to study signalling and metabolic complementarity pathways between members of the rice microbiome and the pathogen.

2   Methodology

Italian rice samples have been harvested from two fields of the SAPISE Italian rice cooperative organization located in the rice producing area of Vercelli (Italy), where *D. zeae* represents an emerging issue for rice farmers. Two rounds of sampling in two years representing two rice growing seasons have been carried out with the aim of increasing the statistical power of our analysis. During the first sampling (July 2017), healthy and *D. zeae* infected rice plants have been collected from the same field; samples belong to two Italian rice varieties, which show different sensibility
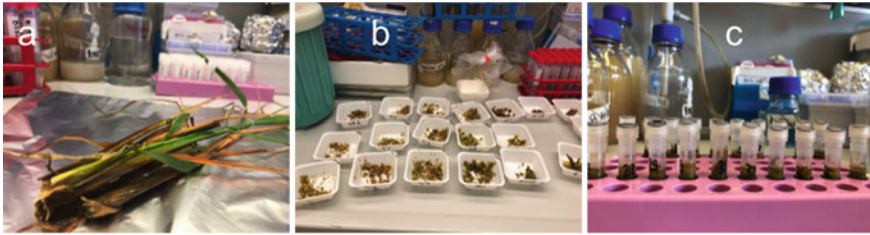
Fig 1 **a** Rice plant affected by foot rot disease caused by *Dickeya zeae*. **b** Samples treatments stage. **c** Bacterial DNA extraction phase

to the pathogen. The selection of the infected plants was performed by two plant pathologists, who recognized the typical symptoms of bacterial foot rot disease which are characterized by yellow and dry leaves, black rot and foul-smelling base and roots. The second round of sampling (July 2018) has been performed in the same fields of the first, following the same procedure protocol.

In Vietnam, three rice cultivars from three representative rice cultivating areas (Southern, Middle and Northern parts in Vietnam) were selected for collecting the samples. Samples had bacterial blast *Xoo*-infection symptoms and in the same fields samples of healthy plants were also taken. The selection of the *Xoo*-infected plants was carried out by plant pathologists: symptoms of this disease appear on the leaves of young plants as yellowish-with and wavy edges. In fact, these bacteria infect rice plants leaf veins as well as the xylem causing blockage and plant wilting.

Once in the laboratory, the plant material was washed and for each infected sample a small region between the lesion and the green part of the plant stem has been selected and weighed. The same region from healthy plant stem has been cut and treated in the same way in order to reduce differences. Later, for each sample, half of the plant material was crumbled and ground rapidly to a fine powder in liquid nitrogen using sterilized pestle and mortar for bacterial DNA extraction (Fig. 1) and 16S rRNA gene amplicon library preparation. The residual sample portion was macerated in 4 mL PBS solution and stored with 18% glycerol at −80 °C for bacterial strains isolation to be performed at a later step.

2.1 DNA Extraction, 16S rRNA Gene Amplicon Library Preparation and Sequencing

DNA was extracted using the PowerMax Soil DNA isolation kit (MO BIO Laboratories, Carlsbad, CA, USA) following the manufacturer's protocol and using, as starting material, 0.25 g for each sample. DNA concentration was analysed by using a Nanodrop^{TM} spectrophotometer (Thermofisher Scientific Inc., Waltham, MA, USA) and settled to a concentration of 7 ng/µl for preparing the 16S rRNA gene amplicon library. All DNA samples had absorbance ratios of $A_{260}:A_{230} > 1.7$

and $A_{260}:A_{280} > 1.8$. The 16S rRNA amplicon library was prepared following the Illumina Inc.'s protocol (https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf—Illumina Inc., San Diego, CA, USA). Briefly, individual barcoded libraries are directly prepared by PCR using long primers incorporating the Illumina adapter sequences, which allow to pool together and sequence a large number of libraries (Fig. 2). Samples were amplified in the V3 and V4 regions of the 16S rRNA sequence using degenerated primers (Table 1) selected from Klindworth et al. (2013), in 25 cycles PCR, followed by an AMPure XP beads clean-up (A63880l—Beckman Coulter Inc., Brea, CA, USA). A second PCR reaction was performed to attach dual indexes and Illumina sequencing adapters using the Nextera XT Index Kit (these indexes are used by the Illumina MiSeq sequencing equipment as specific identifiers for each sample) followed by a second AMPure XP bead clean-up. Each sample give rise to a 16S amplicon library.

After the library preparation, each amplification was checked using the Bioanalyzer equipment (Agilent Inc., Santa Clara, CA, USA), which provides information about size, integrity, and purity of the amplified. 16S rRNA gene concentration was measured by fluorimetric quantification using Qubit 2 (Invitrogen Inc., Carlsbad, CA, USA) and each library sample was settled to a concentration of 4 nM. A final
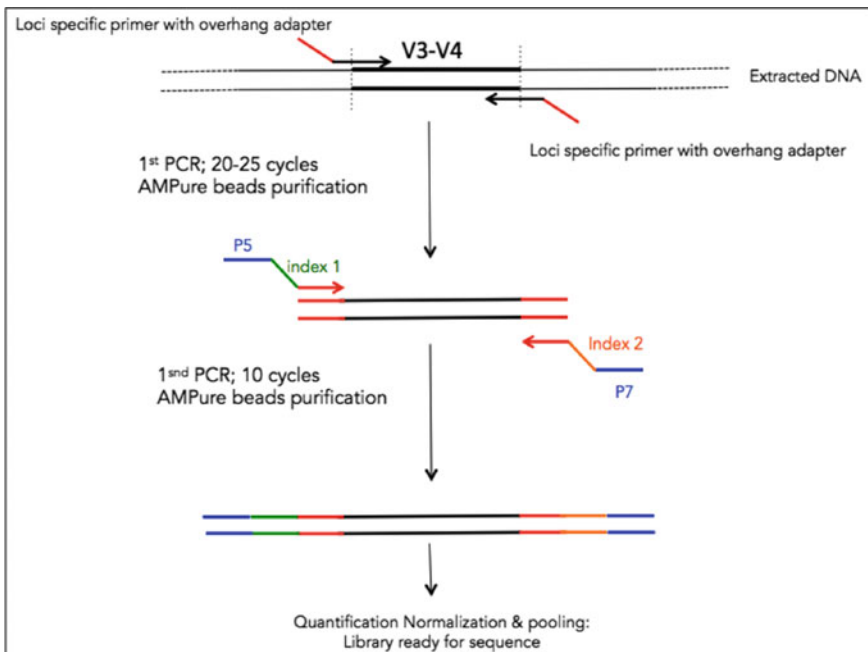
Fig. 2 16S V3-V4 Amplicon Workflow for 16S rRNA amplicon library preparation (extracted and modified from: https://eu.idtdna.com/pages/education/decoded/article/16s-rrna-indexed-primers-amplify-phylogenic-markers-for-microbiome-sequencing-)

Table 1 List of primers used

Name of the primer	Sequence	References
Illumina-16S Amp PCR Fw Primer	5′-tcgtcggcagcgtcagatgtgtataagagacagcctacgggnggcwgcag-3′	Klindworth et al. (2013)
Illumina-16S Amp PCR Rw Primer	5′-gtctcgtgggctcggagatgtgtataagagacaggactachvgggtatctaatcc-3′	Klindworth et al. (2013)
fD1	5′-agagtttgatcctggctcag-3′	Weisburg et al. (1991)
rP2	5′-acggctaccttgttaggactt-3′	Weisburg et al. (1991)
907R	5′-ccgtcaattcmtttragttt-3′	Lane (1991)
785F	5′-ggattagataccctggta	Nakagawa and Kawasaki. (2001)

sample, called "pool of libraries", prepared by adding 5 µl of each sample at 4 nM, was then sequenced.

Sequencing was performed using the Illumina Miseq technology generating 25 million sequencing reads in a single run with a 2 × 250 paired-ends lengths; the Miseq machine parameters were set to produce two separate fastq files for each sample.

2.2 Sequence analysis

Sequence reads obtained via the Illumina Miseq of the 16S amplicon libraries (see above) were processed and analyzed using scripts and programs from DADA2 v1.1.5 (Callahan et al. 2016) and Phyloseq (McMurdie and Holmes 2013). The paired reads were trimmed at position 240 (trimming the last 10 nucleotides) for the forward reads and the reverse reads at position 160 where the quality distribution crashes. Reads were demultiplexed using script from DADA2 pipeline with a minimum phred score of 20. Dereplication was performed combining all identical sequencing reads into "unique sequences" with a corresponding "abundance" equal to the number of reads of that unique sequence. The forward and reverse reads were merged together to obtain the full-denoised sequences with and overlap of at least 12 bases. Subsequently reads were filtered to remove ambiguous reads, low quality sequences, singletons, reads shorter than 200 bp or longer than 275 bp. After removing chimeras, the sequences were grouped into OTUs (Operational Taxonomic Units) with a 97% of identity and each OTU was annotated with RDP reference database, giving a final

OTU table. This table contained the microbiome of each sample classified by taxonomic ranks (Kingdom, Phylum, Class, Order, Family and Genus). Relative abundances of OTUs between samples and the comparative analysis of species richness and diversity indices (Chao1, Shannon, Simpson, ACE and good coverage) among samples were statistically analysed using the vegan package version 2.5-4 (Oksanen et al. 2013) and phyloseq package (McMurdie and Holmes 2013) in R version 3.5.2 (R Core Team 2014).

To explore variation in bacterial community compositions among samples, a correspondence analysis (CA) was conducted in R and the relationships between community structures were studied by principal coordinate analysis (PCoA). To test for differential representation of microbial taxa in different conditions the Deseq2 package (Love et al. 2014) was used.

2.3 Isolation of Culturable Bacteria from Rice Healthy Plants

One of the main aims of this bi-lateral project was to isolate from samples of healthy rice plants, a collection of culturable rice-associated bacteria, which could be involved in disease resistance and be used as natural biocontrol agents against the two studied pathogens.

The macerated plant tissue stored with 18% glycerol at −80 °C was used for the isolation and different dilutions were plated on 1/10 Tryptic Soy Agar (TSA; Difco, BD Laboratories, MD 21152 USA) medium. The choice of a suitable growth medium is crucial during isolation, because different media will obviously affect both the number and diversity of bacteria that can be isolated from a specific plant tissue. Plates were incubated at RT for 5 days. Independent colonies of putative rice associated bacteria showing distinct colony morphology were picked and streaked again on 1/10 TSA plates to ensure purity of the culture. These single colonies were stored at −80 °C in 1/10 TSA and 18% glycerol. The Italian group assembled a bacterial collection obtained from healthy plants collected in *D. zeae* infected field and another collection of bacterial strains was obtained from *Xoo* rice samples (sick and healthy) by the Vietnamese laboratory.

2.4 Screening for Natural Biocontrol Agents Against Dickeya zeae *and* Xanthomonas oryzae *pv*. oryzae *(Xoo)*

Bacterial isolates from the culture collection were tested for antibacterial activity in vitro and isolates with antagonistic capability against *Xoo* and *D. zeae*, using the "agar disk-diffusion method" were identified. Briefly, agar plates were inoculated with a standardized inoculum of the pathogen. Then, filter paper discs (about 6 mm in diameter), containing the test isolate at a desired concentration, were placed on the

agar surface. Antimicrobial agents most commonly diffuse into the agar and inhibit the growth of the pathogen. The Petri dishes were incubated at 28 °C for 24 and 48 h and then the diameters of inhibition growth halos were measured. Amplification of the 16S rRNA gene was then performed on the putative bacterial strains producing anti-microbial compounds by using fD1Funi 16S and rP2Runi 16S primers (Table 1) in order to identify the bacterial isolates. Colony PCR was performed after boiling (10′ at 98 °C) a colony suspension in 50 µl of sterile H_2O. PCR products were purified by using Gel extraction and PCR Clean-Up System purification kit. The sequencing performed with primers 907R and 785F (Table 1) was realized by GATC (Eurofins Genomics Company) and identification of the isolates was obtained by BLAST analysis at NCBI (https://www.ncbi.nlm.nih.gov).

2.5 *Identification of Bioactive Compounds from Selected Candidates*

The candidate bacterial strains showing in vitro antimicrobial properties against *D. zeae* and *Xoo* were grown for 48 h in TSA 1/10 medium. The bacterial cells were pelleted by centrifugation at 10,800 g for 10′ and the cell free supernatant was collected. Briefly, the cell free supernatant was extracted twice by adding an equal volume of organic solvent; different solvents as ethyl acetate, butanol or hexane were used to recover the largest set of compounds. The collected organic layers were combined, washed with distilled water (1 volume), and evaporated to dryness in rotary evaporation at 40 °C. The dried residue was dissolved in methanol and then subjected to HPLC analysis. The analysis for the purification and identification of the compounds are now running in the Vietnamese laboratory.

2.6 **In Planta** *Test on the Selected Bacterial Isolates to Control the Two Bacterial Pathogens*

In *planta* experiments will be set up in order to test the antimicrobial activities of the selected bacterial candidates against the two bacterial pathogens (*D. zeae* in Italy, *Xoo* in Vietnam). The *in planta* assays will follow three different inoculation methods; seed and/or root submersion, spray and microencapsulation (coating of bacteria with a biodegradable polymer to ensure a slow release of the bacteria in the soil by pore-assisted diffusion or degradation of capsules). This last method is particularly promising because encapsulated bacteria will be released slowly and continuously in the plant-soil microenvironment and they are expected to possess an extended shelf life at room temperatures. Furthermore, this method reduces the risk of bacterial transport in water and soil as well as minimizes the spread of hazardous residues in various environments.

3   Results

Similar to the human gut, plants are tightly colonized by complex microbial communities. The plant microbiome plays a crucial role in plant health, nutrient uptake and stress tolerance, but the role of microbes in controlling or aggravating the pathogen attacks remain poorly understood. By studying and comparing the bacterial community composition of two rice plant diseases (rice bacterial foot rot and bacterial blight) these functions of the plant microbiome could be unveiled.

The advent of next generation sequencing has greatly facilitated our ability to study complex microbial communities at an unprecedented scale. Amplicon-based community profiling approaches provide insights on the total hidden community structure and phylogenetic diversity of the plant associated microbes. The challenge now is to understand how to use the generated microbiome and metagenomic datasets to elucidate meaningful relationships between microbes and between microbes and plants.

The large amount of microbiome data obtained during the past 2 years of this bilateral project are now under analysis using the protocols described in "*Methodology*". Until now a preliminary part of the analysis has been performed by the Italian laboratory on a small set of healthy and *D. zeae* infected samples.

Regarding this limited analysis, after the raw sequence reads quality control and filtering, a total of 3,593,080 high quality sequencing reads were obtained with a median read count per sample of 52,030. The high-quality reads were clustered using >97% sequence identity into 6848 bacterial OTUs (Operational Taxonomic Units). Low-abundance OTUs were discarded as well as chloroplast sequences, resulting in 6350 OTUs. The OTUs were clustered at Genus taxonomic level obtaining a final number of 244 different bacterial taxa.

Firstly, the bacterial species richness and the diversity of the two communities (healthy and diseased) were investigated. Measurement of within-sample diversity (α-diversity) was calculated using Shannon H′ index, which takes into account both the richness of a sample and the evenness of taxa.

The alpha diversity index indicates that the numbers of different taxa is significantly higher among the healthy plants compared to the infected one. The decrease in number of diversity in the infected plants could be the result of the pathogen activity, which has the ability to produce a potent antibiotic called zeamine (Zhou et al. 2011) reducing and shaping in this way the microbiome of the plants. The bacterial community of the examined healthy plants is characterized by an extremely high diversity confirming the results of others studies performed on the microbiota composition of healthy rice plants (Edwards et al. 2015; Hardoim et al. 2011).

Successively, a β-diversity analysis, comparing the microbial community composition between samples, was performed to know how similar or dissimilar bacterial communities are in term of phylogenetic distance (Fig. 3).

The comparison of the microbial profile among healthy and infected samples demonstrates that each sample is clustered by symptomatology and that the two groups of samples (healthy and diseased) are clearly separated, indicating significant
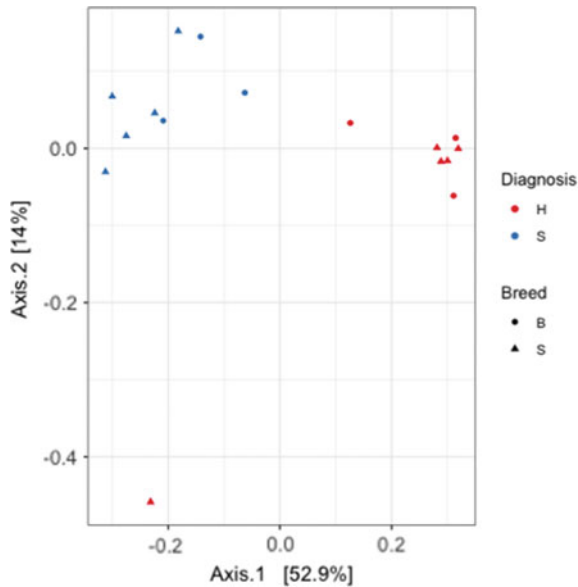
Fig. 3 PCoA analysis of UniFrac distances. In this two-dimensional plot each point represents the entire microbiome of a single sample and each axis reflects the percent of variation between samples, with the x-axis representing the greatest dimension of variation and the y-axis the second greatest dimension. The colors refers to the two different study groups

differences among them. According to the expectations, the most evident difference was the presence and the amount of the pathogen, *D. zeae*: it was found out only in the infected plants and with a high relative abundance (Fig. 4). This result shows that the profile of the bacterial community in rice plants is affected by the presence of the pathogen, which cause a significant change in the community taxonomy composition; even at Phylum level (Fig. 5) there are significant differences in the community composition between healthy and infected samples.

Currently the computational analysis for most of the samples is still in progress and, since data processed are not complete, the core and variable microbiota of the aereal parts of healthy and diseased rice are, at this point, just sketched. Differences between the bacterial communities colonizing diseased and healthy rice have been evidenced and bacterial taxa, putatively responsible for making the plant either more resistant or more susceptible to the bacterial pathogens, have been recognised, but confirmation and validation are needed and will be provided by the analysis of the totality of the samples and by the set-up of in vivo co-inoculation experiments.

These results will highlight the role of the interactions among rice and the microbiome providing information on unexplored diversity. In particular taxa highly present in the pathobiome will provide clues on possible commensal microbes involved in the disease. Isolation of them will be the next challenge together with the investigation of the genetic mechanisms of this pathogen-commensal interaction.
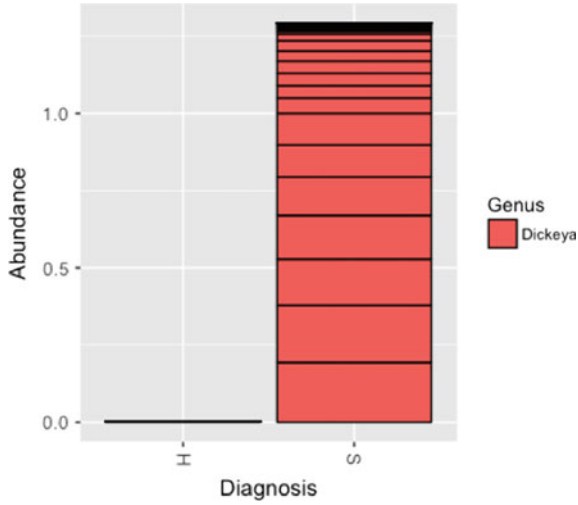
Fig. 4 *Dickeya zeae* relative abundance in healthy samples (H column) and diseased samples (S column).
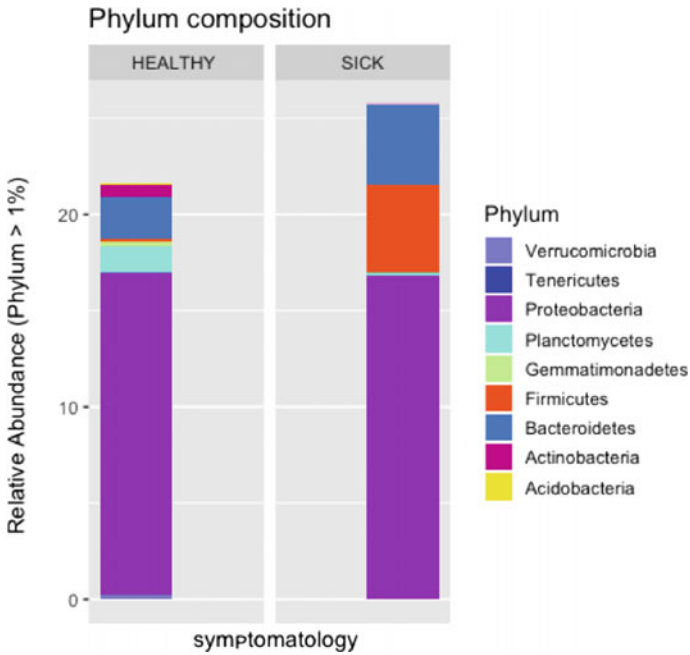
Fig. 5 Taxonomy composition at Phylum level. All the healthy samples are grouped and summed together in the first column (HEALTHY) and the totality of infected plants in the second (SICK)
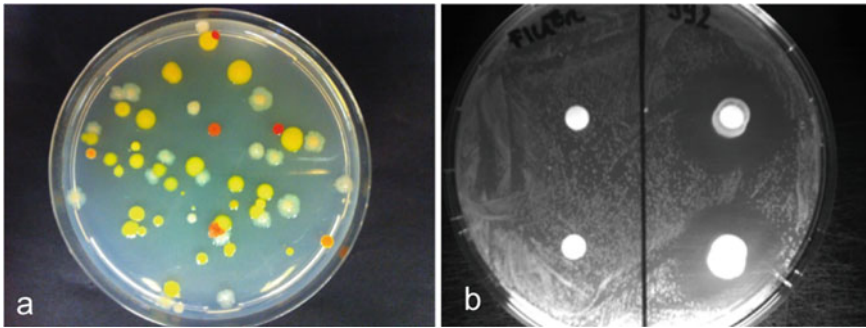
Fig. 6 **a** Isolation of bacteria from plant samples. **b** Screening for antagonistic strains against *Dickeya zeae*

On the other hand, taxa highly present in the microbiome of healthy samples will provide clues on possible microbial antagonists and their isolation will add new members to the bacterial collection from healthy samples already created in Vietnam and Italy (Fig. 6a). The microbial collections are used for screening for antagonistic capability against *Xoo* and *D. zeae* (Fig. 6b). Several strains tested have shown to be antagonist of *Xoo* and on the basis of their 16S rRNA gene sequence, they are members of the *Bacillus* and *Pantoea* genera. Similarly several isolates are antagonists of *D. zeae* and some were identified as belonging to the *Burkholderia* genus. At the moment, we are investigating the physiology of the selected strains with the aim of identify optimum condition for having antagonistic effect against *Xoo* and *D. zeae*. Once identified, these bacteria will be tested *in planta*, via the microencapsulation method, for their ability to control the bacterial pathogens in vivo making the rice plants more resistant.

The final data of this bilateral project will provide avenues for the design of microbial solutions for a more sustainable rice agriculture, with advantages for both the research groups and countries. To conclude, it is important to mention that the collaboration established by the two groups has allowed the scientific formation of two Ph.D. students and the organization of one international workshop on the role of plant microbiome on plant health.

References

Berendsen RL, Pieterse CM, Bakker PA (2012) The rhizosphere microbiome and plant health. Trends Plant Sci 17(8):478–486

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583

de Miranda MS, Fonseca ML, Lima A, de Moraes TF, Rodrigues FA (2015) Environmental impacts of rice cultivation. Am J Plant Sci 6:2009–2018

Doornbos R et al (2012) Impact of root exudates and plant defense signaling on bacterial communities in the rhizosphere. A review. Agron Sustain Dev 32:227–243

Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. Proc Natl Acad Sci 24:112(8)

FAO (2017) The future of food and agriculture. Trends and challenges. FAO, Rome. https://ricepedia.org/rice-as-a-crop/rice-productivity

Guo L, He X, Shi W (2014) Intercellular communications in multispecies oral microbial communities. Front Microbiol 5:328. https://doi.org/10.3389/fmicb.2014.00328

Hardoim PR, Andreote FD, Reinhold-Hurek B, Sessitsch A, van Overbeek LS, van Elsas JD (2011) Rice root-associated bacteria: insights into community structures across 10 cultivars. FEMS Microbiol Ecol 77(1):154–164

Hosni T, Moretti C, Devescovi G, Suarez-Moreno ZR, Barek Fatmi M, Guarnaccia C et al (2011) Sharing of quorum-sensing signals and role of interspecies communities in a bacterial plant disease. ISME J 5:1857–1870. https://doi.org/10.1038/ismej.2011.65

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horm M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 2013:41

Kolenbrander PE, Palmer RJ Jr, Periasamy S, Jakubovics NS (2010) Oral multispecies biofilm development and the key role of cell–cell distance. Nat Rev Microbiol 8:471–480. https://doi.org/10.1038/nrmicro2381

Lamichane JR, Venturi V (2015) Synergisms between microbial pathogens in plant disease complexes; a growing trend. Trens Palnt Sci 6:385

Lane DJ (1991) 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds) Nucleic acid techniques in bacterial systematics. Wiley, New York, pp 115–175

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550

Lugtenberg B, Kamilova F (2009) Plant-growth-promoting rhizobacteria. Annu Rev Microbiol 63:541–556

McMurdie PJ, Holmes S (2013) Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE 8(4)

Nakagawa K, Kawasaki H (2001) Determination method of 16S rRNA gene sequence. In: The Society for Actinomycetes (ed) In the isolation and characterization of actinomycetes. Business Center for Academic Societies, Japan, pp 88–117

Oksanen J, Blanchet FG, Kindt R, Legendre P et al (2013) Vegan: community ecology package, 2013. R-package version 2.0-10

Pretty J, Bharucha ZP (2015) Integrated pest management for sustainable intensification of agriculture in Asia and Africa. Insects 6:152–182

R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

UN (2014) United Nations, Department of Economic and Social Affairs, Population Division World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)

UN (2017) United Nations, Department of Economic and Social Affairs, Population Division World Population Prospects

Van der Ent S et al (2009) Priming of plant innate immunity by rhizobacteria and b-aminobutyric acid: differences and similarities in regulation. New Phytol 183:419–431

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 697–703

Zhou J, Zhang H, Wu J, Liu Q, Xi P, Lee J, Liao J, Jiang Z, Zhang LH (2011) A novel multidomain polyketide synthase is essential for zeamine production and the virulence of *Dickeya zeae*. Mol Plant Microbe Interact 24(10):1156–1164

Sustainable Methods to Control *Pyricularia oryzae*, the Causal Agent of Rice Blast Disease

Luca Sella, Van V. Vu, Alessandra Quarantin, Rocco Caracciolo, Rakshita Govind, Angela Bolzonello, Silvio Tundo, Marta De Zotti, Francesco Favaron, Hoang D. Nguyen, Quynh L. Le, Trung T. Nguyen, Le T. Do, and Hung M. Nguyen

Abstract By 2030, global rice production will need to increase to meet the demand of the growing world population. However, rice is severely affected by blast disease caused by the fungus *Pyricularia oryzae*, which accounts for 10–30% yield losses per year globally, thus posing a threat to the world's most important food crop. The fungus can infect all parts of the rice plant, including leaves, nodes and panicles. In the early stages of the infection process, *P. oryzae* forms an infection structure called appressorium to break the plant cuticle. After an initial biotrophic phase, the fungus kills plant cells thus leading to visible symptoms. Currently, several methods for the management of rice blast disease, such as agronomic practices, use of resistant cultivars and synthetic fungicides, are available and can be exploited in an integrated management approach. However, despite some resistant rice varieties that have been developed by breeding programs, *P. oryzae* is able to rapidly develop new races that overcome resistance genes. Furthermore, the continuous use of fungicides may be harmful to humans and environment and can increase the risk of appearance of resistant fungal races. Biological control using microbial agents or plant extracts with antimicrobial activity is therefore considered a possible alternative and sustainable approach to control rice blast disease. However, biological protection is difficult to achieve in the field as its effectiveness is variable, highly depending on formulations

L. Sella (✉) · A. Quarantin · R. Caracciolo · R. Govind · A. Bolzonello · S. Tundo · F. Favaron
Department of Land, Environment, Agriculture and Forestry (TESAF), University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy
e-mail: luca.sella@unipd.it

V. V. Vu
Nguyen Tat Thanh University, 298-300A Nguyen Tat Thanh Street, District 4, Ho Chi Minh City, Vietnam

M. De Zotti
Department of Chemistry (DISC), University of Padova, Via Marzolo 1, 35131 Padova, Italy

H. D. Nguyen · Q. L. Le
Institute of Tropical Biology, Vietnam Academy of Science and Technology, 9/621 Hanoi highway, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam

T. T. Nguyen · L. T. Do · H. M. Nguyen
Duy Tan University, K7/25, Quang Trung, Da Nang, Vietnam

and climatic conditions. Because of these drawbacks, durable control of rice blast still represents a substantial challenge. Innovative strategies for the identification of new molecules for the sustainable protection of rice are thus highly desirable.

1 Introduction

Rice (*Oryza sativa* L.) is the second largest cereal crop in world production after wheat (FAOSTAT 2014). At global level, paddy production is around 700 million tons (FAOSTAT 2012). Although predominant in Asia, where about 90% of the world's rice is grown and consumed (Pooja and Katoch 2014), this crop is also cultivated in Europe, mainly in Mediterranean countries such as Italy (FAOSTAT 2016). Rice is the most consumed staple food for more than a third of the world's population. In particular, rice accounts for approximately 23% of per capita calorie intake worldwide (Wilson and Talbot 2009), and more than 50% in Asian countries such as Vietnam (Gianessi 2014). By 2050, global rice production will have to double to meet the demands of growing world population, without negatively affecting basic resources (Khush 2005; FAO 2009). The expected increase will have to be obtained with less land and water available for rice users as well as by using a reduced amount of fertilizers and chemicals (Ribot et al. 2008).

This challenge needs the development of high yielding rice varieties tolerant to biotic and abiotic stresses. In fact, diseases reduce the yield potentiality of rice from 10 tons/ha to an average harvesting of about 5 tons/ha (Miah et al. 2013).

In Vietnam, rice production is the key economic sector with approximately 45 million tons per year (FAO 2018). The Mekong Delta is one of the most important rice producing areas of Vietnam and has transformed the country from suffering a rice deficit into a huge rice surplus economy (Demont and Rutsaert 2017). Indeed, Vietnam has become one of the top rice exporters (Berg 2001; Berg and Tam 2012), although 40–50% of the costs of exportable rice are associated with imported fertilizers and agro-chemicals (Demont and Rutsaert 2017).

In Europe, 80% of rice production takes place in Italy and Spain, the former being the largest rice producer. Rice cultivation in Italy is mostly located in Northern regions (Piemonte, Lombardia, and Veneto), with about 250,000 ha and a rough rice production of 1.5 million tons (Kunova et al. 2014).

2 Rice Blast Disease: Pathogen, Geographic Distribution and Economic Importance

Rice can be attacked by many pests and diseases which cause huge losses annually worldwide. In particular, rice productivity is severely affected by a disease

Fig. 1 *Pyricularia oryzae* distribution map. *Source* https://www.plantwise.org/knowledgebank/dat asheet/46103#; accessed on September 3, 2019

caused by the hemibiotrophic fungal pathogen *Pyricularia oryzae* Cavara (anamorph of *Magnapothe oryzae* B. *Couch* sp. nov.) (Wilson and Talbot 2009; Chauhan et al. 2017). *P. oryzae* is a haploid heterothallic fungus, belonging to Ascomycetes, with a single mating type gene having two alleles (MAT1-1 and MAT1-2). The pathogen requires both opposite mating types for sexual reproduction to occur, but limited studies have provided evidence of sexual reproduction occurring under field conditions (Pagliaccia et al. 2018).

The disease caused by *P. oryzae*, named "brusone" in Italy in 1828, is currently known as "blast", first described by Metcalf after a strong *P. oryzae* infection that affected rice in the United States in 1906 (Nunes Maciel 2011).

The disease is very common in temperate rice-growing regions (Yan and Talbot 2016), being distributed worldwide in more than 85 countries (Fig. 1) of South Asia, Europe and America (Martin-Urdiroz et al. 2016) including Italy (Titone et al. 2015), Vietnam and United States (Talbot 2003; Wilson and Talbot 2009) and reduces annual total rice production by 10–30% (Talbot 2003). These losses exceed 70 billion US dollars and would be enough to feed more than 60 million people (Miah et al. 2013; Nalley et al. 2016). Regional epidemics can be devastating; if susceptible cultivars are grown, environmental factors favour disease development and adequate management measures are not adopted, then losses in productivity can range from 60 to 100% (Chauhan et al. 2017), thus impacting upon global food security (Skamnioti and Gurr 2009). Indeed, *P. oryzae* is considered scientifically and economically the most important fungal plant pathogen (Dean et al. 2012).

Under climate change conditions, an increased impact of *P. oryzae* due to lengthening of the growing season and warmer temperatures is expected. By 2030, a homogeneous increase in the number of infection events is predicted throughout Europe (+5 to 20%) and mostly in Italy (even above 20%; Bregaglio et al. 2013).

3   Infection Cycle and Symptoms

Usually the infection process occurs overnight and can be severe during periods of warm temperatures (25–30 °C) and relative air humidity values (above 85–89%). The presence of wet organs for 4–10 h is sufficient for the onset of the disease (Bregaglio et al. 2013). The fungus overwinters in rice straw and spreads rapidly by airborne spores. Upon landing on leaf surface, the three-cell conidia bind to the leaf surface using special adhesive released from the apical compartment of the spores (Hamer et al. 1988) and within 2 h produce a germination tube in a water droplet (Fig. 2a).

By recognizing the hydrophobic and hard surface of the rice leaves, in less than 3 h the short germ tube differentiates at the tip by forming a specialized infection structure called appressorium (Wilson and Talbot 2009; Kuroki et al. 2017). This dome-shaped structure generates enormous cellular turgor pressure (up to 8 MPa or 80 bars) by accumulating concentrations up to 3 M of glycerol (de Jong et al. 1997; Foster et al. 2017; Zhou et al. 2017) and this mechanical force is able to break the rice cuticle allowing invasion of the underlying epidermal cells (Howard et al. 1991; de Jong et al. 1997; Wilson and Talbot 2009). Key factors favoring the accumulation of this extremely high turgor pressure are a differentiated cell wall rich in chitin and the deposition of a thick layer of melanin on the inner side of the appressorium cell wall, adjacent to the plasma membrane. Melanin creates a barrier that facilitates the accumulation of glycerol and the subsequent generation of the high turgor pressure (Chumley and Valent 1990; Money and Howard 1996; Martin-Urdiroz et al. 2016).

Exploiting this mechanical force, a narrow penetration hypha emerges from the base of the appressorium, which has a thinner cell wall and lacks melanin, breaks the leaf cuticle and grows inter- and intracellularly in the epidermal cells, initially maintaining the integrity of the plant plasma membrane.
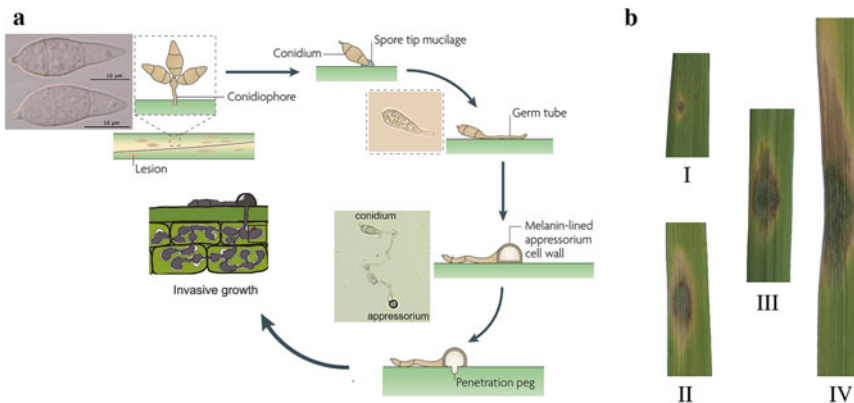
Fig. 2   a Infection cycle of *Pyricularia oryzae* adapted from Wilson and Talbot (2009). **b** Development (from I to IV) of typical eye-shaped blast symptoms on rice leaf surface

The cellular processes required for appressorium formation and function, as well as the recent progress in understanding the biology of host penetration and colonisation of rice tissue are described in detail by Martin-Urdiroz et al. (2016).

At 96 h after the initial infection, the first visible symptoms become apparent. *P. oryzae* can infect all above-ground parts of the rice plant, including leaves, nodes and panicles (Wilson and Talbot 2009). In particular, infections on leaves cause typical leaf spot, eye-shaped symptoms characterized by large, spreading lesions with a necrotic center and a chlorotic margin (Fig. 2b; Wilson and Talbot 2009). Leaf infections reduce the photosynthetic efficiency of plants.

When rice blast infects rice seedlings, the plant often dies. In mature plants, the fungus can infect stems, nodes or the emerging rice panicle resulting in the nearly complete loss of grains (Talbot 2003). In particular, nodal blast can cause necrosis and frequently the breakage of the panicles (Chauhan et al. 2017). In Italy, environmental conditions for rapid growth of the pathogen mainly occur between the second half of July and the first half of August. The most severe outbreaks are observed starting from inflorescence stage and cause neck blast symptoms (Titone et al. 2015).

Infections at the reproductive stage (flowering and mature panicles) cause the contamination of rice seeds. In this case, planting infected seeds results in poor germination and abnormal seedlings (Faivre-Rampant et al. 2013).

Under high humidity conditions, the appearance of necrotic lesions is accompanied by the development of aerial conidiophores releasing asexual conidia that are spread by wind and/or rain splash and serve as a secondary inoculum allowing further spreading of the disease (Wilson and Talbot 2009).

4 Pathogenetic Mechanisms of *Pyricularia oryzae*

The pathogenic mechanisms of *P. oryzae* have been extensively studied. Genes differentially expressed during appressorium formation have been identified (Soanes et al. 2012). In particular, *P. oryzae* activates dynamic changes of its cell wall components during both appressorium formation and infection. The fungal cell wall is largely composed of branched β-1,3 and β-1,6 glucans cross-linked to chitin. Fungal glucan and chitin synthases, chitinases and glucanases may play important roles by maintaining cell wall plasticity (Kong et al. 2012; Samalova et al. 2017; Nguyen and Nguyen 2017). Indeed, the differentiation of the *P. oryzae* cell wall structure during appressorium formation and infection process has been shown to be mediated by the activity of different chitin synthases and β-1,3-glucanosyltransferases and appears a key factor to produce the required mechanical force to break the plant cuticle (Odenbach et al. 2009; Kong et al. 2012; Samalova et al. 2017).

Besides this, pathogenic fungi have evolved strategies to avoid host recognition, protecting themselves during infection against host innate immunity. For instance, the accumulation of α-1,3-glucan on the surface of the *P. oryzae* cell wall could mask β-1,3-glucan and/or chitin, thus protecting the fungal cell wall and overcoming plant defense mechanisms (Mochizuki et al. 2011; Fujikawa et al. 2012).

P. oryzae is considered a hemibiotrophic fungus (Talbot 2003). After a significant biotrophic phase at the early stages of infection, in which the fungus grows within living rice cells surrounding the invaginated plant plasma membrane, and depends on them for nutrition, the fungus reprograms its transcription and switches to a necrotrophic phase, killing the plant cells and leading to visible lesions (Ribot et al. 2008). At this stage, *P. oryzae* expresses many genes encoding secreted proteins, mainly of unknown function. Indeed, *P. oryzae* possesses a complex secretome (Dean et al. 2005), with 739 proteins predicted to be secreted and 163 putatively secreted proteins occurring in families. Several of these families are predicted to encode enzymes possibly involved in the degradation of the plant cell wall and cuticle (Dean et al. 2005). Many of these genes are significantly up-regulated during the necrotrophic stage, suggesting that the fungus degrades the plant cell walls or remodels its own cell wall. In particular, the simultaneous silencing of *P. oryzae* endo-β-1,4 xylanases or cellulases belonging to the glycoside hydrolase families 6 and 7 has shown that these cell wall degrading enzymes clearly contribute to fungal virulence (Vu et al. 2012; Nguyen et al. 2011).

Additionally, *P. oryzae* secretes in the plant apoplast and cytoplasm several effectors that remain largely uncharacterized in terms of their biological role (Martin-Urdiroz et al. 2016).

Although rapid progress has recently been made in the characterization of *P. oryzae* pathogenetic mechanism as well as in the identification of fungal virulence factors, many questions remain unsolved, including the mechanism inducing the necrotrophic switch.

5 Disease Management Strategies

Given the importance of rice as a staple food source and the yield losses caused by *P. oryzae*, the availability of methods to control rice blast disease is crucial to ensure global food security. However, the control is expensive and extremely difficult (Talbot 2003; Nalley et al. 2016).

Integrated Pest Management (IPM) control strategies, including chemical, cultural, genetic and biological methods (Table 1), aim at improving rice farming systems, producing higher yield in a sustainable way.

5.1 Disease Forecasting Models

Disease forecasting allows researchers to simulate the potential risk of infections or to forecast the evolution of disease epidemics, providing information about whether, when and where to apply a particular disease management practice, such as a fungicide treatment.

Table 1 Currently applied measures to control Rice Blast Disease in the field

Chemical	De methylation inhibitors—Triazoles (e.g. tebuconazole, propiconazole and difenoconazole) Quinone outside inhibitors (e.g. azoxystrobin, trifloxystrobin and picoxystrobin) Melanin biosynthesis inhibitors—Reductase (e.g. tricyclazole)
Cultural	Redce Nitrogen fartilization Silicon fertilization (e.g. calcium or magnesium silicate, silicagel) Prevent water stress Incorporation of remains in the soil Wider spacing of plants Early planting
Genetic	Use of resistant cultivars (pyramided resistance genes)
Biological	*Bacillus subtilis* (QST 713 stain)

Disease forecasting is potentially one of the most effective and economical strategies against blast disease. To date, 52 studies have been published: the majority of these systems are able to predict only leaf blast and are mostly based on input variables such as air temperature, relative humidity and rainfall (Katsantonis et al. 2017). Other critical factors for the onset of the disease, such as leaf wetness, nitrogen fertilization and variety resistance, have been integrated only in a few models (Katsantonis et al. 2017) but should be taken into consideration in future developments of rice blast prediction systems.

However, few rice blast forecasting models are currently in use for rice growers that can be used for many years or in different geographical regions. Most of the available blast disease models are calibrated and evaluated in Asian rice cultivation areas. One of the four models currently in use is available in Europe as a module implemented in the EU service "Monitoring Agricultural ResourceS" (MARS), operated by the Joint Research Center at Ispra in Italy (Katsantonis et al. 2017). A dynamic deterministic model (SiRBInt) simulating the rice-blast interaction and including both weather dependent-crop and pathogen growth patterns is currently available in Italy after a ten-year validation (Biloni et al. 2006).

The major limitations of the disease forecasting approach are the requirement of powerful computers and advanced networks and servers with extensive database capabilities, and the need to validate the models making them more reliable, by correcting current inaccuracies and uncertainties in the predictions.

5.2 Chemical Control

Chemical control based on the use of fungicides is the most common and reliable management strategy, especially once blast symptoms are evident in the field (Kunova et al. 2014; Chen et al. 2015; Pak et al. 2017). Chemicals commonly used to control rice blast are systemic fungicides with different site-specific modes of

action and include tebuconazole, propiconazole and difenoconazole (sterol demethylation inhibitor; Pak et al. 2017); azoxystrobin, trifloxystrobin and picoxystrobin (mitochondrial respiration inhibitors, derived from naturally occurring strobilurins, belonging to the quinone outside inhibitor group; Kunova et al. 2012; Chen et al. 2015); tricyclazole (Kunova et al. 2014; Pak et al. 2017). The latter inhibits melanin biosynthesis in the appressorium (Woloshuk et al. 1983), a key structure for fungal penetration in the host tissue, and is one of the most used fungicides, being the most effective in controlling the disease (Kumar et al. 2013; Katsantonis et al. 2017). In particular, tricyclazole shows high efficacy by inhibiting sporulation and secondary infections (Kunova et al. 2012), a single spray being effective in reducing disease incidence (Titone et al. 2015). Besides, field trials have also demonstrated the high efficacy of mixtures containing difenoconazole and azoxystrobin as well as tebuconazole and trifloxystrobin (Usman et al. 2009).

Although in poor countries farmers can not meet the expense of pesticides, in many Asian countries such as Vietnam the fungicide use on rice has increased drastically in the past decades (Berg and Tam 2012). In recent years, blast has also caused severe losses in Italy, thus forcing farmers to apply fungicides on more than 75% of the rice cultivated area (Food Chain Evaluation Consortium 2011).

However, the increased and continuous reliance on fungicides has proved to be unsustainable and cost-inefficient, having negative impacts on environment and health of users and consumers and also increasing the risk of appearance of resistant strains, favored by the strong selective pressure. Indeed, to reduce the development of resistant strains of the pathogen, the use of site-specific fungicides is often recommended in mixture or alternation.

Recently, information campaigns have resulted in a substantial reduction in pesticide use by modifying Vietnamese farmers' attitude and management strategies (Berg and Tam 2012).

Fungicide restrictions such as the gradual withdrawal of synthetic formulations constitute further limits on the chemical approach. Indeed, European legislation on plant protection products is imposing more stringent rules for their approval (EC1107/2009) and sustainable use (Directive128/2009). In particular, tricyclazole has been banned from use in European and also Italian rice cultivations, thus limiting the number of active substances still authorized on rice (Kunova et al. 2014).

5.3   Cultural Control

Blast disease occurrence and development are significantly affected by nutrients, in particular Nitrogen (N) and Silicon (Si).

Severe blast disease is often associated with excessive use of N fertilizers, which in turn promotes higher usage of fungicides (Berg and Tam 2012), as recently observed in Vietnam (Demont and Rutsaert 2017). Therefore, highly active fertilizers are not recommended in conventional rice cropping due to their breakdown effects on field blast resistance (Freitas et al. 2010). Besides, plants fertilized with high amount

of N have fewer silicated epidermal cells and thus have lower resistance (Pooja and Katoch 2014). Indeed, as shown in field plot experiments in Vietnam, Si enhances rice plant physiological strength, improving defense against abiotic and biotic stresses. However, Si availability can differ greatly in different paddy soils. Therefore, fertilization with calcium or magnesium silicate or with silica gel (SiO_2) in paddy fields with low available Si might be an option to increase rice yields and mitigate blast disease (Klotzbücher et al. 2018). In particular, calcium silicate as foliar application enhances rice resistance to *P. oryzae* infection through accumulation and fortification of Si in the epidermal cell wall and increased lignin content in leaf (Ng et al. 2019).

Rice grown under upland conditions is more susceptible to blast disease than when grown in flooded soil. Besides, under upland conditions, plants subjected to drought stress are more susceptible to blast disease. For this reason, it is highly recommended to prevent water stress conditions by ensuring irrigation uniformity and maintaining a layer of standing water in the field (Nunes Maciel 2011; Chauhan et al. 2017).

Planting rice in sandy or muddy soils is not recommended because these soils create favorable conditions for the onset of the disease. Since the fungus overwinters in straw and other rice residues (Kunova et al. 2014), incorporation of remains of rice into the soil by deep ploughing is also an effective measure to reduce the source of inoculum in the field (Ogoshi et al. 2018).

Planting time and sowing density also affect the occurrence of blast disease; early planting and a wider spacing of plants are suggested in order to escape or reduce infections. The management of weeds and collateral plant hosts also contribute to further reduce the severity of blast disease in the field (Chauhan et al. 2017).

5.4 *Genetic Resistance*

Genetic plant resistance is the most economical and environmentally friendly disease management strategy. Both qualitative and quantitative resistance genes have been reported in the rice germplasm (Miah et al. 2013; Srivastava et al. 2017). Qualitative resistance is governed by major gene(s) and leads to complete resistance, while quantitative resistance is governed by many minor genes, leading to partial resistance. Usually, rice resistance to *P. oryzae* is dominant and controlled by one or few major resistance (R) genes. To date, more than 100 R genes (*Pi* genes) and 350 quantitative trait loci (QTLs) have been identified from diverse germplasm. In particular, 22 major R genes, mainly encoding for nucleotide binding - leucine rich repeat (NB-LRR) proteins, have been successfully cloned and characterized (Ballini et al. 2008; Zhai et al. 2011; Srivastava et al. 2017). Although many of these R genes have been transferred using conventional breeding methods to cultivated rice varieties (Korinsak et al. 2011; Miah et al. 2013), the presence of many races/strains of the pathogen in the field and the high genetic variability of the fungal population allows *P. oryzae* to rapidly adapt by mutating and evolve new and more virulent races able to overcome genetic resistance, especially if based on major R genes. Indeed,

genetically resistant rice varieties become ineffective within 2/3 growing seasons (Urso et al. 2016; Pagliaccia et al. 2018).

Currently, among Italian rice varieties, 15% show resistance to the disease and 17% moderate resistance under field conditions (Titone et al. 2015), with serious economic consequences.

On the whole, the best way for controlling the disease is to incorporate both qualitative and quantitative genes in resistant varieties. In particular, plant-breeding and marker-assisted selection strategies that integrate in cultivated crop varieties several major resistance genes by pyramiding could represent an effective approach to produce broad spectrum and durable resistant rice cultivars (Urso et al. 2016). However, considerable effort is needed to develop a pyramided rice cultivar, and the resistance phenotype could be associated with a yield penalty.

5.5 Biological Control

Recently, consumers have become more aware of the negative effects of pesticide use. In fact, the use of fungicide has harmful effects on the environment and puts food safety at risk through the possible presence of chemical residues in grains.

Given the need of reducing reliance on synthetic pesticides to protect human health and the environment and the short-time effectiveness of genetic resistance, it becomes crucial to investigate and develop new effective, durable and sustainable alternative methods for rice blast disease management.

Recent studies have been therefore focused on identifying biocontrol agents able to control *P. oryzae* growth and development (Table 2). In particular, several beneficial soil microorganisms have been tested.

Treatments with antagonistic *Bacillus* spp. enhance the antioxidant defense activities in shoots and roots of rice in response to *P. oryzae* infection, thus alleviating the induced oxidative damage and suppressing blast disease incidence (Rais et al. 2017).

Application of culture filtrates of a strain of *Bacillus methylotrophicus*, producing phenaminomethylacetic acid, 24 h before inoculum with *P. oryzae* reduced blast symptoms by about 85% in both greenhouse and field trials (Shan et al. 2013).

Pre-treatments of rice plants with a strain of *Streptomyces* spp. have been able to reduce by about 65% the progression of blast disease severity compared with other treatments. In this case, the potential biocontrol effect of *Streptomyces* spp. depends not only on its ability to inhibit *P. oryzae* but also on its capacity to promote plant growth and resistance (Awlaa et al. 2017).

Also, treatments of rice seeds with a strain of the endophytic bacterium *Achromobacter xylosoxidans* have been shown to significantly increase the activity of defense related enzymes and thus to reduce blast disease incidence by about 40% (Joe et al. 2012).

To date, only one product based on the biocontrol agent *Bacillus subtilis* (strain QST 713) is registered for rice treatment in Italy against blast disease. In particular,

Table 2 Biological control agents recently tested against Rice Blast Disease and their mode of action

Bacteria	*Bacillus* spp. *Streptomyces* spp. *Achromobacter xylosoxidans*	Triggering of rice defenses Production rice growth and resistance Promoting rice growth and resistance Activation of rice defense enzymes	Rais et al. (2017) Shan et al. (2013) Awlaa et al. (2017) Joe et al. (2012)
Yeast	Several species	Delay fungal spore germination by reducing appressorium formation	Kunyosying et al. (2018)
Plant extracts	Saponins *Angelica gigas* *Chrorrolaena odorata*	Alteration of fungal membrane Leading to cell rupture Inhibition of spore germination reduction of foliar and seed infections	Abbruscato et al. (2014) Yoon et al. (2011) Khoa et al. (2011)
Nano-particles	*Chaetomium lucknowense*	Inhibition of spore germination	Song et al. (2018)

this treatment is allowed also in organic farming from the panicle formation stage to the harvest.

A first evidence that yeasts can also reduce rice blast symptoms has been recently obtained (Kunyosying et al. 2018). Several antagonistic yeast isolates have been found to delay spore germination and significantly reduce appressorium formation. Greenhouse and field experiments have confirmed the ability of these yeast isolates to significantly reduce disease incidence to about 15% when compared to 80% of disease incidence of the control treatment.

Additional investigations have assessed the fungicidal activity of plant extracts, including essential oils, phenolic compounds, coumarin derivatives and other metabolites such as saponins.

A mixture of saponins (triterpenoid glycosides, able to alter fungal membrane permeability leading to cell rupture) extracted from alfalfa (*Medicago sativa*) has shown promising antifungal activity against *P. oryzae* when used to treat different Italian rice cultivars (Abbruscato et al. 2014).

Crude root extracts of *Angelica gigas*, containing antifungal coumarin compounds such as decursin and decursinol angelate, suppress the development of rice blast disease by strongly inhibiting *P. oryzae* spore germination (Yoon et al. 2011). Additionally, aqueous extracts of a wild herbal Vietnamese plant, *Chromolaena odorata*, are able to reduce by 45% blast symptoms by different application methods (foliar spray and seed treatment) in semi-field conditions in Vietnam (Khoa et al. 2011). Nano-particles derived from *Chaetomium lucknowense* extracts are able to antagonize *P. oryzae*, reducing blast symptoms by about 50% when compared to the fungicide tricyclazole (Song et al. 2018).

6 Novel Approaches for Blast Management

Despite all the control measures currently applied in the field by farmers, blast disease is still an important constraint to rice production and its durable control still represents a substantial challenge.

Therefore, the development of innovative strategies for the identification of new molecules for the sustainable protection of rice are highly desirable.

One of the major limits to the use of antagonistic biocontrol agents against crop pathogens in open field is the strong influence of different environmental conditions on their effectiveness (Boland 1997). Additionally, the efficacy of biological control based on the use of plant extracts with antimicrobial activity is often affected by the significant variability usually observed in their composition and formulation (Avasthi et al. 2010).

Since the beneficial modes of action of antagonistic microorganisms against plant pathogens, such as antibiosis and the induction of systemic resistance, are known to be mostly activated by their secondary metabolites, an innovative approach is the use of the pure active substances responsible for their beneficial action. With this approach, the health hazards (recently highlighted also by EFSA, 2016) connected with the use of living fungi as biological control agents (for instance, *Trichoderma* spp. are well-known human opportunistic species and new pathogens of plants and even humans could possibly evolve due to horizontal gene transfer among microorganisms; Harman et al. 2004; Koonin et al. 2001) and the spread of useless products in the environment would be drastically reduced.

Although rapid progress has been recently made in the identification of virulence factors of *P. oryzae*, genetic studies able to better define the key steps for attachment, penetration and invasive growth of *P. oryzae* on rice tissues are still necessary and could provide new methods to control blast disease. In particular, a more integrated understanding of the mechanism responsible for appressorium formation, for instance by characterizing the role of genes actively and differentially expressed during this stage (Soanes et al. 2012), is certainly required and could contribute to identifying key enzymes essential for pathogenesis on rice. This knowledge could represent a useful approach in order to identify new targets for developing antifungal chemicals or inhibitor molecules able to effectively inhibit the activity of the essential fungal enzymes, thus reducing growth and infection rate of *P. oryzae*.

References

Abbruscato P, Tosi S, Crispino L, Biazzi E, Menin B, Picco AM, Pecetti L, Avato P, Tava A (2014) Triterpenoid glycosides from *Medicago sativa* as antifungal agents against *Pyricularia oryzae*. J Agric Food Chem 62:11030–11036

Avasthi S, Gautam AK, Bhadauria R (2010) Antifungal activity of plant products against *Aspergillus niger*: A potential application in the control of a spoilage fungus. Biol For Int J 2(1):53–55

Awla HK, Kadira J, Othmanc R, Rashida TS, Hamide S, Wonga M-Y (2017) Plant growth-promoting abilities and biocontrol efficacy of *Streptomyces* sp. UPMRS4 against *Pyricularia oryzae*. Biol Control 112:55–63

Ballini E, Morel JB, Droc G, Price A, Courtois B, Notteghem JL, Tharreau D (2008) A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides new insights into partial and complete resistance. Mol Plant-Microbe Interact 21:859–868

Berg H (2001) Pesticide use in rice and rice-fish farms in the Mekong Delta, Vietnam. Crop Prot 20:897–905

Berg H, Tam NT (2012) Use of pesticides and attitude to pest management strategies among rice and rice-fish farmers in the Mekong Vietnam. Int J Pest Manage 58:153–164

Biloni M, Rodolfi M, Picco AM (2006) SiRBInt, a new simulation model to forecast rice blast disease. Ital J Agrometeorol 3:58–62

Boland GJ (1997) Stability analysis for evaluating the influence of environment on chemical and biological control of white mold (*Sclerotinia sclerotiorum*) of bean. Biol Control 9:7–14

Bregaglio S, Donatelli M, Confalonieri R (2013) Fungal infections of rice, wheat, and grape in Europe in 2030–2050. Agron Sustain Dev 33:767–776. https://doi.org/10.1007/s13593-013-0149-6

Chauhan BS, Jabran K, Mahajan G (2017) Rice production worldwide. Springer, Switzerland

Chen Y, Yang X, Yuan SK, Li YF, Zhang AF, Yao JC, Gao TC (2015) Effect of azoxystrobin and kresoxim-methyl on rice blast and rice grain yield in China. Ann Appl Biol 166:434–443

Chumley FG, Valent B (1990) Genetic analysis of melanin-deficient, nonpathogenic mutants of *Magnaporthe grisea*. Mol Plant-Microbe Interact 3:135–143

de Jong JC, McCormack BJ, Smirnoff N, Talbot NJ (1997) Glycerol generates turgor in rice blast. Nature 389:244–244

Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeyer C, Li W, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan EJ, Birren BW (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434:980–986

Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD (2012) The top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol 13:414–430

Demont M, Rutsaert P (2017) Restructuring the vietnamese rice sector: towards increasing sustainability. Sustainability 9:325. https://doi.org/10.3390/su9020325

EFSA Panel on Biological Hazards (2016) Scientific opinion on the risks for public health related to the presence of *Bacillus* spp. in foodstuffs. EFSA J 14:4524

Faivre-Rampant O, Genies L, Piffaneli P, Thareau D (2013) Transmission of rice blast from seeds to adult plants in a nonsystemic way. Plant Pathol 62:879–887

FAO (2009) FAO's director-general on how to feed the world in 2050. Insights Expert Meet FAO 1:1–35

FAO (2018) GIEWS—global information and early warning system, country briefs: Vietnam. Available online at https://www.fao.org/giews/countrybrief/country.jsp?code=VNM

FAOSTAT (2012) Food and agricultural organization of the United Nations. Available at https://FAOSTAT.fao.org

FAOSTAT (2014) Food and agriculture organization of the United Nations. Available at https://faostat.fao.org

FAOSTAT (2016) Food and agriculture organization of the United Nations. Available at https://fao stat.fao.org

Food Chain Evaluation Consortium (2011) Economic damage caused by lack of plant protection products against rice blast (*Pyricularia grisea*) in rice in Italy. Study on the establishment of a European fund for minor uses in the field of plant protection products: Final report, pp 159–164. https://ec.europa.eu/food/plant/protection/evaluation/study_establishment_eu_fund.pdf

Foster AJ, Ryder LS, Kershaw MJ, Talbot NJ (2017) The role of glycerol in the pathogenic lifestyle of the rice blast fungus *Magnaporthe oryzae*. Environ Microbiol 19(3):1008–1016

Freitas JC, Angelimalavolta VM, Salomon MV, Cantarella H, Castro LHSM, Azzini L (2010) Nitrogen fertilization and blast incidence in upland rice. Bragantia 69:173–179

Fujikawa T, Sakaguchi A, Nishizawa Y, Kouzai Y, Minami E, Yano S, Koga H, Meshi T, Nishimura M (2012) Surface alpha-1,3-glucan facilitates fungal stealth infection by interfering with innate immunity in plants. PLoS Pathog 8:e1002882

Gianessi L.P. (2014). Importance of pesticides for growing rice in South and South-East Asia. Crop Life Foundation. International Pesticide Benefit Case Study 108

Hamer JE, Howard RJ, Chumley FG, Valent B (1988) A mechanism for surface attachment in spores of a plant pathogenic fungus. Science 239:288–290

Harman GE, Howell CR, Viterbo A, Chet I, Lorito M (2004) Trichoderma species–opportunistic, avirulent plant symbionts. Nat Rev Microbiol 2(1):43–56

Howard RJ, Ferrari MA, Roach DH, Money NP (1991) Penetration of hard substrates by a fungus employing enormous turgor pressures. Proc Natl Acad Sci USA 88:11281–11284

Joe MM, Islam MR, Karthikeyan B, Bradeepa K, Sivakumaar PK, Sa T (2012) Resistance responses of rice to rice blast fungus after seed treatment with the endophytic *Achromobacter xylosoxidans* AUM54 strains. Crop Prot 42:141–148

Katsantonis D, Kadoglidou K, Dramalis C, Puigdollers P (2017) Rice blast forecasting models and their practical value: a review. Phytopathologia Mediterranea 56(2):187–216

Khoa ND, Thuy PTH, Thuy TTT, Collinge DB, Jørgensen HJL (2011) Disease-reducing effect of *Chromolaena odorata* extract on sheath blight and other rice diseases. Phytopathol 101:231–240

Khush GS (2005) What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol Biol 59:1–6

Klotzbücher A, Klotzbücher T, Jahn R, Xuan LD, Cuong LQ, Chien HV, Hinrichs M, Sann C, Vetterlein D (2018) Effects of Si fertilization on Si in soil solution, Si uptake by rice, and resistance of rice to biotic stresses in Southern Vietnam. Paddy Water Environ 16:243–252. https://doi.org/10.1007/s10333-017-0610-2

Kong LA, Yang J, Li GT, Qi LL, Zhang YJ, Wang CF, Peng YL (2012) Different chitin synthase genes are required for various developmental and plant infection processes in the rice blast fungus *Magnaporthe oryzae*. PLoS Pathog 8:16

Koonin EV, Makarova KS, Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 55:709–742

Korinsak S, Sirithunya P, Meakwatanakarn P, Sarkarung S, Vanavichit A, Toojinda T (2011) Changing allele frequencies associated with specific resistance genes to leaf blast in backcross introgression lines of Khao Dawk Mali 105 developed from a conventional selection program. Field Crops Res 122(1):32–39

Kumar MKP, Gowda DKS, Moudgal R, Kumar NK, Gowda KTP, Vishwanath K (2013) Impact of fungicides on rice production in India. In: Nita M (eds) Fungicides: showcases of integrated plant disease management from around the world. InTech, pp 77–98

Kunova A, Pizzatti C, Cortesi P (2012) Impact of tricyclazole and azoxystrobin on growth, sporulation and secondary infection of the rice blast fungus, *Magnaporthe Oryzae*. Pest Manag Sci 69:278–284. https://doi.org/10.1002/ps.3386

Kunova A, Pizzatti C, Bonaldi M, Cortesi P (2014) Sensitivity of nonexposed and exposed populations of *Magnaporthe oryzae* from rice to tricyclazole and azoxystrobin. Plant Dis 98:512–518

Kunyosying D, To-anun C, Cheewangkoon R (2018) Control of Rice Blast Disease Using Antagonistic Yeasts. Int J Agric Technol 14(1):83–98

Kuroki M, Okauchi K, Yoshida S, Ohno Y, Murata S, Nakajima Y, Nozaka A, Tanaka N, Nakajima M, Taguchi H, Saitoh KI, Teraoka T, Narukawa M, Kamakura T (2017) Chitin-deacetylase activity induces appressorium differentiation in the rice blast fungus *Magnaporthe oryzae*. Sci Rep 7:9697. https://doi.org/10.1038/s41598-017-10322-0

Martin-Urdiroz M, Oses-Ruiz M, Ryder LS, Talbot NJ (2016) Investigating the biology of plant infection by the rice blast fungus *Magnaporthe oryzae*. Fungal Genet Biol 90:61–68

Miah G, Rafii MY, Ismail MR, Puteh AB, Rahim HA, Asfaliza R, Latif MA (2013) Blast resistance in rice: a review of conventional breeding to molecular approaches. Mol Biol Rep 40:2369–2388

Mochizuki S, Saitoh K-I, Minami E, Nishizawa Y (2011) Localization of probe accessible chitin and characterization of genes encoding chitin-binding domains during rice–*Magnaporthe oryzae* interactions. J Gen Plant Pathol 77:163–173

Money NP, Howard RJ (1996) Confirmation of a link between fungal pigmentation, turgor pressure, and pathogenicity using a new method of turgor measurement. Fungal Genet Biol 20(3):217–227

Nalley L, Tsiboe F, Durand-Morat A, Shew A, Thoma G (2016) Economic and environmental impact of rice blast pathogen (*Magnaporthe oryzae*) alleviation in the United States. PLoS ONE 11:1–15. https://doi.org/10.1371/journal.pone.0167295

Ng LC, Hafiz ES, Sariam O, Razi RIM (2019) The effect of calcium silicate as foliar application on aerobic rice blast disease development. Eur J Plant Pathol 153:533–543. https://doi.org/10.1007/s10658-018-1580-y

Nguyen BQ, Nguyen NBC (2017) The role of cell wall degrading enzymes in pathogenesis of *Magnaporthe oryzae*. Curr Prot Pept Sci 18:1–16

Nguyen QB, Itoh K, Vu VB, Tosa Y, Nakayashiki H (2011) Simultaneous silencing of endo-β-1,4 xylanase genes reveals their roles in the virulence of *Magnaporthe oryzae*. Mol Microbiol 81:1008–1019

Nunes Maciel JL (2011) *Magnaporthe oryzae*, the blast pathogen: current status and options for its control. CAB Rev Perspect Agric Vet Sci Nutr Nat Resour 6:050

Odenbach D, Thines E, Anke H, Foster AJ (2009) The *Magnaporthe grisea* class VII chitin synthase is required for normal appressorial development and function. Mol Plant Pathol 10:81–94

Ogoshi C, Carlos FS, Ulguim AR, Zanon AJ, Bittencourt CRC, Almeida RD (2018) Effectiveness of fungicides for rice blast control in lowland rice cropped in Brazil. Trop Subtrop Agroecosyst 21:505–511

Pagliaccia D, Urak RZ, Wong F, Douhan LI, Greer CA, Vidalakis G, Douhan GW (2018) Genetic structure of the rice blast pathogen (*Magnaporthe oryzae*) over a decade in North Central California Rice Fields. Microb Ecol 75:310–317. https://doi.org/10.1007/s00248-017-1029-4

Pak D, You MP, Lanoiselet V, Barbetti MJ (2017) Azoxystrobin and propiconazole offer significant potential for rice blast (*Pyricularia oryzae*) management in Australia. Eur. J. Plant Pathol. 148:247–259

Pooja K, Katoch A (2014) Past, present and future of rice blast management. Plant Sci Today 1(3):165–173

Rais A, Jabeen Z, Shair F, Hafeez FY, Hassan MN (2017) *Bacillus* spp., a bio-control agent enhances the activity of antioxidant defense enzymes in rice against *Pyricularia oryzae*. PLoS ONE 12(11):e0187412. https://doi.org/10.1371/journal.pone.0187412

Ribot C, Hirsch J, Balzergue S, Tharreau D, Nottéghem J-L, Lebrun M-H, Morel J-B (2008) Susceptibility of rice to the blast fungus, *Magnaporthe Grisea*. J Plant Physiol 165:114–124

Samalova M, Mélida H, Vilaplana F, Bulone V, Soanes DM, Talbot NJ, Gurr SJ (2017) The β-1,3-glucanosyltransferases (Gels) affect the structure of the rice blast fungal cell wall during appressorium-mediated plant infection. Cell Microbiol 19:e12659. https://doi.org/10.1111/cmi.12659

Shan H, Zhao M, Chen D, Cheng J, Li J, Feng Z, Ma Z, An D (2013) Biocontrol of rice blast by the phenaminomethylacetic acid producer of *Bacillus methylotrophicus* strain BC79. Crop Prot 44:29–37

Skamnioti P, Gurr SJ (2009) Against the grain: safeguarding rice from rice blast disease. Trends Biotechnol 27(3):141–150. https://doi.org/10.1016/j.tibtech.2008.12.002 (PMID: 19187990)

Soanes DM, Chakrabarti A, Paszkiewicz KH, Dawe AL, Talbot NJ (2012) Genome-wide transcriptional profiling of appressorium development by the rice blast fungus. PLoS Pathog 8:e1002514

Song JJ, Soytong K, Kanokmedhakul S, Kanokmedhakul K (2018) Nano-particles from *Chaetomium lucknowense* to inhibit rice blast pathogen caused by *Pyricularia oryzae* in pot experiment. Int J Agric Technol 14(7):1961–1968

Srivastava D, Shamim M, Kumar M, Mishra A, Pandey P, Kumar D, Yadav P, Siddiqui MH, Singh KN (2017) Current status of conventional and molecular interventions for blast resistance in rice. Rice Sci 24(6):299–321

Talbot NJ (2003) On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. Annu Rev Microbiol 57:177–202

Titone P, Mongiano G, Tamborini L (2015) Resistance to neck blast caused by *Pyricularia oryzae* in Italian rice cultivars. Eur J Plant Pathol 142:49–59

Urso S, Desiderio F, Biselli C, Bagnaresi P, Crispino L, Piffanelli P, Abbruscato P, Assenza F, Guarnieri G, Cattivelli L, Valè G (2016) Genetic analysis of durable resistance to *Magnaporthe oryzae* in the rice accession Gigante Vercelli identified two blast resistance loci. Mol Genet Genomics 291:17–32. https://doi.org/10.1007/s00438-015-1085-8

Usman GM, Wakil W, Sahi ST, Saleemil Y (2009) Influence of various fungicides on the management of rice blast disease. Mycopathol 7(1):29–34

Vu BV, Itoh K, Nguyen QB, Tosa Y, Nakayashiki H (2012) Cellulases belonging to glycoside hydrolase families 6 and 7 contribute to the virulence of *Magnaporthe oryzae*. Mol Plant-Microbe Interact 25(9):1135–1141

Wilson RA, Talbot NJ (2009) Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. Nat Rev Microbiol 7:185–195

Woloshuk C, Sisler H, Vigil E (1983) Action of the antipenetrant, tricyclazole, on appressoria of *Pyricularia oryzae*. Physiol Plant Pathol 22:245–259

Yan X, Talbot NJ (2016) Investigating the cell biology of plant infection by the rice blast fungus *Magnaporthe oryzae*. Curr Opin Microbiol 34:147–153

Yoon MY, Kim YS, Ryu SY, Choi GJ, Choi YH, Yang KS, Cha B, Han SS, Kim JC (2011) In vitro and in vivo antifungal activities of decursin and decursinol angelate isolated from *Angelica gigas* against *Magnaporthe oryzae*, the causal agent of rice blast. Pestic Biochem Physiol 101:118–124

Zhai C, Lin F, Dong Z, He X, Yuan B, Zeng X et al (2011) The isolation and characterization of Pik, a rice blast resistance gene which emerged after rice domestication. New Phytol 189:321–334

Zhou T, Dagdas YF, Zhu X, Zheng S, Chen L, Cartwright Z et al (2017) The glycogen synthase kinase MoGsk1, regulated by Mps1 MAP kinase, is required for fungal development and patho

Collaborative RTD for Precise and Reliable GNSS Based Positioning for Land Management

Tung Hai Ta, Vinh The La, Hiep Van Hoang, Matteo Vannucchi, Gabriella Povero, and Gustavo Belforte

Abstract Global Navigation Satellite Systems (GNSS) easily allow precise determination of position, velocity and time (PVT) enabling hundreds of different applications in almost any field of modern life, from transportation to logistics, from surveying to disaster management, from natural resources monitoring to services for citizens. In this field, the collaboration involving the NavSAS group at Politecnico di Torino and Fondazione LINKS (former ISMB) together with the NAVIS group at HUST is lasting since more than 10 years. Besides focusing on basic research aspects that provide knowledge and skills, which are essential for achieving excellence and expertise, a considerable effort has been devoted to the development of technological solutions that allow precise and reliable PVT determination, which is crucial in land management and environmental applications. In particular, a low-cost prototype GNSS receiver named NavisA has been developed. The equipment can ensure an accuracy in the range of some centimeters and is therefore suited for land management applications. The equipment was tested in different conditions by independent investigators that verified its performances confirming that it can contribute to the modernization of the land management sector in Vietnam, where traditional time-consuming total station measurements are still widely used.

T. H. Ta · V. T. La · H. Van Hoang
NAVIS Center, Hanoi University of Science and Technology, Hanoi, Vietnam

M. Vannucchi · G. Povero (✉)
Fondazione LINKS, Turin, Italy
e-mail: gabriella.povero@linksfoundation.com

G. Belforte
Politecnico di Torino, Turin, Italy

1   Introduction

International collaboration has been widely promoted at Politecnico di Torino since long time. In a first stage, it was mainly directed towards European and North American institutions. However, with the general progress of culture and science in a globalizing world, at the end of last century, collaborations with institutions of other regions became more frequent. This change was also facilitated by the European Union that started funding educational and RTD collaborative projects involving institutions based in countries that were not traditional partners of previous RTD activities. It is in this framework that, in the late nighties, just at the time in which Vietnam was restoring its links with western economies, Politecnico di Torino started its collaboration with the Hanoi University of Science and Technology (HUST) in the framework of different EU funded projects.

Almost in the same period in which the first connections between Politecnico di Torino and HUST were setup, the European Union (EU) decided to develop its own Global Navigation Satellite System (GNSS) Galileo. Several reasons suggested this move that had a very important practical impact on RTD in Europe, since it became urgent to increase the European expertise in GNSS technology to support the considerable effort needed to setup the new system. Important funding was allocated to promote the growth in Europe of this field of study and technology development. To support the setup of the Galileo System and facilitate its future adoption by potential users, the EU supported also several awareness raising actions aimed at providing better knowledge of the new system under construction, of its features and advantages. Preliminary studies and testing of new applications were also promoted to prepare market opportunities for the time in which Galileo would have been fully operational. Europe has obviously been the first and more important recipient for these actions; however, considering the global nature of the foreseen Galileo system and the European positive attitude for international cooperation, considerable energies and resources were also devoted to international collaboration.

In the first years 2000, at Politecnico di Torino and at Istituto Superiore Mario Boella (ISMB), a research institution strictly related with Politecnico that became operational at that time, a very active joint research group (NavSAS) focusing on GNSS technology and related applications was setup and started working actively in research and education. Leveraging the existing links with HUST and the European support for international collaboration, in 2004 the first collaborative project in the field of GNSS was setup. It involved ISMB, Politecnico di Torino and HUST together with Universitat Politecnica de Catalunya, and Southeast University in Nanjing, China.

In following years, the collaboration in the field of GNSS between Politecnico di Torino and ISMB on the Italian side, and HUST on the Vietnamese side, strengthened and grew quite considerably. With the support of the SEAGAL Project, an initiative funded by the EU in the FP7 framework, the NAVIS Centre was setup at HUST and started its operations in October 2010. Its mission was, and is, to act as linking entity between Europe and South-East Asia (SEA) in the field of Global Navigation Satellite

Systems, promoting European technology by reinforcing cross-links between EU and SEA actors and facilitating international collaboration among players.

From 2010 onwards, the collaboration between the NavSAS Group in Torino and the NAVIS Centre in Hanoi has continuously developed and increased, also thanks to different EU funded projects as well as to activities funded by the Italian and Vietnamese Governments in the framework of the Executive Programme of Scientific and Technological Cooperation between the Italian Republic and the Socialist Republic of Vietnam. A timeline reporting the main research projects in which the NavSAS Group and the NAVIS Centre have been involved is reported in Fig. 1, where also the main steps in the setup of the Galileo System are reported. While developing its activities, the NAVIS Centre has setup important research links with outstanding institutions such as the European Space Agency (ESA), the Institute for the Protection and Security of the Citizen of the European Commission Joint Research Centre (JRC-IPSC), the Japanese Space Agency (JAXA), the Australian Centre for Space Engineering Research (ACSER) and has reached some important technical achievements like being one of the first 50 users in the world of the Galileo system, as certified by ESA.

Besides being engaged in research projects, the NAVIS staff has always devoted keen attention to technology transfer actions that can contribute to the growth of the Vietnamese society and to its technological development. Among them an important activity has been the development of a low cost GNSS receiver which anyway provides good accuracy and can be customized to fit the needs of local Vietnamese users.
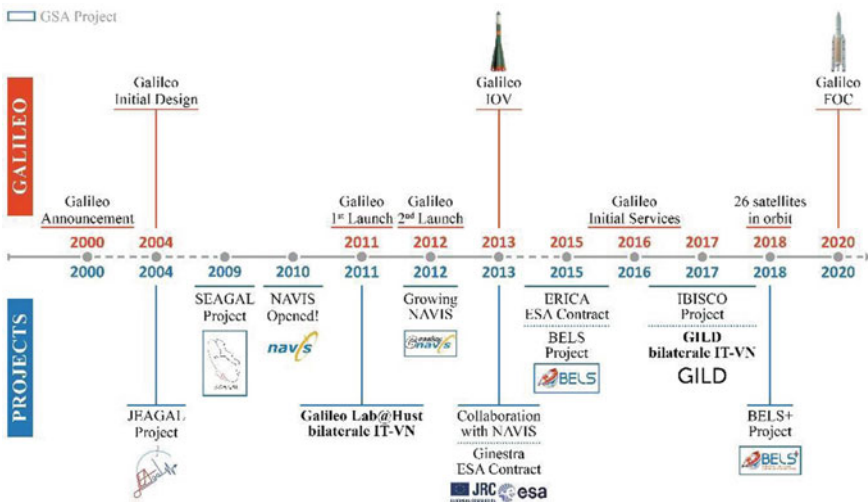

Fig. 1 Main joint research activities linked to the major steps in the development of the Galileo System

Different prototypes of this receiver, named NavisA, have been developed in past years always improving its functionalities. The target applications for this equipment are those in the field of surveying and mapping as well as land monitoring in which centimeter level accuracy is commonly required and large amount of measurements are often needed. The NavisA has been also used for setting up a small research-oriented CORS network before the Department of Surveying, Mapping and Geo-Information of Viet Nam started deploying the Vietnamese CORS network that is now operational. With this equipment and with its related applications, the NAVIS staff won the first prize at the Vietnam Talent Award 2015.

In this paper, the NavisA GNSS receiver is presented together with the research oriented NaviNet CORS network that was setup for research purposes at the NAVIS Centre. The outcomes of some tests performed on the NavisA receiver for testing its performances are reported as well.

2   Background

In Vietnam, surveying, mapping and land monitoring are still conducted to a large extent through traditional methods (total stations and trilateration). With this approach, the points where measurements are performed need to be in view from each other. This is indeed a quite strict requirement that can increase the complexity of a survey measurement. On top of this, considerable time is needed to collect measurements and to elaborate them. When instead GNSS technology is used, often this is done with static receivers that need to stay for long time in each measurement point in order to ensure acceptable accuracy. This solution is obviously time consuming and not efficient.

However, GNSS technology nowadays provides better solutions based on the use of receivers which can provide good accuracy also in dynamic conditions in which the receiver is moved around almost continuously. Surveying is obviously much faster and more efficiently conducted when using this kind of receivers. However, while there are several receivers of this kind on the market, their price is quite high for the Vietnamese contest, and this fact has been preventing their wide adoption.

Noting these facts and recognizing the interest in the Vietnamese society for the development of technological solutions in line with the demands of the local society, the NAVIS Centre started to work on a project for the development of a low cost GNSS receiver providing centimeter level accuracy using up-to-date technological solutions but standard components available on the market. This activity has also been partly supported by some Vietnamese national research funding.

With GNSS technology, the position on earth of a GNSS receiver is obtained first computing its distance from different GNSS satellites, whose position in sky is known, and then deriving its position by trilateration. The final accuracy and precision of the computed position depends on whether the distance from satellites is derived using only the code transmitted to the receiver or also the phase of the carrier signal. Furthermore, it is affected by different error components related to the satellite orbits,

the clocks on board the satellites, the effect of the ionosphere and of the troposphere on the signals traveling through them.

To achieve high positioning accuracy, it is required on the one side to use carrier phase measurements, on the other side to reduce to the largest possible extent the measurement errors (Subirana et al. 2013).

Two main approaches have been developed in GNSS technology to perform this last task: Real Time Kinematics (RTK) and Precise Point Positioning (PPP).

2.1 Real Time Kinematics

Real Time Kinematics (RTK) is the most commonly used precise positioning method providing decimeter (or even centimeter) level precision in real-time. Besides using carrier phase measurements, in this approach most of the error components in the distance determination between receiver and satellites are eliminated implementing a differential measurement strategy. It consists in using two receivers: one in a fixed position (the so-called base station) while the second (the so-called rover) is moving around to take measurements in different points. Assuming that the errors in the determination of the distance between satellites and receivers are almost the same for the rover and the base station (which is true if the distance between the two does not exceed 20 km), the differential measurement allows deleting most of the error components providing a very accurate determination of the relative position between rover and base station. If the position of the base station is precisely known, which can be the case if the base station is in a known reference point or if it has been in position for a sufficiently long time allowing a very accurate determination of its position, then also the absolute position of the rover is very accurately known.

Indeed, the main disadvantage of RTK is the need of a reference station closed by. The requirement can be overcome if a CORS network of reference stations is available. This infrastructure consists in a network of reference stations interconnected through NTRIP servers to a NTRIP Caster that can broadcast corrections through the internet to any connected rover operating within the area covered by the RTK network. Indeed, the cost of such infrastructure is related to its extension, and to the number of connected reference stations but in general is quite large. Interested readers can refer to (Mekik and Arslanoglu 2009) and (Gakstatter 2013) for more details on RTK.

2.2 Precise Point Positioning

Precise Point Positioning (PPP) is a different approach to precise position determination. Using dual frequency receivers, the important error component connected with the ionospheric delay, which is frequency dependent, can be removed using a linear

combination of measurements at the two frequencies. In addition, accurate corrections provided through the internet are used for orbits and clocks. These corrections are derived with different models and solutions by the International GNSS Service (IGS) which makes use of the data collected by a global network of receivers covering the whole world. Since 2010, IGS is running a pilot project in which registered users can access free of charge its real-time correction data. Interested readers can refer to Grinter and Roberts (2013) and Rizos et al. (2012) for more details on PPP.

3   The NavisA Receiver

At the NAVIS Centre keen attention has been devoted in past years to the development of a low cost GNSS receiver, the NavisA, that provides good accuracy and can be customized to the needs of the Vietnamese users. Different prototype versions of it have been developed, constantly improving the characteristics and the features of the equipment. All the versions have been designed using off-the-shelf components and chipsets. At the beginning in 2013 a very simple single-frequency receiver was built as training prototype used for familiarizing with all practical problems related to the construction of this kind of equipment. Following versions were dual-frequency receivers in which both technical performances and friendliness of use were constantly improved. With these receivers both the RTK and the PPP solutions have been implemented. The latest version of this receiver is shown in Fig. 2 and is described in more detail hereafter. It should be noted that in this version the RTK solution is implemented relying on the corrections provided by the CORS network recently setup in Vietnam by the Department of Survey and Mapping while the PPP solution is derived using IGS NTRIP streams available at the following link https://

Fig. 2   NavisA presented to Prime Minister of Vietnam Mr. Nguyen Tan Dung

Fig. 3 GNSS receiver module of NavisA: it can be connected to an android tablet through Bluetooth or with a cable to a PC

www.igs.org/rts/products. However, the receiver can also work together with a base station that has also been developed at the NAVIS Centre and was commonly used with previous prototype versions.

The latest NavisA is split in two parts and the one with the GNSS receiver (see Fig. 3) is quite compact, light with limited power consumption. It is implemented with the dual-frequency Hemisphere (P306) chipset integrated with a Bluetooth transceiver module that ensures the connection with an android-based tablet to which it provides the raw pseudorange and phase measurement data. On the tablet, all the software programs that implement RTK or PPP solutions are allocated and run. These programs elaborate the raw pseudorange measurements together with relevant information retrieved through internet using the GPRS connection of the tablet. When implementing the RTK solution, the information retrieved through internet typically consists in the corrections received from the CORS network setup and managed by the Department of Survey and Mapping and Geo-Information of Viet Nam. Of course, it is possible to use also other networks, or the data provided by a base station. When implementing the PPP solution, IGS corrections are retrieved from internet. The Graphic User Interface (GUI) presents all relevant information on the tablet screen in a user-friendly way. Obviously, still an antenna must be connected to the receiver. With this solution, the whole equipment results to be of quite easy use and turns out to be very portable.

To test the performances of this equipment, different pilot experiments have been conducted and are described in Sect. 5.

4   NaviNet: A Small CORS Network of NavisA Receivers

In 2015, when the Department of Surveying, Mapping and Geo-Information of Viet Nam had not yet deployed its CORS network, the NAVIS Centre has been running for some months a small research-oriented CORS network, named NaviNet, covering a limited area in the Hanoi city. The stations of this network were equipped with NavisA receiver modules (like the one shown in Fig. 3) each one connected to a PC and through internet to the server at the NAVIS Centre. The locations of the stations together with some information about each one of them are shown on in Fig. 4.

The server at the NAVIS Centre, besides storing the data from all the stations, made them available to registered users for research purposes. An experimental service providing post processing corrections was also operational for some time.

Since the CORS network managed by the Department of Surveying, Mapping and Geo-Information of Viet Nam has become operational, this network was dismounted and only data from the receiver at the NAVIS Centre can still be accessed by registered users for research purposes.

The experience gained with this experimental CORS network has been very valuable. It has supported research and demonstration activities allowing also to show to different authorities the advantages of this kind of solution. It has also facilitated some demonstration activities conducted in the framework of the European funded project BELS (Povero, et al. 2015).

Fig. 4 NaviNet, a small-scale network of NavisA receivers to provide differential correction for precise positioning

5 Experiments and Pilot Testing of NavisA Receiver

Several experiments and tests were conducted to assess the performances of the NavisA receiver and to compare them with those provided by traditional methods employing total stations, the conventional equipment still widely used for this kind of task in Vietnam. This comparison had also the scope to allow potential users of the NavisA receiver to verify that its performance is as good as that of traditional methods but much less effort and time is required to acquire measurements and results.

5.1 Experiment at the Hanoi University of Science and Technology

A first test to evaluate the performances of the NavisA was conducted in the premises of the Hanoi University of Science and Technology where the NAVIS Centre is located. A pattern, to be followed by the receiver antenna, was drawn on the floor of the court of the university and then it was carefully followed with the NavisA antenna.

The positions provided by the NavisA were reported on a grid with centimeter-resolution and are shown in Fig. 5 both for the RTK solution (green points) and for the PPP solution (orange points).

From this experiment it can be noticed that in both cases a centimeter-level precision is continuously obtained.
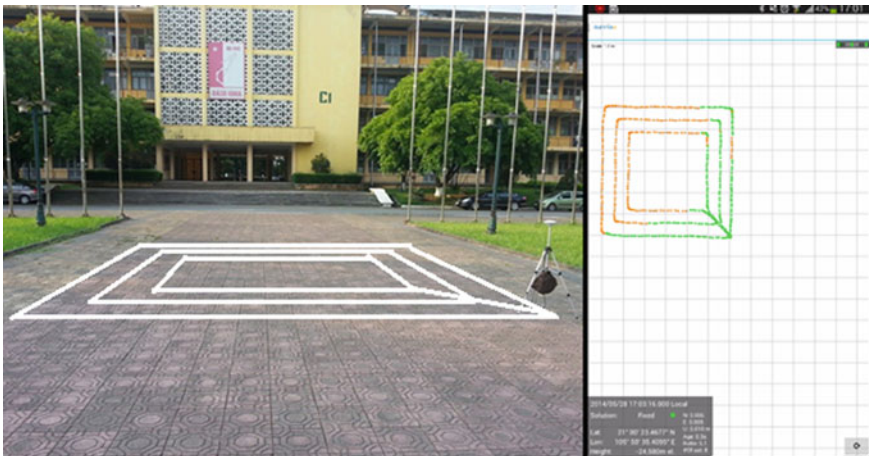
Fig. 5 Experiment for testing the NavisA receiver at HUST campus

5.2   Surveying Angular and Distance Framework Measurement

In this experiment, a set of seven points placed on the perimeter of a construction site had to be accurately located to be used as reference points for the construction surveying. The position of each of these points was determined with the NavisA receiver as well as with the traditional total station method. The two methods provided almost the same outputs. The differences in the distances between adjacent points evaluated with the two methods were all under 1.5 cm.

5.3   Measurements of Earth Crust Movements

The NavisA receiver has been used by officers of the Surveying, Mapping and Geo-Information Department to test whether it could be used for monitoring slow land movements like the ones occurring in landslides (Quang 2018). In this study, investigators evaluated the height retrieved with the NavisA receiver in a set of points for which the official height was also available from the database of the Surveying, Mapping and Geo-Information Department. One of the points was then chosen as reference and the differences in height between all the tested points and the reference one have been computed using the NavisA measurements as well as using the official data. The difference in the obtained results with the two data sets was typically in the range of 1 cm and in any case not larger than 2.5 cm. Thus, the surveyors, considering usual practices in Vietnam for landslide monitoring, found the results satisfactory and concluded that the NavisA could be used for monitoring landslide movements.

5.4   Topographical Survey for Construction Site

Another test was autonomously performed by a construction company that borrowed the NavisA equipment and tested it. In the test, 3D positions of some test points were derived with the NavisA and with the traditional total station method. The results of their comparison were provided to the NAVIS and are reported in Table 1.

It can be noted that the differences between the measurements performed with the two methods are rather limited, in the range of some centimeters, with maximum values of 9.8 cm for X direction, 5.1 cm for Y direction and 5.0 cm for the height. It should be noted that the measurements were obtained with two independent methods, thus their individual errors can add up when computing their difference.

Being satisfied with the performances of the NavisA, the construction company used it together with a laser scanner in order to derive a high-resolution topographic mapping to be used in official documents related to the construction of a highway near Chu Lai airport in Quang Nam Province, Vietnam.

Table 1 Point measurement with NavisA and total station equipment

Point	Difference NavisA-Total Station		
	X(m)	Y(m)	H(m)
DC5-11	0.000	0.000	0.000
DC5-12	-0.052	0.051	0.050
DH1	-0.098	-0.042	0.011
DH2	-0.057	-0.012	0.018
DH3	-0.064	-0.010	0.007
DH4	-0.064	-0.016	0.022
DH5	-0.058	0.006	0.016
DH6	-0.069	-0.020	0.024
DH7	-0.065	-0.021	0.004
DH8	-0.083	-0.003	0.014
DH9	-0.007	-0.004	0.010
DH10	-0.048	-0.001	0.024
DH11	-0.067	-0.003	0.022
DH12	-0.040	-0.012	0.017

6 Conclusions

In this work, some of the RTD activities conducted at the NAVIS Centre in collaboration with the NavSAS Group of Politecnico di Torino and Fondazione LINKS (former ISMB) have been presented. Specifically, the GNSS based solutions oriented to fulfill needs of the Vietnamese society in the field of surveying, mapping and land monitoring have been presented. The main product produced at the NAVIS Centre for these activities is the NavisA GNSS receiver, a prototype that presents good performance as proved with several experiments but has the advantage of being less expensive than other receivers available on the market. On top of this, the equipment can be easily customized to the needs of Vietnamese users.

References

Gakstatter E (2013) RTK GNSS receivers: a flooded market? GPS World, 21 Mar 2013
Grinter T, Roberts C (2013) Real time precise point positioning: are we there yet? In: Proceeding of the international global navigation satellite systems society, 16–18 July 2013

Mekik C, Arslanoglu M (2009) Investigation on accuracies of real time kinematic GPS for GIS applications. Remote Sens 2009(1):22–35

Povero G et al (2015) Building links between Europe and South-East Asia in the field of EGNSS: the BELS project and the Navis Centre. In: 21st Ka and broadband communications conference, Bologna, Italy, 12–14 Oct 2015

Quang VT (2018) Application of CORS network in subsidence monitoring and adjustment of height values periodically. In: Proceedings of the 2018 Vietnamese national conference on surveying and mapping science and technology, pp 2–8

Rizos C, Janssen V, Roberts C, Grinter T (2012) PPP vs DGNSS. Geomatics World

Subirana JS, Zornora JM, Hernandez M (2013) GNSS data processing. ESA Communications. ISBN: 9292218867

Energy for Vietnam

Sustainable Energy Supply in Vietnam

Eleonora Riva Sanseverino, Ninh Quang Nguyen, Gaetano Zizzo, and Rossano Musca

Abstract The steep growth of electricity demand in Vietnam (11% per year on average since 2006) calls for a deep rethinking of the Vietnamese electrical power system. Its topological structure indeed may not be well suited for a deep penetration of renewable energy sources and a predictable increase of reactive industrial loads. The Vietnamese power system is described and a few analytical methods are considered for its study both in steady state and in dynamic conditions in different scenarios. The power system simulation software Neplan is used for the implementation of the system's model and the performing of the dynamic analysis.

1 Introduction

The Vietnamese power system is rapidly changing, following the steep growth of electricity demand and the current worldwide need for clean energy. The need for the system is to expand, in order to host more capacity and deliver more energy, and at the same time to integrate an always bigger share of renewable energy sources (RES), in order to reduce pollution and increase production from clean energies. The geographical conditions of the Country play a key role in both requirements, reinforcement and RES integration: very long 500-kV transmission lines connect North and South, with a significant amount of power flowing through the Center. The generated power is then concentrated in specific sites where power plants are built, and it comes from different energy sources (Fig. 1): hydropower accounts for around 41% of total installed capacity in 2018, while thermal generation covers mostly the rest of the country's generation and it runs primarily on domestic coal (about 67.8% of total thermal power, concentrated in the North) and natural gas (around 15% of total, mostly installed in the South). Oil covers only the 1% of the generation mix.

E. Riva Sanseverino (✉) · G. Zizzo · R. Musca
University of Palermo, Palermo, Italy
e-mail: eleonora.rivasanseverino@community.unipa.it

N. Q. Nguyen
Institute of Energy Science, Hanoi, Vietnam

Fig. 1 Energy mix in the
Vietnamese power system
for the year 2018 (%)
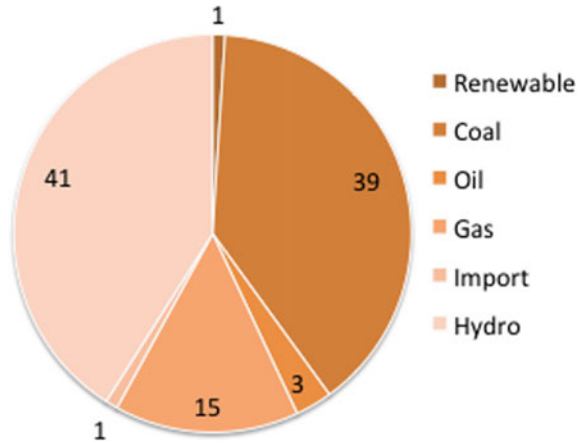
In the North, power is generated by hydro-generators and coal-fired power plants. The center of Vietnam has the lowest electricity demand and consequently also the lowest installed generation capacity. In the South, power is mainly generated by several natural gas power plants systems, installed along the coasts and supplied by offshore deposits. This energy mix has been relatively stable in the last years.

However, the energy mix is expected to radically change in the very next years. If the government plans will be realized, the coal-fired capacity will become 4 times the current capacity, reaching more than 60 GW and thus covering more than 40% of the total demand. The share of hydro-generation capacity will drop to 22%, as it already reached almost its entirely potential. Open- and combined-cycle gas turbines will account for 20% of the overall capacity.

By 2030, the total capacity from renewables is expected to reach 20 GW, with annual rates averaging to 25%. The 7th Power Development Plan announced that about 60% will be solar photovoltaic, 30% wind, and 10% biomass (Gerner et al. 2017).

At the same time, the growth of electricity demand is asking for an expansion of the current power system structure: the interconnections between the three areas will be strengthened by four more 500-kV lines, two of them linking North and Center and bringing transmission capacity to 3.6 GW, and the other two lines linking Center and South and bringing capacity to 6.5 GW. As the transmission network is reinforced, the import of electricity from Lao People's Democratic Republic (PDR) will also increase substantially, with the level of imports through the interconnection reaching up to 3 GW by 2025 and 5 GW by 2030 (Gerner et al. 2017).

The power system of Vietnam is therefore strongly characterized by the peculiar topology of the system, and it is going to experience a steep increase of the share from RES power plants. The illustrated situation arises naturally a series of technical issues, which must be faced and analyzed: one of this is surely the occurrence of power system oscillations between the two areas, North and South, where the generation is concentrated. The appearance of the so called inter-area oscillations at low frequency

is undesired and detrimental for the overall system operation. Moreover, the situation might be likely worsened by the introduction into the system of a large amount of RES power plants, which are typically interfaced to the network via static power converters and do not provide any damping or inertia to the system.

In this document, the phenomenon of inter-area oscillations in the 500-kV system of Vietnam is investigated. A comprehensive analysis is performed both with time-domain and modal analysis, in order to identify the oscillation modes of the system and verify the appearance of inter-area oscillation between North and South. The system will be modeled in detail, and it will be simulated in perturbed condition, assuming the outage of a big power plant. The general problem of inter-area oscilla-tions has been already investigated for several different cases (Xia et al. 2017; Khaji and Aghamohammadi 2017; Pentayya et al. 2014), including also the case of RES integration into the system (Liu et al. 2004). At the moment, there are no studies or papers regarding the problem of inter-area oscillations in the Vietnamese's power system.

2 Description of 500 kV Vietnamese Power System

The Vietnamese power system extends from the North to the South, with very long transmission lines running all along the Country and connecting North and South through the Center. It is interesting to observe that from this point of view this struc-ture is similar to the Italian one (Favuzza et al. 2018). The long 500 kV transmission lines cover in total around 2000 km. As stated in the introduction, the power plants are concentrated in the North and in the South: the total installed capacity is 8622 MW in the North and 5594 MW in the South (Fig. 2).
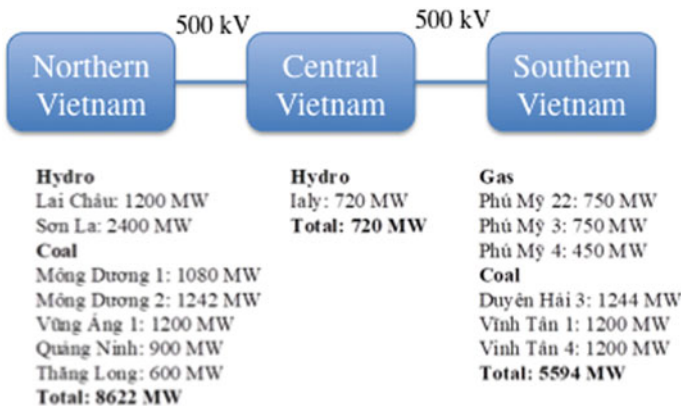
Fig. 2 Total installed capacity of the three areas: North, Center and South
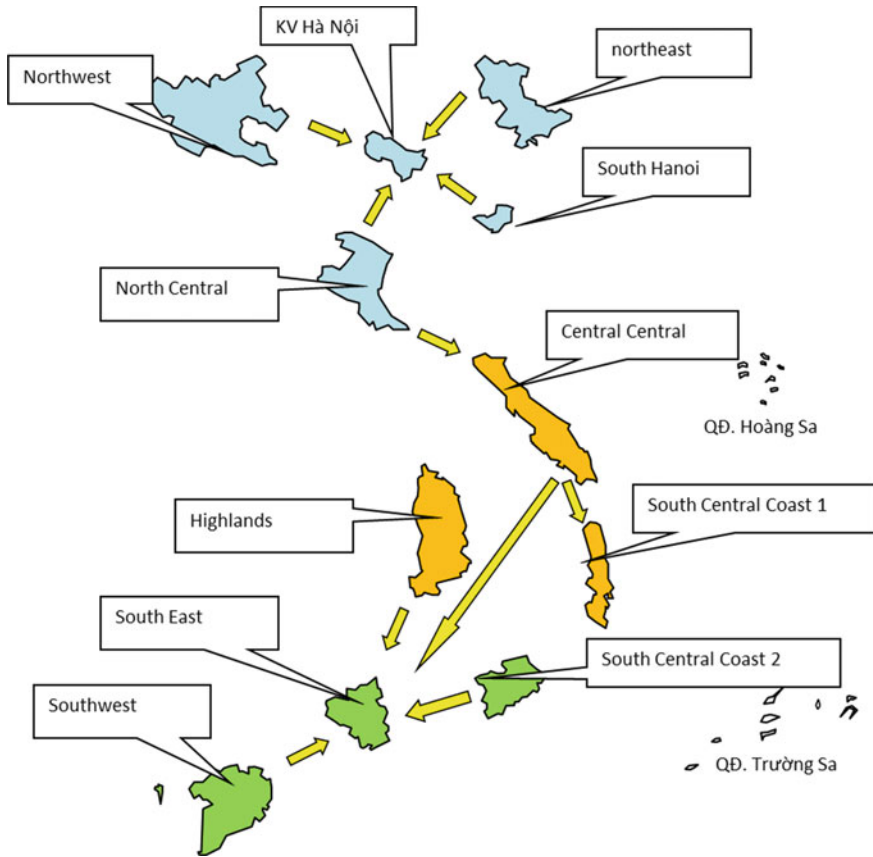
Fig. 3 Inter-regional transmission trend for the period 2015–2030

Considering the distribution of the energy sources within the geographical regions, the expected new power plants and the expected load increase, it is possible to display the diagram of the inter-regional power flows as in Fig. 3. In the figure, the indicated directions of the mainstream capacity are according to the planning period 2015–2030. As it can be seen, the power flows through the long interconnecting corridor composed of 500 kV transmission lines, going from the North area to the South area.

3 Simulation Model

In order to perform the dynamic analysis of the system, a suitable simulation model of the 500-kV Vietnamese power system is developed. The model is intended to be valid both for small signal and transient stability simulations: the combinations of the two analyses allow a complete investigation of the oscillation modes of the

system, with the direct identification of coherent clusters of generators and the check of the inter-area oscillations phenomena. The model is implemented within the power system simulation software Neplan, developed in collaboration with the Institute of Energy Science (IES-VAST) of Vietnam, and validated with the data provided by the National Load Dispatch Centre (NLDC) of Vietnam.

Elements without any dynamic such as lines and transformers are simulated with the corresponding impedances and shunt capacitances. Loads are also implemented as constant impedances, as they are assumed voltage independent. In order to fulfill the load flow constraints, the interconnections with the neighboring countries are simulated with simple equivalents. The synchronous machines of all the power plants are modeled in detail with 6th order sub-transient model: each synchronous machine is equipped with proper controllers for voltage and frequency primary regulation. Secondary regulation is not included as it is not going to act within the considered simulated time (up to 30 s). The parameters of element models are reported in the Appendix.

The simulation model of the system is suitable for both time-domain and modal analysis, and it is based on the Differential Algebraic Equation (DAE) mathematical structure: this allows the formulation of a mathematical model with a set of nonlinear equations. Under the hypothesis of small perturbation on the state variables, the resulting non-linear system can be linearized (Kundur 1994), obtaining the state matrix A required for the calculation of the eigenvalues of the system. This model formulation allows then a comprehensive dynamic analysis of the system.

For the time domain simulations, the system is considered subjected to a sudden generation loss in the Center area: the perturbation represents the outage of one generator and it corresponds approximately to a power unbalance of approximately 450 MW.

For a more convenient visualization of the system topology and power plants displacement, the model is developed directly on the geographical map of Vietnam (Fig. 4), as for the available feature in the software Neplan.

4 Dynamic Analysis

The modal analysis is performed on the linearized model of the system: the eigen-values of the state matrix A are reported graphically in the σ-ω plane of Fig. 5, together with maximum damping ratio $\zeta_{max} = 5\%$ (red line). As it can be clearly observed, the system is stable in the small, as all eigenvalues lie on the left-hand side of the σ-ω plane. The eigenvalues located around the x-axis are known as controller modes, as they express the regulation action of controllers on the system dynamic response. The remaining eigenvalues are responsible of the oscillatory modes, and they can be distinguished according to the associated oscillation frequency as follows (Kundur 1994; Alawasa et al. 2012): in the frequency range 0.7–2 Hz, they are defined as local modes; in the frequency range 0.1–0.7 Hz, they are defined as inter-area modes.
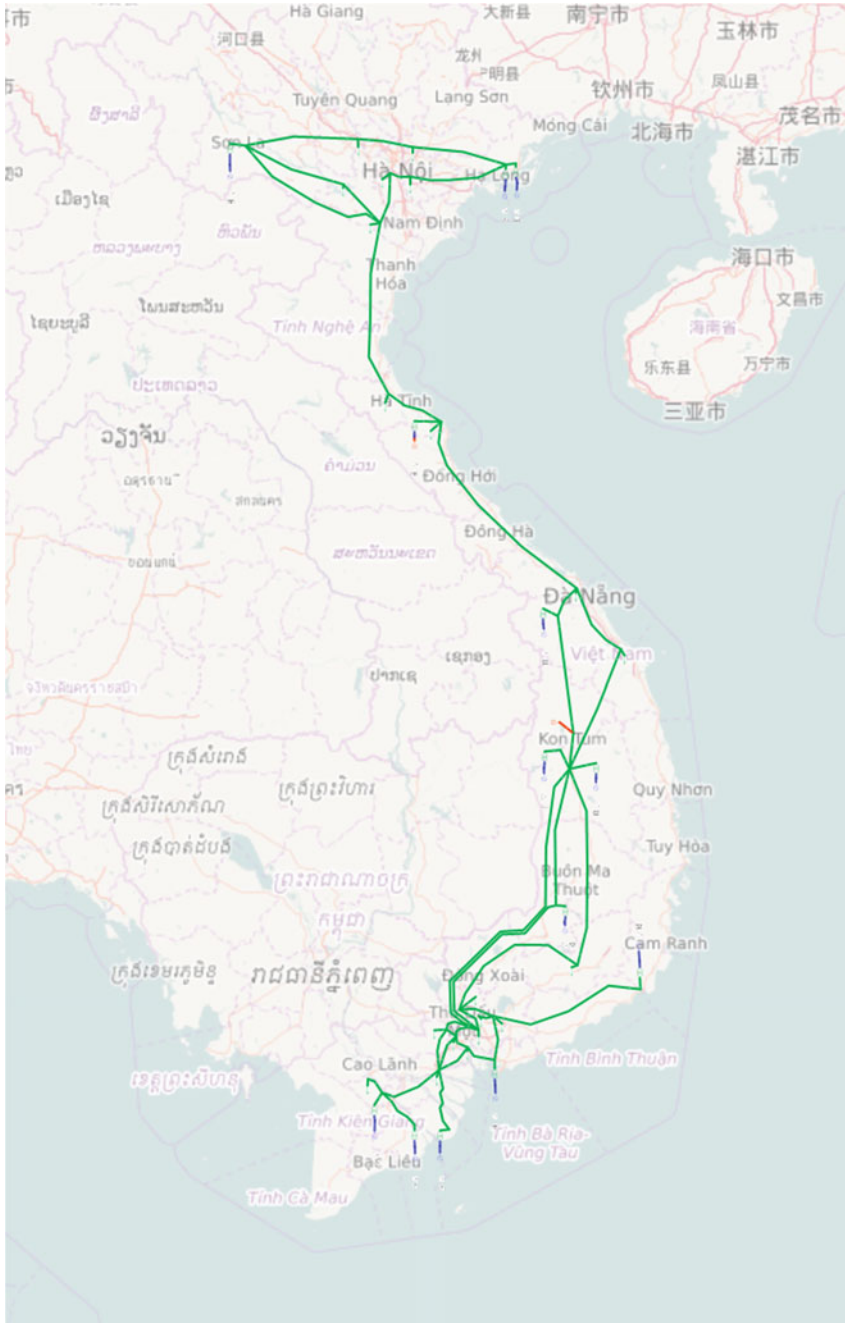
Fig. 4 Simulation model of the 500-kV Vietnamese power system, directly displayed on geographical map in Neplan
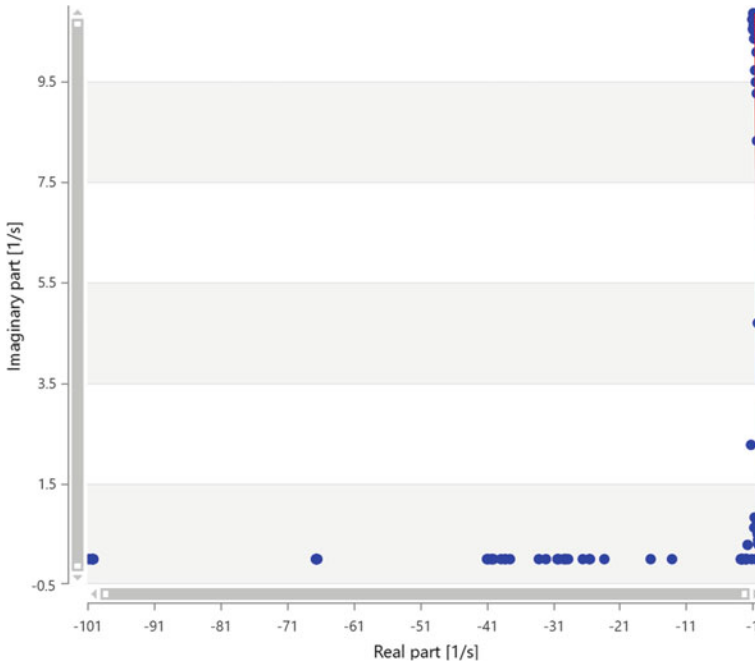
Fig. 5 Eigenvalues of the implemented power system model

With the help of Fig. 5, two inter-area modes are recognized: $\lambda = -0.23 \pm j4.68$ (damping ratio of 0.05, oscillation frequency of 0.74 Hz) and $\lambda = -1.24 \pm j2.27$ (damping ratio of 0.48, oscillation frequency of 0.36 Hz). According to the corresponding participation factors, the mode $\lambda = -0.23 \pm j4.68$ is recognized as the dominant inter-area mode, with a poor damping ratio and an oscillation frequency of 0.74 Hz. This oscillation mode involves the state variables of the dynamic models corresponding to the following power plants: Son La, Mong Duong (North); Ialy (Center); Vinh Tan, Duyen Hai, Phu My (South). This inter-area mode suggests the presence of oscillations at the frequency of 0.75 Hz between the groups of generators in the North and in the South, oscillating across the long 500 kV interconnection lines through the Center.

The result of the modal analysis is then checked with the time-domain analysis of the complete non-linear model of the system. The dynamic simulation gives the results shown in Fig. 6. The generation outage in the Center area produces a significant power unbalance, with the consequent under-frequency transient: the system reacts with primary regulation and it reaches a new steady-state condition within the first 20 s after the disturbance. The secondary regulation is appointed for the elimination of the steady-state frequency error ($\Delta f = 97.3$ MHz). Figure 6 shows the expected classification of two different clusters of synchronous machines, composed of machines oscillating coherently with each other. For a better visualization of the results, a detail of the first seconds after the occurrence of the disturbance is shown
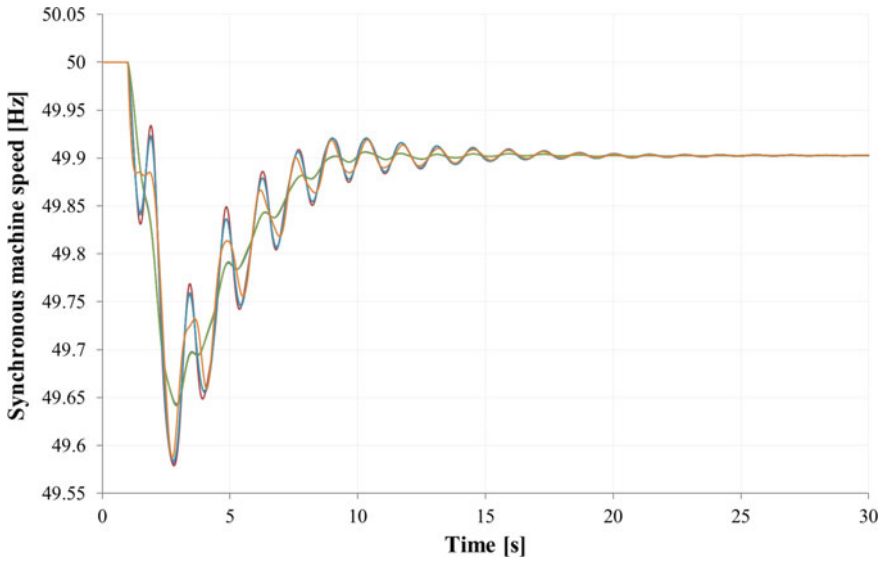
Fig. 6 Speed of synchronous machines in the North, Center and South areas

in Fig. 7. As expected, generators of the North and generators of the South manifest a tight coherency within each area of belonging; generators of the Center (orange curve in Figs. 6 and 7) initially oscillate between the two clusters of North and South,
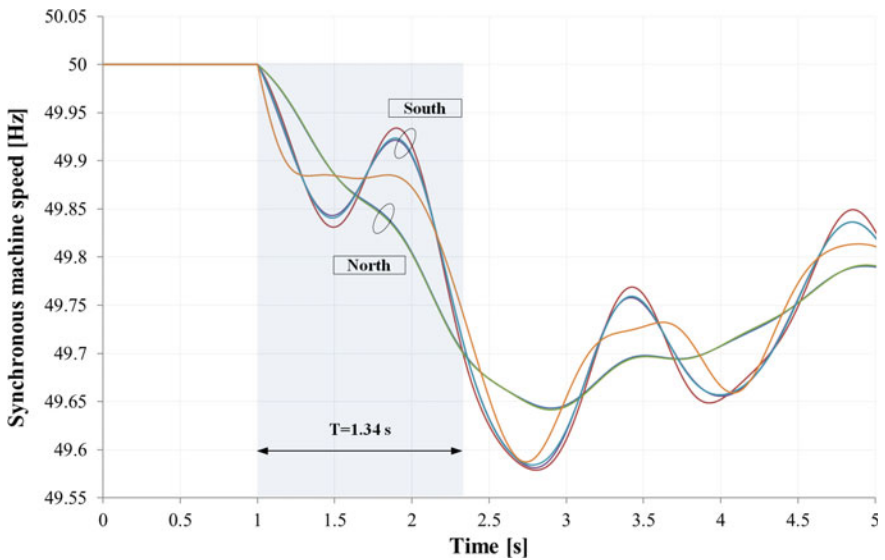
Fig. 7 Inter-area oscillations between North and South in Vietnamese power system

and after few seconds they become coherent to the cluster of the South. The detected coherency in the oscillatory response is conveniently adopted as a criterion for the identification of the generator clusters (Sotiropoulos 2011). The plots of Fig. 7 show the North and South generators swinging against each other, clearly involved in the examined inter-area oscillations. The detected period of these oscillations is T = 1.34 s, so accordingly the frequency is f = 1/T = 0.746 Hz (Fig. 7). The results of the time-domain analysis support then the results obtained in the modal analysis, confirming the appearance of poorly damped oscillations between North and South in the 500 kV Vietnamese power system.

5 Perspectives

The document describes the initial study of the 500 kV power system of Vietnam. The concentration of power plants in the North and in the South area of the country, together with the long interconnection of overhead transmission lines, makes the system inclined to the appearance of the so called inter-area oscillations between the power plants of the two areas. The appearance of these oscillations is an undesired phenomenon and it is detrimental for the operation of the system. This phenomenon is investigated for 500-kV Vietnamese power system through comprehensive dynamic analysis: a detailed model of the system is conveniently developed to perform both small signal and transient stability analysis. The results of the two analyses agree in pointing out the appearance of inter-area oscillations between North and South at the frequency of 0.74 Hz. Participation factors and dynamic results clearly indicate the existence of two clusters of synchronous machines, swinging against each other and belonging respectively to the North and to the South area of the country.

In the perspective of the actions included in the Power Development Plan 7, with the expected installation of a significant amount of RES power plants into the system, the analyzed phenomenon of poorly damped oscillations could be worsened. The expected reinforcement of the power system, with upgrading projects of building new transmission lines and expanding the system capacity, could also complicate the situation, especially considering possible delays in the progress of the projects and the selection of different scenarios.

The appearance of inter-area oscillations in the Vietnamese power system is then an issue which needs to be investigated, especially in the perspective of the expansion of the high voltage transmission network and the installation of the new RES power plants interfaced via power converters to the system.

First simulations of Vietnamese power system with integration of RES have been performed. A power plant of 1 GW has been considered connected to the node "Da Nang" (Center), supplying the power of P = 800 MW. The results are reported in Figs. 8 and 9.

The results show the following:

- the North-South oscillations still occur;
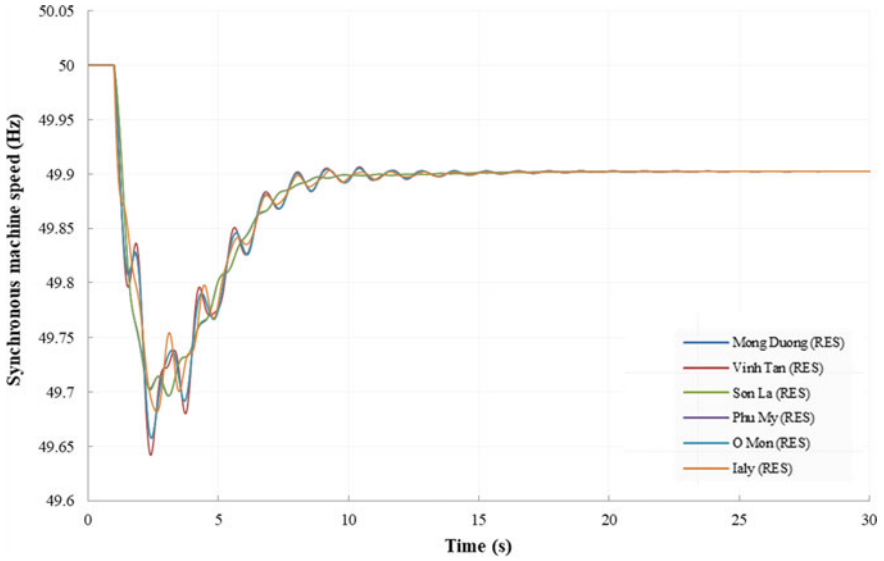
Fig. 8 Speed of synchronous machines in the North, Center and South areas – RES power generation of 800 MW in the Center area
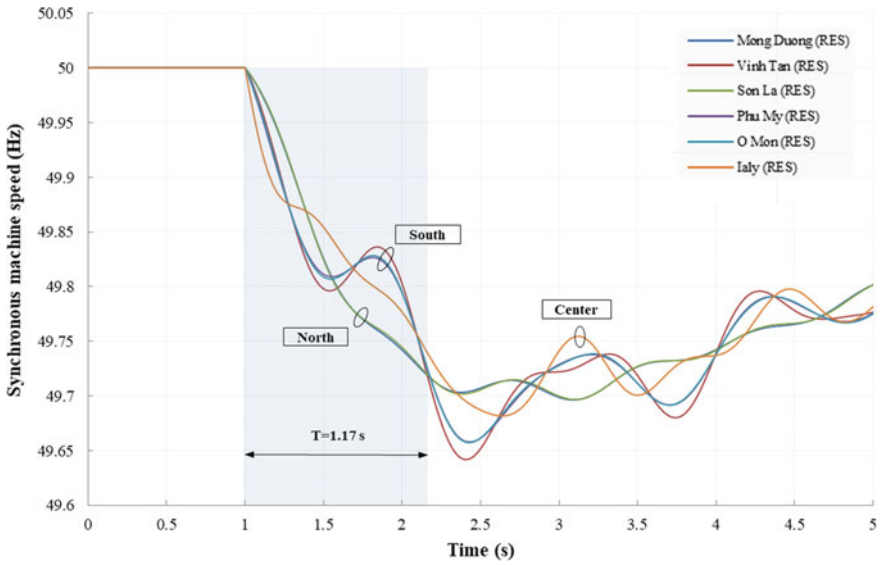
Fig. 9 Inter-area oscillations between North and South in Vietnamese power system – RES power generation of 800 MW in the Center area

- the oscillations have higher frequency and shorter duration;
- the power oscillations North-Center and Center-South are slightly more intense (as expected, due to the additional generation in the Center).

These conclusions are however to be considered as partial, requiring more complete and deeper investigations. The main conclusion is that simulations need to be carried out to assess inter-area oscillations damping and worst-case scenarios for all the planned RES installations in the Vietnamese power system.

Appendix

See Tables 1, 2 and 3.

Table 1 Synchronous machines parameters

Parameter	Value
H (s)	3.5
R (pu)	0.003
Xl (%)	15
Xd (%)	181
Xq (%)	176
Xd$'$ (%)	30
Xq$'$ (%)	65
Xd$''$ (%)	23
Xq$''$ (%)	25
Td0$'$ (s)	8.0
Tq0$'$ (s)	1.0
Td0$''$ (s)	0.03
Tq0$''$ (s)	0.07

Table 2 Exciter simple with limits parameters

Parameter	Value
KA (pu)	200.0
TA (s)	0.015
TB (s)	10.0
TC (s)	1.0
EFMAX (pu)	5.0
EFDMIN (pu)	-5.0

Table 3 Turbine Tgov1 parameters

Parameter	Value
R (pu)	0.05
T1 (s)	0.5
T2 (s)	3.0
T3 (s)	10.0
VMAX (pu)	1.0
VMIN (pu)	0.0

References

Alawasa KM, Mohamed YA-RI, Xu W (2012) Comparative study of the impact of full scale wind turbines on inter-area oscillations. In: 2012 IEEE power and energy society general meeting, San Diego, USA, 22–26 July 2012

Favuzza S, Ippolito MG, Massaro F, Mineo L, Musca R, Zizzo G (2018) New energy corridors in the euro-mediterranean area: the pivotal role of sicily. Energies 11(6):1415

Gerner F, Chattopadhyay D, Bazilian M, Tran KH (2017) Is pumped storage hydroelectric power right for Vietnam? World Bank 2017/75, Washington, DC

Khaji M, Aghamohammadi MR (2017) Emergency transmission line switching to suppress power system inter-area oscillation. Int J Electr Power Energy Syst 87:52–64

Kundur P (1994) Power system stability and control. McGraw Hill, New York

Liu G, Xu Z, Huang Y, Pan W (2004) Analysis of inter-area oscillations in the South China interconnected power system. Electr Power Syst Res 70(1):38–45

Pentayya P, Anumasula R, Patil S, Pandey V, Kumar C, Chitturi S (2014) Spontaneous oscillations and modal resonance in Indian grid: a case study. In: 18th national power systems conference, NPSC 2014, Guwahati, India, 18–20 Dec 2014

Sotiropoulos E (2011) Evaluation of coherency-based aggregation methods. Semester Project Thesis, EEH Power Systems Laboratory, Swiss Federal Institute of Technology (ETH), Zurich, Nov 2011

Xia X, Li C, Ni W (2017) Dominant low-frequency oscillation modes tracking and parameter optimisation of electrical power system using modified Prony method. IET Gener Transm Distrib 11(17):4358–4364

Challenges and Opportunities for Renewable-Based Microgrids Integration in Vietnam

Eleonora Riva Sanseverino, Quynh Thi Tu Tran, Binh Doan Van, Hang Thi Thuy Le, and Ninh Nguyen Quang

Abstract Vietnam is among the South-Asian regions the one that better supplies remote areas. However, many islands and remote areas are still not connected to the main grid and this fact jeopardizes their development, while forcing people to move to urban areas. With a high potential from renewable energy sources and a lot of islands, Vietnam has thus many favorable environmental features for developing the microgrids technology. In this chapter, a detailed analysis about opportunities and challenges for widespread deployment of microgrids technology in Vietnam is considered. Such analysis is based on the assessments of the potential from renewable energy sources in the country and of the national policy framework for the development of renewable energy sources. In the application section, results from the test of improved optimized control strategies to one microgrid system are presented. The results show that such cutting-edge technologies can improve the applicability of the microgrid technology thus reducing its Return on Investment cost.

1 Introduction

Vietnam governmental strategies are focusing on exploration and development of renewable energy sources to supply electricity to consumers. This strategy is also devoted to supply areas where the national grid is not available, such as remote islands and high mountains; these are called communes. The documents (IE 2016; Sinh et al. 2017) indicate the following data about electricity supply to final users: by 2016, the national power grid was covering all 63 provinces/big cities and 713 districts.[1] The rate of communes having access to electricity through the main grid was 98.6%

---

[1] https://www.gso.gov.vn/.

E. Riva Sanseverino (✉) · H. T. T. Le
University of Palermo, Palermo, Italy
e-mail: eleonora.rivasanseverino@community.unipa.it

Q. T. T. Tran · B. D. Van · N. N. Quang
Institute of Energy Science, VAST, Hanoi, Vietnam
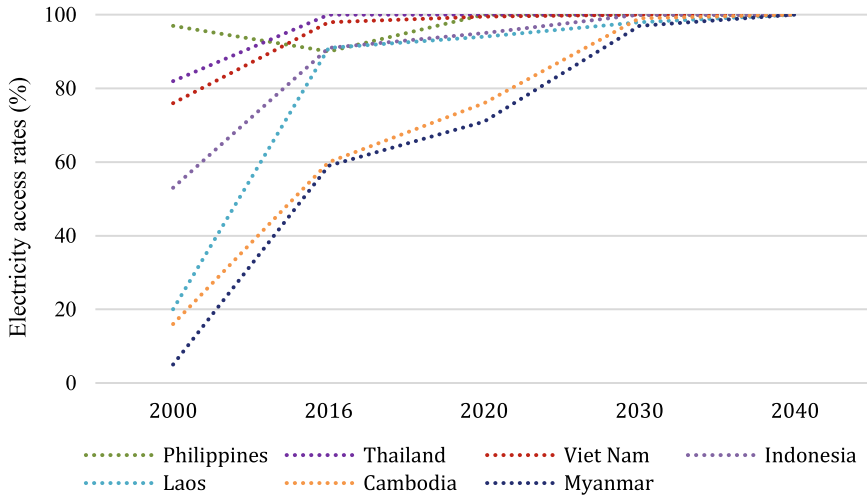
Fig. 1   Rates of Electricity Access in Southeast Asia in fact and under Policies Scenario

and for rural households the rate was was 97.3%. The remaining communes and households were concentrated in remote rural areas and were supplied by local diesel generators, off-grid solar/wind power systems and small hydro power plants (Sinh et al. 2017; Cattelaens et al. 2015).

By the end of 2020, most rural households should have access to electricity (Prime Minister 2016; Prime Minister 2013). The electricity for rural, mountainous and island areas is planned to come from both the national grid and renewable energy sources.

According to the report of IEA (2017), the population that do not have access to grid electricity in Southeast Asian countries is shown in Fig. 1.

From the above Fig. 1, as compared to other countries in the region, Vietnam electrical system supplies almost all population.

However, in Vietnam investments for main grid reinforcement is lagging behind and most of the households in remote areas still depend on conventional fuel burning that is polluting and damaging, both for the environment as well as for humans' health. Generation of electricity through renewable sources such as wind, solar, hydro and even geothermal could emerge as a great alternative in providing clean energy even in remote areas.

With the high potential for renewable energy generation in Vietnam, supplying electricity in remote areas is quite possible, if the development policy on renewable energy power is implemented appropriately. Microgrids are the technology that is considered the most suitable for implementing renewable energy in remote areas with a good power quality. Microgrids are defined as small or medium scale electric distribution networks that include multiple loads, distributed energy resources and storage systems. These energy resources can either be generation units such as fuel
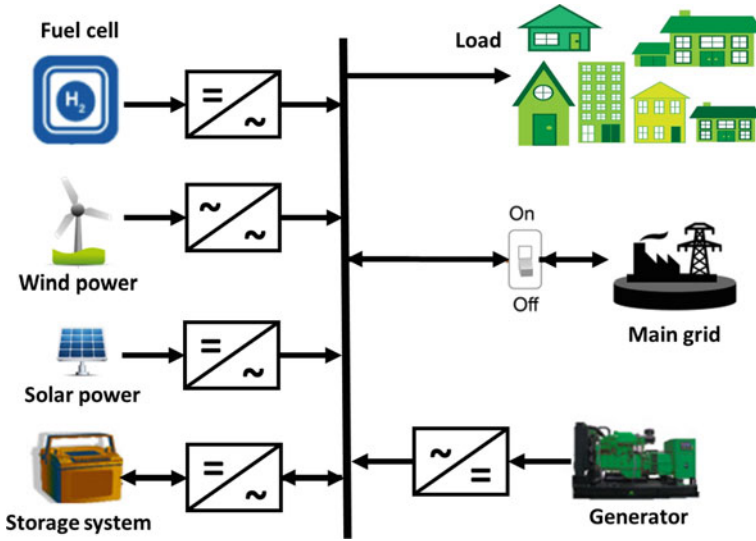
Fig. 2 Sample microgrid system

cells, or renewable energy systems, as well as storage or controllable loads, as shown in Fig. 2 (Soshinskaya et al. 2014).

Therefore, research methods aiming at optimized operation in microgrid systems play an important role to support the widespread adoption of renewable energy in Vietnam as well as to reduce environmental impact, costs and losses operation for Vietnam's final users supply regardless their geographic location.

In this chapter, after a review about the potential of renewable energy in Vietnam, a methodology to reduce power losses for voltage and frequency control in microgrids is described and some results are presented.

2 The Potential of Renewable Energy in Vietnam

Vietnam is one of the Asian countries having the highest renewable energy potential throughout the country. According to (MoIT&DEA 2017; EREA&DEA 2019), the total technical potential of biomass is about 7 GW; biogas is approximately 5.3 GW; municipal waste is around 1.5 GW; small hydro power (<30 MW) is 6.7 GW, solar power is 300 GW, and wind power is 26.7 GW. These renewable energy sources when being used will contribute to meeting the rapidly growing demands of electricity (Bich Huy 2015). Up to 2019, 5 types of technologies have been exploited to produce electricity in Vietnam from renewable energy with a total installed capacity of about 8250 MW, these including: small hydropower (3000 MW), biomass (300 MW),

solar (4500 MW) and wind (450 MW) (EREA&DEA 2019). The current status of renewable energy just obtained about 2.5% of the overall renewable energy potential.

The considered energy installations however are grid connected and mostly at medium and low-voltage. Microgrids technology is another way to exploit renewable energy both in grid connected and in remote areas at low cost.

Barriers on microgrid project development in Vietnam can be listed as follows:

- High investment costs of renewable energy technology solutions;
- Difficult access to financial resources at different stages of project/enterprise development;
- Limited access to tailor-made financing for end-users;
- Legal and institutional framework and uniform legal framework;
- Lack of adequate incentive mechanism;
- Lack of national/province RE development master plan.

2.1 Solar Energy

Vietnam has great potential of solar energy with about 300 GW (MoIT&DEA 2017). The studies (Hoat et al. 2007; Polo et al. 2015) showed that the highest solar potential is concentrated in the South and Highlands. The solar radiation in these regions is the most suitable for photovoltaic installations, reaching an average solar intensity level of 5 kWh/m^2/day.

Currently, the total capacity of grid-connected solar PV power in Vietnam is around 5 GWp, including solar power plants (SPP) and rooftop solar systems (RTS), with about 90 projects of SPP and over 22,000 systems of RTS in operation (EREA&DEA, 2019; VEPG, 2019). The country is preparing for large investments in the solar energy sector, and through the revised Power Development Plan 2011–2020 (PDP7 rev) aims to increase solar power installed capacity up to 12 GW in 2030, the latter accounting for 3.3% of the total electricity production in that year (Prime Minister 2016).

2.2 Wind Energy

Vietnam is considered to have the best wind resources in Southeast Asia. Located in the monsoon climate zone, and shaped by its over 3000 km long coastline, Vietnam´s potential to develop and generate wind power is large. The World Bank recently identified great pote 2017ntial for harnessing wind energy in Vietnam´s south central regions and the Mekong River Delta. The estimations of wind potential is around 26.7 GW (MoIT&DEA ; Cattelaens et al. 2015).

According to the Renewable Energy Development Strategy (REDS) up to 2030 (Prime Minister 2015), Vietnam will promote on-shore wind power and will assess the potential for off-shore wind resources as an electricity solution after 2030. The targets

for the electricity production from wind energy are 16 billion kWh and 53 billion kWh in 2030 and 2050, covering 2.7% and 5% of the total electricity production, respectively.

2.3 Hydroelectric

The river and stream systems of Vietnam are dense, and distributed in the whole country with more than 2200 rivers and streams stretching over 10 km long (Cattelaens et al. 2015). The hydropower potential in Vietnam is up to 26.7 GW, including about 20 GW for medium and large scale, and 6.7 GW for small scale (EREA&DEA 2019). In the PDP7 rev (Prime Minister 2016), it is planned to extend the electricity production up to 88.7 billion kWh by 2030, covering 17.7 billion kWh for small hydropower. So far, most of medium and large hydropower projects have been fully exploited.

Classification of hydropower projects in Vietnam is considered based on the total generation capacity (WB 2017):

– Large hydropower with a capacity of over 100 MW/plant;
– Medium hydropower with a capacity of under 100 MW/plant;
– Pumped-storage hydropower with any capacity;
– Small hydropower plant (SHP) with a capacity below 30 MW.

According to (WB 2017) 1000 SHP sites have been identified as potentially exploitable for electricity generation. Out of these, about 190 sites are in operation, 180 are under construction, and over 600 sites are under investment assessment or planned for installations.

– The category of SHP is also divided into different scales:
– Micro-hydropower systems, owned by households in mountainous rural areas, having capacities in range of 200–1000 W for lighting during flood season.
– Grid-connected hydroelectric systems provide power only to independent mini grid systems with a typical power output ranging from 1 kW to 1 MW.
– Grid-connected hydro systems have power ranges from 1 to 30 MW.

2.4 Biomass Energy

Being an agricultural country, Vietnam has a high potential of biomass energy production with about 7 GW (EREA&DEA 2019). The agricultural waste is most abundant in the Mekong River Delta with over 21 million tons and is responsible for approximately 50% of the main biomass residues in the whole country (Cattelaens et al. 2015). The next, in terms of agricultural waste production, is the Red River Delta with about 15% of the country's main residues. The main biomass residues hereby come from bagasse, rice husk, rice straw, maize trash and wood fuel. These residues

are used mostly for energy supply in local areas, including heat and electricity, such as fuel for cooking, brick burning, crop drying, energy generation, or cogeneration plants. According to REDS (Prime Minister 2015), the targets of electricity production from biomass are up to 37 billion kWh in 2030 and 85 billion kWh in 2050. There are still many obstacles, such as high investment and low returns. The government is working on more supporting policies and land and tax incentives to reduce the costs of importing spare parts and equipment.

3 National Policy Framework

3.1 Policies for Remote Areas and Rural Electrification in Vietnam

Policy framework has a great influence on the development and implementation of microgrid networks in Vietnam. Vietnamese Government has built a clear policy framework with a set of principles, long-term goals, and national commitments to a defined program as described below.

In the PDP7 rev, the strategy to develop and create rural power supply is indicated. Government will stimulate Vietnam Electricity to develop the national power grid to supply power to most rural households by 2020. The aims are:

– To further develop the national power system to supply efficiently and with highly quality sufficient electricity to meet the power demand for production and residential purposes in rural areas. In case areas cannot meet conditions to access the national power grid, the Government provides investment and support policies for development of local power resource to ensure that by 2020 the ratio of households where electricity is available is 100%.
– To provide support policies to help developing socio-economic situations, including developing power supply system for provinces and poor households in remote areas, especially if it concerns ethnic minorities, in order to strengthen ethnic's solidarity, maintain defense security, ensure the living and production of people, and improve physical and mental living.
– To have a Governmental program to development investment and power supply to every hamlets and minority ethnics of Tay Nguyen.
– To upgrade the rural power grid to increase supply capability and electricity quality and reduce power loss in power lines.

The rural electrification (REII) project is expected to fund such developments. The project was a prominent development cooperation between the Vietnamese government and the World Bank, in effect since 2005 (Sinh et al. 2017). The total loan capital of the project was 420 million USD between 2005 and 2014. The objective of the project was to invest in expanding the power grid to meet the electricity needs for socio-economic development in rural areas. Strengthening the reliability

and quality of low-voltage grid, reducing power loss in the grid, lowering the price of electricity for rural households, improving business efficiency and electricity management capacity in rural communes were some of the achieved objectives.

The main legislations in Vietnam related to rural electrification laws are listed below. These policy measures have supported the development of the rural and remote areas through electrification.

- Decision 22/1999/QD-TTg on approval of the rural electricity scheme by 2000, dated 13 February 1999. The goal of the scheme was to bring electricity to all provinces and districts nationwide by 2000, as well as 80% of communes and 60% of farmers having electricity for living and work.
- Law 28/2004/QH11 on Electricity, dated 14 December 2004. The law was dealing with contents such as: policies for developing rural, mountainous and island electricity; invest in developing electricity and retail price of electricity in rural, mountainous, island.
- Decision 110/2007/QD-TTg on approval of the national power development planning in the period of 2006–2015, with a vision to 2025 (called the Power Development Planning VI—PDP6), dated 18 July 2007. The document mentioned the continued implementation of rural electricity development programs approved by the Prime Minister, with a target of 95% of communes accessed to electricity by 2010.
- Decision 1855/QD-TTg on approval of Vietnam's national energy development strategy up to 2020, with a vision to 2050 dated 27 December 2007. The document mentioned the promotion of the development of new and renewable energy sources, bioenergy, nuclear power to meet the needs of socio-economic development, increasing the total rural households using trade energy for cooking to 80% and most rural households should access electricity by 2020.
- Decision 177/2007/QD-TTg approving the "Development scheme of biofuels by 2015, with a vision to 2025″ dated 20 November 2007;
- Decision 21/2009/QD-TTg on electricity sale prices in 2009 and market mechanism in period of 2010–2012, dated February 12, 2009.
- Decision 1208/QD-TTg on the approval of the national power development planning in the period of 2011–2020 with a vision to 2030 (Master Plan VII), dated 21 July 2011. The document referred to issues such as: new investment on the national grid or on-site power sources (small and micro hydropower; solar cells, wind sources combined with diesel) with the goal of having most rural households supplied with electricity by 2020; renovating and upgrading the rural electricity grids in order to efficiently supply electricity with assurance quality for daily-life, production and development needs of rural areas.
- Circular 97/2008/TT-BTC, by Ministry of Finance about "Circular on implementing state support policies for investment, development of electricity in rural, mountainous and island areas" dated on 28 October 2008;
- Decision 8217/QD-BCT on approval of Renewable Energy Development Plan for Delta, midland area up to 2020 and a vision to 2030, dated 28 December 2012. The document highlighted the issues: promoting the maximum and effective

exploitation and use of renewable energy sources in the region; fulfilling the objective of rural electrification of the Government on electricity supply to remote, mountainous, island areas.

– Decision 2081/QD-TTg on approval the electricity supply program for rural, mountainous areas and islands in the period of 2013–2020, dated 8 November 2013. The goal of the program was to bring electricity to most rural households across the country; investing in the development of grids to supply electricity from the national grid to remote rural communes with a not too high investment ratio; investing in power supply with on-site energy sources (renewable energy sources, battery charging stations …).

– Decision 428/QD-TTg of approval of the Revised National Power Development Master Plan for the 2011–2020 Period with the vision to 2030, dated 18 March 2016. The document mentioned the continuation of the Program on power supply for rural, mountainous and island areas under the Prime Minister's approval decision; also in this law renewable energy sources are mentioned as amin tool for supplying rural, mountainous and island areas so that most rural households have access to electricity by 2020.

– Decision 1740/QĐ-TTg of approval of the electricity supply program for rural, mountainous areas and islands in the period of 2016–2020, dated 13 December, 2018. The decision mentioned the supply of electricity from the national grid and/or new and renewable energy sources to rural, mountainous and island areas, in order to achieve the goal by 2020 for most rural households accessing electricity. Specifically, 100% of communes will have access to electricity; 9890 villages and 1,055,000 households will have to be reached from the national grid; and about 21,000 households could get electricity from renewable energy sources outside the national grid.

It can be seen that the opportunities for renewable energy (biomass, wind, solar, hydro and geothermal) have a high potential in Vietnam. Hydro energy is already extensively implemented in Vietnam, solar and wind project for off-grid communities are also developed but on a less commercial basis. It can be concluded that there is a strong interest and need for improved enabling technologies, consisting of—but not limited to—a further development policy framework that supports the (commercial) development of access to renewable energy. The optimization research for microgrid networks needs to be carried out in parallel to promote the effectiveness of existing projects as well as being applied for future projects.

3.2   Challenges

By taking priority in investment and use of renewable energy sources, the renewable energy market is being developed in Vietnam (Vietnam Electricity 2018a) at all scales. Organizations and individuals are given financial support for the development and use of renewable energy and their legitimate benefits are protected under laws

and regulations. But currently, the renewable energy market has not established yet in the country and the policies for attracting investors to develop renewable energy have not been successful enough.

On the other hand, the Energy Service Companies (ESCOs), providing energy and energy-efficiency improvement solutions to customer (Vietnam Electricity 2018b), have important role in facilitating the deregulation of energy market. However, they are still a few (only 6 by 2015), and their development is still limited. Access to finance is certainly one of the significant barriers for the energy efficiency market. Only some of the local financial institutions have created green financing business line for lending, and this is a limited part of the portfolio. Many Energy service companies have a small size and do not hold the same creditworthiness or resources as the larger companies, thus they have only limited access to capital from the local banking sector. It will take several years to promote the Energy Efficiency market and relevant operating companies.

Currently, the electricity tariff is pretty low, only around at 6–9 UScents/kWh (Vietnam Electricity, 2019). Increasing the electricity price would encourage the development of technologies as well as attract the investors into the renewable energy resources expansion. But with the current low tariffs, it is hard to find loans from banks for renewable energy projects.

Feed-in-tariffs (FiTs) have been set for wind in 2011 and have been revised in 2018 (Prime Minister 2011; Prime Minister 2018); for biomass and waste incineration they have been defined in 2014 and have been revised in 2020 (Prime Minister 2014; Prime Minister 2020); for solar PV farms and rooftop solar PV systems they have been set 2017, revised in 2019 and again revised in 2020 (Prime Minister 2017; Prime Minister 2019; Prime Minister 2020). However, they seem not to have produced significant results. Vietnam is one of the countries with the lowest FiTs in the world (Vietnam Briefing 2019). The FiTs is currently under the electricity Long Run Marginal Cost curve. It means that these tariff levels are not economically practicable and, in the long term, the best economic option for plant owners' is to shut down the plants.

Regarding wind power projects, the investors need to have at least 20% of the total investment, the rest can be loans. Therefore, for large wind power projects, the mobilization of investment capital is challenging for some developers. Also, dealing with the banks about arranging loans is also the major obstacles as well despite there is some preferential loan program for developing renewable energy resources run by the banks. As a consequence of the fact that the FiT rates are relatively low, investors cannot demonstrate the rate of return of the projects to the banks. Hence, moneylending for wind power project are not being approved. Although the goverment has regulated the subsidy mechanism for wind power price in (Prime Minister 2011; Prime Minister 2018), the selling price for on-shore wind power has been set to 8.5 UScents/kWh and for off-shore wind power to 9.8 UScents/kWh. At this stage, it is however still too soon for the wind power developers to see the effectiveness of the new mechanism in Vietnam.

Regarding solar power projects, the current FIT price for floating solar power plants is 7.69 UScents/kWh, for ground solar power plants is set to 7.09 UScents/kWh and for rooftop solar systems, 8.38 UScents/kWh (Prime Minister 2020). Although

this price is higher than the other generation renewable energy resources price, this price is still not economically viable for all solar power projects. In Vietnam, only solar water heating can bring high profits for solar investors, especially when the equipment can be produced domestically. While the concentrated solar power projects for the national grid is less profitable because their production cost is higher than other generation sources, the rooftop solar photovoltaic is economically not convenient because of the high cost and low generation quality. However, it would take some years to see any effect from this supportive mechanism.

4 Prospects and Barriers for Microgrid Development in Vietnam

Electricity consumption in Vietnam increased about 13% per year in the period of 2006–2010 and about 10.5% per year from 2011 to 2018. In 2018, the total energy consumption reached 192.93 billion kWh (Ministry of Industry and Trade of Vietnam 2018). Forecast of commercial electricity consumption in 2020 is about 235–245 billion kWh; in 2025 is about 352–379 billion kWh; in 2030 about 506–559 billion kWh; Production and import electricity in 2020 is about 265–278 billion kWh; in 2025, about 400–431 billion kWh and about 572–632 billion kWh in 2030 (Vietnam Electricity 2016b).

In order to meet the electricity power demand till 2030, Vietnam must find the way to effectively use domestic primary energy sources in combination with rationally import of electricity and fuel to diversify energy sources. The renewable energy sources should be prioritized to develop. This would help to ensure the national energy security, contribute to the conservation of energy resources and minimize negative impacts on the environment in electricity production activities.

However, the approved solar and wind power projects are concentrated in areas with high potential: the South-Central Coast and the Central Highlands where, due to the limitation of the regional transmission network, many technical problems may occur, while other areas where large amounts of electricity can be consumed have not been proposed by investors.

This makes the level of voltage harmonic distortion on power grid areas focusing on renewable energy sources higher than the allowable values. In case all renewable energy sources would have been added to maximize capacity, the rate of harmonic pollution would be exceeded in many 110 kV substations in Ninh Thuan and Binh Thuan provinces, the highest Total Harmonic Distorsion being 4.19%. Besides, the frequency regulation reserve capacity of the power system to meet the frequency adjustment process of the national power system cannot even respond well to these changes. Therefore, Vietnam Power System needs to significantly improve the transmission and subtransmission network infrastructure to receive a large amount of renewable energy as planned. To reduce the pressure on the transmission grid and to

promote the advantage of renewable energy sources, Vietnam is focusing on developing distributed renewable power sources, forming small grids like microgrids to supply electricity demand for industrial, commercial and residential customers. For this reason, there are many great opportunities for growing the microgrid network in Vietnam.

Besides the benefits, the rapidly increasing trend of microgrid integration into electricity grid presents technical, regulatory and economic barriers and technical problems, such as grid synchronization, voltage stability, control and protection (Vandoorn et al. 2013; Majumder 2013; Basak et al. 2012).

The barriers to microgrid promotion and deployment can be summarized as follows:

– Matching solutions to the local needs: to access the benefits of flexibility and scalability solutions need to be designed, which match the individual needs of a community. Whereas a centralized grid provides energy for whatever need, decentralized energy systems must be designed to match the local needs. This design process requires detailed understanding of load use, load specification, ability to pay and other site-specific factors, adding complexity to design.
– Efficient loads for productive use: to access the benefits of flexibility, particularly the ability to power livelihoods, more efficient loads are required to ensure that powering these loads from renewable energy is affordable. However, for non domestic loads these efficiency improvements are yet to be easily available for endusers.
– In addition, there are a lot of barriers that relate to bad policy design, or discontinuity and/or insufficient transparency of policies and legislation, the institutional and administrative barriers, market barriers, such as inconsistent pricing structures that disadvantage are renewables, asymmetrical information, market power, subsidies for fossil fuels, and the failure of costing methods to include social and environmental costs.
– Need of local technical capacity to develop microgrids projects. Still there is not enough investment from the government in Research and Development projects. Only recently and with large delays, some projects on renewable energy development have been supported by the Vietnamese government within the Bilateral Cooperation with Italy.

5 Regulating Voltage and Frequency for Minimum Power Losses in Microgrids

As already said, microgrids are small or medium scale electric distribution networks that include multiple loads, distributed energy resources and storage systems (Soshinskaya et al. 2014). Microgrids can be operated either in grid-connected or islanded modes (Guerrero et al. 2011). In these systems, the frequency and voltage are the

main features to be controlled so as to support the optimal power sharing between the generators (Yunwei et al. 2004; Katiraei 2006; Hatziargyriou et al. 2007).

In Microgrids, a three-level control hierarchical architecture (Guerrero et al. 2011) allows providing good power quality. The primary control is usually designed to use droop-control method to stabilize the voltage and frequency for microgrids and regulate the power sharing between distributed generators. The secondary control is designed to compensate the voltage and frequency deviation caused by primary control. The tertiary control is the last control level that controls the power flow in the microgrid and between the microgrid and the grid. To develop microgrid network, there are 3 main technical challenges need to be addressed:

– Voltage and frequency control: Voltage and Frequency should be maintained in desire limits. A system with variable power sources should have their frequency and voltage droop control for the proper operation of loads those are sensitive to frequency and voltage changing (Bollen et al. 2009; Tran et al. 2018).
– Protection: to let microgrid operate flexibly and be on-off to the main grid, create a goof protection system is the most challenging task for implementation. The protection coordination should be deployed and be equally capable to work for the interconnected system as well as isolated DGs operation.
– Load sharing: The load demand can increase or decrease in different areas. The microgrid would be supplying power to the users even when the real load demand requirement of the system is very less. A proper controller would be required for scheduling the generations from DGs in case of low load demand.

In this part, we proposed some application that would help to support for developing renewable energy integration and widespread microgrid in Vietnam as well. A droop control methodology that is proposed to regulate voltage and frequency in system aim to reduce the power losses for microgrids are presented. Based on the OPF solution, the droop coefficients in the droop control function are adjusted following the variation of the loading conditions, and a unique piecewise droop curve for the controlled generator is built from a set of optimized operating points. In this way, the distributed generator's power output can be adjusted more flexibly, the operational efficiency of the microgrid is increased, and the power losses are reduced.

With the networks considered to be lossy, its operation principle and control method are explained. Simulation results of a 9-node 3-inverter system with the proposed control, with optimal selection of the parameters, enhance the output voltage of inverters and validate the effectiveness of the proposed control method.

5.1 Problem Formulation

Consider a microgrid system with i DG units connected to the loads through resistive-inductive lines. With conventional linear droop control, the frequency and voltage primary adjustment of DGs can be implemented by linear droop control technique as already outlined in Eqs. (1) and (2) below.

$$P_{Gi} = -K_{Gi}(f - f_{0i}) \tag{1}$$

where K_{Gi} is the frequency droop coefficient chosen based on active rating of DGs; f_{0i} is the rated frequency.

$$Q_{Gi} = -K_{di}(|V_i| - V_{0i}) \tag{2}$$

where K_{di} is voltage droop coefficients, chosen based on the reactive rating of DGs; V_{0i} is the rated voltage.

In linear droop control, the values of K_{Gi} and K_{di} are usually constant and equal to the maximum value. In the improved method here proposed, the droop parameters K_{Gi} K_{di} of generators are chosen optimally in the range $[K_{Gmini}; K_{Gmaxi}]$ and $[K_{dmini}; K_{dmaxi}]$, respectively.

This droop control method could be implemented in a system which includes fast response generators like wind generators, PV systems which have fast response storage. The standards (IEEE Std 122-1985:0_1 1985) and (Zhao et al. 2016) prove that the droop parameters of generators can be adjusted. The different K_G and K_d values should be chosen under different generated output power. In this case, different droop relations are needed for the resynchronization with the grid in different operating conditions.

The Otimal Power Flow, OPF, for each loading condition, outputs a minimum loss operating state. The latter is an expanded state comprising system's frequency and the droop parameters. The OPF underlying the identification of the piecewise linear droop law is expressed as shown in (Quang et al. 2014).

The active power and reactive power injected at bus i can be expressed as follows:

$$P_i = \sum_{i=1}^{n_{\ln s}} \sum_{j=1}^{n_{h_r}} |V_i||V_j||Y_{ij}| \cos(\theta_{ij} - \delta_i + \delta_j) \tag{3}$$

$$Q_i = -\sum_{i=1}^{n_{005}} \sum_{j=1}^{n_n} |V_i||V_j||Y_{ij}| \sin(\theta_{ij} - \delta_i + \delta_j) \tag{4}$$

where:

n_{br} is the number of branches which are connected to bus i.

n_{bus} is the number of buses in the microgrid.

V_i, δ_i are the voltages and phase angle of the voltage at bus i; V_j, δ_j δ_{i_h} are the voltage and phase angle of the voltage at bus j. These values depend on K_d and K_G at droop buses;

Y_{ij} is the admittance of branch ij; θ_{ij} is the phase angle of Y_{ij};

The objective function of the OPF, which is the total real power losses, can be expressed as follows:

$$\text{Min} \ \ OF_{(K_G)} = P_{\text{Loss}} = \sum_{i=1}^{n_{\ln m}} P_{i(K_G)} \tag{5}$$

Assuming that the generator Gi is available to regulate droop coefficients for the n_G generators in an islanded microgrid system and to drive the contribution of each source to the overall load demand, the other generators have the fixed droop coefficients K_{Gi} and K_{di}. The objective function (6) must be subjected to the following constraints:

$$\begin{cases} \sum_{i=1}^{n_6} P_{Gi} = \sum_{i=1}^{n_d} P_{Li} + P_{\text{loss}} \\ \sum_{i=1}^{n_6} Q_{Gi} = \sum_{i=1}^{n_d} Q_{Li} + Q_{\text{lass}} \\ K_{\text{Gimin}} \le K_{Gi} \le K_{\text{Gimax}} \\ K_{\text{dimin}} \le K_{di} \le K_{\text{dimax}} \\ f_{\text{min}} \le f \le f_{\text{max}} \\ V_{\text{min}} \le V_i \le V_{\text{max}} \end{cases} \tag{6}$$

The solutions, for each loading condition, are one steady state operating point with a single frequency and voltage assuming as reference bus (with zero displacement) one of the grid forming units buses. The problem is solved by a Glow-worm Swarm Optimization (GSO) method, a heuristic method (Krishnanand and Kaipa 2017) which has been implemented within Matlab environment.

For example, consider an islanded microgrid system including 2 generators. DG1 has the ability to adjust the droop coefficients, DG2 has fixed droop coefficients. The model of droop coefficients regulation method can be shown in Fig. 3.

5.2 Case Study and Results

A test system which includes 9 nodes is considered to illustrate the effectiveness of the improved regulation methodology (Fig. 4).

Two scenarios are considered:

•Scenario 1, implementing the conventional droop control method.

•Scenario 2, implementing the proposed optimized control method with KG1 selected optimally in the proposed range.

To emphasize the effectiveness of the regulation method for power sharing between generators, assume that the largest load in the network at bus 9 varies in 8 steps from 0.64 to 1.34 pu and assume that the loads at the other buses do not change. After the calculation, the improvement of power losses and other results are shown in pu as reported in the Tables 1, 2, 3 and 4.
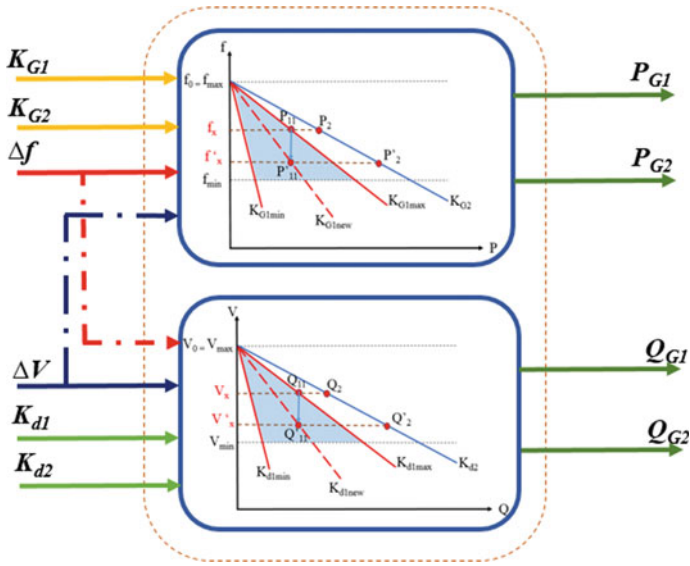
Fig. 3 Model of droop coefficients regulation method
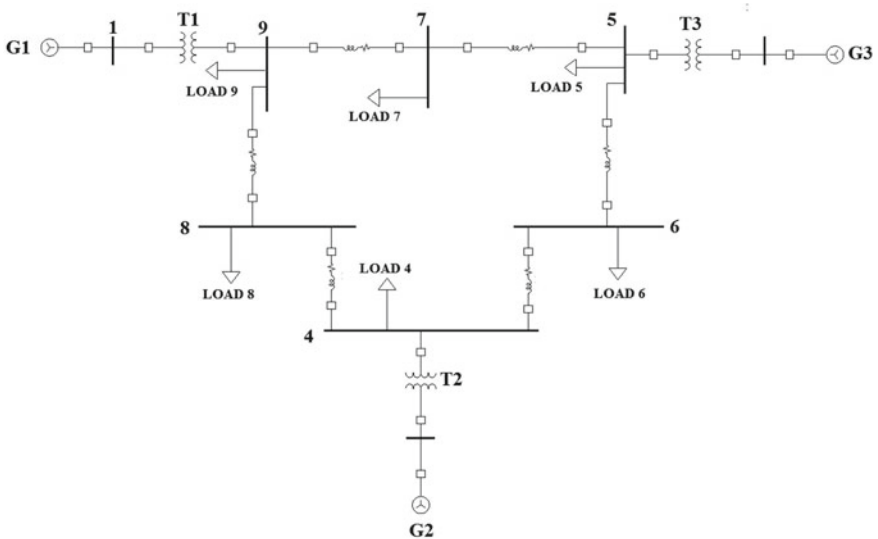
Fig. 4 9-bus test system

Table 1 The results in pu with linear droop regulation

Case	KG2	Kd2	PlOSS*100	f
1	20.0	6.0	0.5125	0.9952
2	20.0	6.0	0.5396	0.9943
3	20.0	6.0	0.5704	0.9933
4	20.0	6.0	0.643	0.9913
5	20.0	6.0	0.7296	0.9894
6	20.0	6.0	0.8313	0.9875
7	20.0	6.0	1.0136	0.9846

Table 2 The results in pu with linear droop regulation

Case	VI	V2	V3	V4	V5	V6	V7	V8	V9
1	1.015	1.0215	0.9989	1.0127	0.9945	0.9988	0.9864	1.0013	0.9974
2	1.0141	1.021	0.9979	1.0121	0.9935	0.9979	0.9852	1.0003	0.9959
3	1.0132	1.0206	0.9968	1.0114	0.9924	0.997	0.9S4	0.9991	0.9944
4	1.0112	1.0195	0.9945	1.01	0.9901	0.995	0.9814	0.9967	0.9913
5	1.0095	1.0186	0.9925	1.0087	0.988	0.9932	0.9792	0.9946	0.9885
6	1.0075	1.0174	0.9901	1.0072	0.9S57	0.9912	0.9766	0.9921	0.9S55
7	1.004	1.0153	0.9857	1.0042	0.9S12	0.9872	0.9717	0.9S75	0.9798
8	0.9997	1.0125	0.98	1.0002	0.9755	0.9822	0.9655	0.9817	0.9727

Table 3 The results in pu with improved droop regulation

Case	KG2	Kd2	Ploss*100	f
1	12.3	4.1	0.4248	0.9909
2	12.1	4.2	0.4262	0.9896
3	12.0	4.9	0.4295	0.9884
4	12.0	4.4	0.4399	0.9861
5	12.1	4.1	0.4582	0.9839
6	12.0	4.5	0.4842	0.9816
7	14.0	5.0	0.6313	0.9800
8	17.8	5.5	1.0348	0.9800

It can be observed that the power losses were reduced of about 18% as compared to the conventional droop control method. For optimizing the regulation, it can be seen that the frequency and voltage values are regulated to their limit at every loading condition.

Table 4 The results in pu with improved droop regulation

Case	VI	V2	V3	V4	V5	V6	V7	V8	V9
1	1.0136	1.0148	0.9946	1.0063	0.9902	0.9937	0.9826	0.9974	0.9956
2	1.0128	1.0150	0.9940	1.0062	0.9896	0.9934	0.9818	0.9968	0.9945
3	1.0124	1.0169	0.9945	1.0075	0.9901	0.9942	0.9819	0.9972	0.9941
4	1.0100	1.0141	0.0912	1.0047	0.0868	0.0911	0.9785	0.9938	0.9903
5	1.0080	1.0119	0.0886	1.0025	0.0841	0.0886	0.9757	0.0911	0.9871
6	1.0064	1.0126	0.0876	1.0025	0.0831	0.0880	0.9742	0.0900	0.9850
7	1.0036	1.0128	0.9850	1.0019	0.9805	0.9862	0.9710	0.9872	0.9803
8	0.9995	1.0114	0.9799	0.9994	0.9754	0.9819	0.9654	0.9817	0.9730

6 Conclusion

From the general overview of the renewable energy resources in Vietnam, it can be seen that the renewable energy options have the potential to become the lowest-cost option for the country to meet its energy needs. The development of microgrid network would help to increase the renewable energy penetration and, at the same time, increase the reliability and power quality of the Vietnam's power system in the future. This also helps to improve the livelihood of the rural communities and promote decentralization of electric power in Vietnam.

In this chapter, an enhanced operation mode for islanded microgrids is proposed. Through an improved droop regulation methodology, the droop coefficients are changed, and new feasible and optimized set points for distributed generators are found to minimize the microgrid operation losses and satisfy the grid and generators constraints. In this way, the frequency is kept within the desired limits. A case study is implemented to demonstrate the effectiveness of the proposed method in different loading conditions. The achieved results were compared to those obtained with the conventional droop regulation. It was shown that, when a load variation occurs, the proposed droop regulation adjusts the droop coefficients, minimizes the power losses, and maximizes the efficiency of power sharing between distributed generators.

References

Basak P, Chowdhury S, Halder nee Dey S, Chowdhury SP (2012) A literature review on integration of distributed energy resources in the perspective of control, protection and stability of microgrid renewable and sustainable energy reviews. Sciencedirect 16:5545–5556

Bich Huy N, Anh Tuan N (2015) Current situations and solutions for renewable energy development in Vietnam. Asia Pacific J Sustain Agric Food Energy (APJSAFE) 3:20–23

Bollen M, Zhong J, Lin Y (2009) Performance indices and objectives for microgrids. In: CIRED 2009—20th international conference and exhibition on electricity distribution—Part 1, 8–11 June 2009, p 1–4

Cattelaens P, Limbacher E-L, Reinke F, Stegmueller FF, Brohm R (2015) Overview of the Vietnamese power market—a renewable energy perspective. GIZ Energy Support Programme

Guerrero JM, Vasquez JC, Matas J, Vicuna LGd, Castilla M (2011) Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization. IEEE Trans Industr Electron 58:158–172

Hatziargyriou N, Asano H, Iravani R, Marnay C (2007) Microgrids. IEEE Power Energ Mag 5:78–94

Hoat DD, Tuyen TK, Hang LTT, Thanh NV, Thanh TQ, Quoc TH, Minh NT (2007) Research overview of new and renewable energy in Vietnam and development orientation. Vietnam Academy of Science and Technology

IE (2016) Detail report of the revised national power development master plan for the period 2011–2020 with a vision to 2030. Institute of Energy

IEA (2017) Southeast Asia Energy Outlook (2017) International Energy Agency

IEEE recommended practice for functional and performance characteristics of control systems for steam turbine-generator units (1985) IEEE Std 122-1985:0_1

Katiraei F, Iravani MR (2006) Power management strategies for a microgrid with multiple distributed generation units. IEEE Trans Power Syst 21:1821–1831

Krishnanand N, Kaipa DG (2017) Glowworm swarm optimization: theory, algorithms, and applications. Studies in Computational Intelligence. Springer. Berlin

Majumder R (2013) Some aspects of stability in microgrids. IEEE Trans Power Syst 28:3243–3252

Ministry of Industry and Trade of Vietnam (2018) Report on activities of industry and trade in 2018 (publication in Vietnamese). https://moit.gov.vn/web/guest/bao-cao-tong-hop1. Accessed 19 April 2019

MoIT&DEA (2017) Vietnam Energy Outlook 2017. Danish Energy Agency

Polo J, Martínez S, Fernandez-Peruchena CM, Navarro AA, Vindel JM, Gastón M, Ramírez L, Soria E, Guisado MV, Bernardos A, Pagola I, Olano M (2015) Maps of solar resource and potential in Vietnam. Vietnam Ministry of Industry and Trade & Spanish Agency for International Development Cooperation

Prime Minister (2016) Decision 428/QD-TTg of Approval of the Revised National Power Development Master Plan for the 2011–2020 period with a vision to 2030. Vietnam Government

Prime Minister (2015) Decision 2068/QD-TTg of Approving the Viet Nam's Renewable Energy Development Strategy up to 2030 with an outlook to 2050. Vietnam Government

Prime Minister (2013) Decision 2081/QD-TTg of Approving the Program on electricity supply in rural, mountainous and island areas in period of 2013–2020. Vietnam Government

Prime Minister (2011) Decision 37/2011/QD-TTg on the mechanism to support the development of wind power projects in Vietnam. Vietnamese Government

Prime Minister (2018) Decision 39/2018/QD-TTg on amending and supplementing a number of articles of Decision 37/2011/QD-TTg on mechanisms to support the development of wind power projects in Vietnam. Vietnamese Government

Prime Minister (2014) Decision 24/2014/QD-TTg on the mechanism to support the development of biomass power projects in Vietnam. Vietnamese Government

Prime Minister (2020a) Decision 08/2020/QD-TTg on amending and supplementing a number of articles of Decision 24/2014/QD-TTg on mechanisms to support the development of biomass power projects in Vietnam. Vietnamese Government

Prime Minister (2017) Decision 11/2017/QD-TTg on Mechanism for Encouragement of the Development of Solar Power Projects in Vietnam. Vietnam Government

Prime Minister (2019) Decision 02/2019/QD-TTg on amendments and supplements to certain articles of Decision No. 11/2017/QD-TTG on the mechanism for encouragement of development of solar power in Vietnam. Vietnam Government

Prime Minister (2020b) Decision 13/2020/QD-TTg on Mechanism for Encouragement of the Development of Solar Power in Vietnam. Vietnam Government

Quang NN, Riva Sanseverino E, Di Silvestre ML, Madonia A, Li C, Guerrero JM (2014) Optimal power flow based on glow worm-swarm optimization for three-phase islanded microgrids. In:

AEIT Annual conference—from research to industry: the need for a more effective technology transfer (AEIT), pp 1–6

Sinh TD, Long, NH, Dung NTM (2017) Report on rural electrification policies and supporting mechanisms for off-grid community. Green Innovation and Development Centre

Soshinskaya M, Crijns-Graus WHJ, Guerrero JM, Vasquez JC (2014) Microgrids: experiences, barriers and success factors. Renew Sustain Energy Rev 40(C):659–672 (Elsevier)

The Prime Minister of Vietnam (2017) Decision on mechanism for encouragement of development of solar power in Vietnam (publication in Vietnamese). https://thuvienphapluat.vn/van-ban/Thuong-mai/Quyet-dinh-11-2017-QD-TTg-co-che-khuyen-khich-phat-trien-du-an-dien-mat-troi-tai-Viet-Nam-345919.aspx. Accessed 19 April 2019

Tran TTQ, Di Silvestre ML, Riva Sanseverino E, Zizzo G, Pham T (2018) Driven primary regulation for minimum power losses operation in islanded microgrids. In: Energies, vol 11

United Nations Environment Programme (UNEP) and Global Environment Facility (GEF) (2014) Report on the off-grid lighting status for Southeast Asia and the Pacific https://www.lites.asia/downloads/off-grid-lighting. Accessed 19 April 2019

Vandoorn TL, De Kooning JDM, Meersman B, Vandevelde L (2013) Review of primary control strategies for islanded microgrids with power-electronic interfaces renewable and sustainable energy reviews. Sciencedirect 19:613–628

VEPG (2019) Vietnam factsheet on rooftop solar development 2019. Viet Nam Energy Partnership Group

Vietnam Briefing (2019) Renewables in Vietnam: current opportunities and future outlook. https://www.vietnam-briefing.com/news/vietnams-push-for-renewable-energy.html/. Accessed 19 April 2019

Vietnam Electricity (2016a). Revision of national power development master plan. https://en.evn.com.vn/d6/news/Revision-of-National-Power-Development-Master-Plan--66-163-414.aspx. Accessed 19 April 2019

Vietnam Electricity (2016b). The "dream" of rural electrification (publication in Vietnamese) https://www.evn.com.vn/d6/news/Hien-thuc-hoa-giac-mo-dien-khi-hoa-nong-thon-6-12-18032.aspx. Accessed 19 April 2019

Vietnam Electricity (2018a) National renewable energy development strategy to 2030 with an outlook to 2050 (publication in Vietnamese). https://www.evn.com.vn/d6/news/Chien-luoc-phat-trien-nang-luong-tai-tao-cua-Viet-Nam-den-nam-2030-tam-nhin-den-nam-2050-141-169-22544.aspx. Accessed 19 April 2019

Vietnam Electricity (2018b) Energy service company in Vietnam (publication in Vietnamese). https://www.evn.com.vn/d6/news/Mo-hinh-cung-cap-dich-vu-nang-luong-ESCO-tai-Viet-Nam-141-170-22536.aspx. Accessed 19 April 2019

Vietnam Electricity (2019) Electricity price (publication in Vietnamese). https://en.evn.com.vn/c3/gioi-thieu-l/Electricity-Price-9-28.aspx. Accessed 19 April 2019

Vietnamnet News (2017) Vietnam urged to explore renewable energy. https://english.vietnamnet.vn/fms/environment/174706/vietnam-urged-to-explore-renewable-energy.html. Accessed 19 April 2019

WB (2017) Small hydro resource mapping in Vietnam: final report. World Bank Group

Winrock International Institute for Agricultural Development and SNV Netherlands Development Organisation (2014) Off-grid opportunities and challenges in Vietnam. https://www.snv.org/public/cms/sites/default/files/explore/download/2014-01-24_off-grid_final_report.pdf. Accessed 19 April 2019

Yunwei L, Vilathgamuwa DM, Chiang LP (2004) Design, analysis, and real-time testing of a controller for multibus microgrid system. IEEE Trans Power Electron 19:1195–1204

Zhao J, Lyu X, Fu Y, Hu X, Li F (2016) Coordinated microgrid frequency regulation based on dfig variable coefficient using virtual inertia and primary frequency control. IEEE Trans Energy Convers 31:833–845

Study of Interconnections in Vietnam Power System with Asynchronous Back to Back HVDC Links

Tran Quoc Tuan, Le Quoc Anh, Nguyen Van An, Duong Viet Duc, and Nguyen H. Long

Abstract The Vietnamese power system is growing rapidly both in terms of scale and complexity. Three links with China network in three independent areas with the Vietnam network. This chapter presents a study interconnection possibility between the Vietnam network and the China network by asynchronous back to back high voltage direct current links (B2B HVDC). Firstly, the analysis and the evaluation of the necessity of the B2B HVDC are investigated. Then the simulations are performed in EMTP-RV to identify technical constrains and benefits for two links. Voltage Source Converter (VSC)—Modular Multilevel Converter (MMC) and Line Control Converter (LCC) are used to study. Several scenarios are studied such as voltage control, active and reactive control in normal operation and fault conditions. Finally, a comparison of two technologies (VSC, LCC) is carried out and recommendations are given.

1 Introduction

The Vietnamese power system is growing rapidly both in terms of scale and complexity. The development of the Extra High Voltage network of 500 kV and interconnections with neighboring grids are the main characteristic of such

T. Q. Tuan (✉)
Univ Grenoble Alpes, CEA-LITEN, INES (National Institute for Solar Energy), Le Bourget-du-Lac, France
e-mail: quoctuan.tran@cea.fr

INSTN (Paris Saclay University), Gif-sur-Yvette, France

L. Q. Anh · N. Van An
PECC1, Electricity of Vietnam-EVN, Hanoi, Vietnam

D. V. Duc
National Power Transmission-NPT, Electricity of Vietnam-EVN, Hanoi, Vietnam

N. H. Long
PTC1, Electricity of Vietnam-EVN, Hanoi, Vietnam

growth. The Greater Mekong Sub-region (GMS), includes six countries, namely Lao PDR, Cambodia, Myanmar, Thailand and Vietnam located in South East Asia and China (Yunnan Province and Guangxi Zhuang Autonomous Region of the People's Republic of China). To ensure continuous supply of electrical energy for every member country in the GMS, the cooperation through integration of energy infrastructures becomes essential. Since 2004, interconnections with China network have been realized. In over thirteen years of operation, these links have been changing to adapt the actual situation to the master plan of the National electrical system. Currently, there are three interconnection links between the Vietnam network and the China network: 220 kV Xinqiao—Lào Cai; 220 kV MaLutang–Hà Giang and 110 kV Thâm Câu–Móng Cái. Today, these links with China are connected to two or three areas isolated from the Vietnam network. In function of the purchase power periods, the isolated areas are varied. Figure 1 shows these isolated areas in 2010, 2013, 2015 and 2016, respectively. When these isolated areas are very large, that affects the flexibility and the secure operation of the national grid.

In today's electricity industry, in view of the liberalization, the high penetration of renewable energy and increased effects to conserve the environment, the HVDC (High Voltage Direct Current) solutions have become more desirable for the following reasons (Rudervall et al. 2000; Wikipedia; Mukhedkar; Wikipedia; Sellick and Åkerberg 2012; Oni et al. 2016; Kirby 2012; Kontos et al. 2015; Elliott et al. 2016):
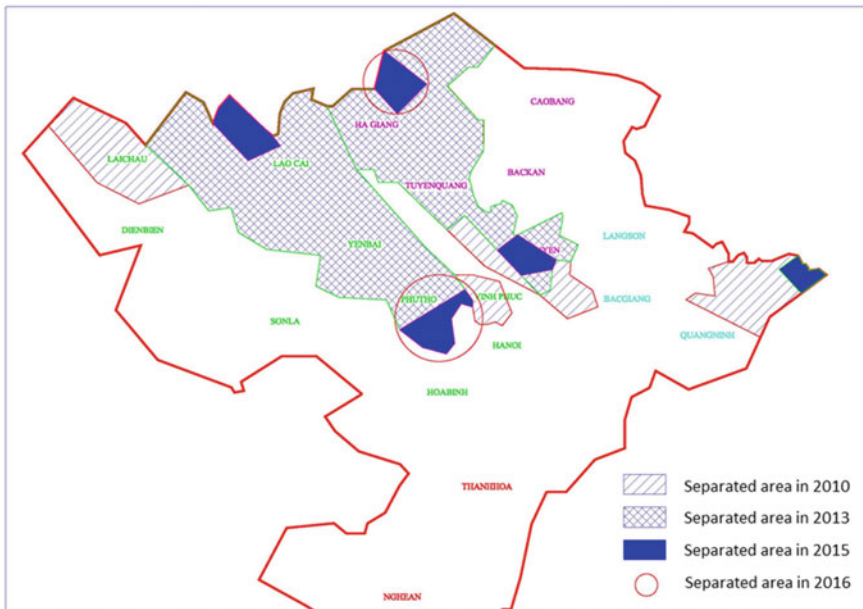

Fig. 1 Isolated areas formed by interconnections between Vietnam and China
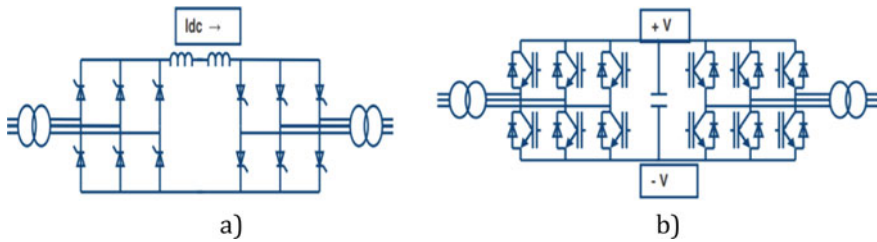
Fig. 2 Architecture of LCC HVDC and VSC HVDC

- Environmental advantages
- Economical issues (appears the cheapest solution)
- Possibility of asynchronous interconnections
- Power flow control
- Stability improvement
- Losses reduction as compared to an equivalent AC transmission scheme (for long connections).

Most of the HVDC systems in operation today are based on line-commutated converters (LCC). The line-commutated indicates that the conversion process relies on the line voltage of the AC system to which the converter is connected in order to affect the commutation from one switching device to its neighbor. The architecture of LCC HVDC is presented in Fig. 2a.

Voltage source converter uses insulated gate bipolar transistor IGBT technology. The current in this technology can both be switched on and off at any time independent of the AC voltage. The architecture of VSC HVDC is presented in Fig. 2b. Recently, a new alternative of VSC-HVDC circuits was proposed and known as modular multi-level converter (MMC). These converter topologies is based on series-connection of several sub-modules of two semiconductor switches and a capacitor.

Back to back connection has both the inverter and the rectifier in the same location, and the valves are normally in the same building. It therefore has a short DC line of few meters located inside the same environment. Back-to-back HVDC is used (Rudervall et al. 2000; Wikipedia; Mukhedkar; Wikipedia):

- to connect asynchronous high-voltage power systems or systems with different frequencies;
- to stabilize weak AC links;
- to supply more active power where the short-circuit capability of the AC system already is limited;
- to control power-flow within synchronous AC systems.

This chapter presents a study of interconnection possibility between the Vietnamese power network and the Chinese power network by asynchronous back-to-back HVDC links (B2B HVDC). Firstly, the analysis and the evaluation of the needs

for the B2B HVDC links are investigated. Then the simulations are performed in EMTP-RV to identify technical constraints and benefits for two links:

- 220 kV, 400 MW B2B HVDC between Xinqiao and Lào Cai;
- 220 kV, 300 MW B2B HVDC between Malutang and Hà Giang.

Voltage Source Converter (VSC)—Modular Multilevel Converter (MMC) and Line Control Converter (LCC) are considered in this study.

Several scenarios are also studied. Such as voltage control, active and reactive control in normal operation and fault conditions.

The comparison of three technologies (VSC, MMC and LCC) is carried out and recommendations are given.

2 Why to Replace the Existing AC Links by B2B HVDC Links?

In this section, the analysis of the motivations behind the replacement of the existing AC links connecting Chinese and Vietnamese power grids by B2B HVDC links is presented. Different studies by using the Siemens software PSS/E for simulating the existing AC links are carried out in order to identify possible weaknesses in future developments.

The analysis of simulation results and the actual operations show that due to the appearance of large isolated areas during the high-power purchase periods, flexible, secure and reliable operation of the national grid can be affected.

During the low purchase power periods, the isolated areas are narrow. However, the power transmission distances at 220 kV isolated areas are too long that increases power losses and makes also difficulties for the operation.

Figure 3 shows the PV curve for 220 kV Yên Bái–Việt Trì transmission lines during a high-power purchase period. The limit capacity for voltage stability is about 26%. When the power purchase is increased, it can cause voltage collapse and increasing power losses.

For this reason, in Việt Trì and Thái Nguyên substations, a Static VAR Compensator system (−50/+50 MVAR) was installed at each substation.

Several steady state and dynamic simulations for different scenarios were carried out to identify disadvantages of these interconnection links and what would happen if these isolated areas cannot be synchronized with the national grid.

For these reasons, the use of B2B HVDC links to replace these AC links is considered to avoid these separations. The following part shows simulations in EMTP-RV to identify technical constrains and benefits for these links are presented.
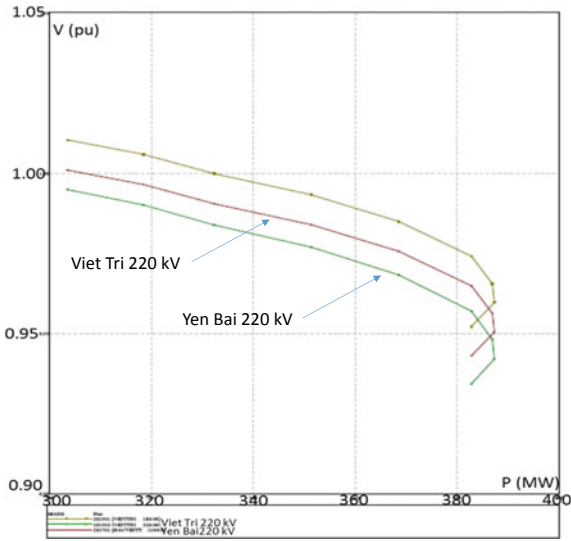
Fig. 3 PV curves for 220 KV Yên Bái–Việt Trì transmission lines during a high electricity purchase period

3 B2B HVDC Solutions

In this section, LCC HVDC and VSC HVDC is modelised in EMTP RV. Figures 4 and 5 show the modelling of LCC and VSC HVDC systems. Figure 6 presents the control system of VSC HVDC.

Figure 7 shows 500 and 220 kV existing AC Northern transmission networks (B2B HVDC included) simulated in EMTP_RV to identify technical constraints and benefits for two new links:
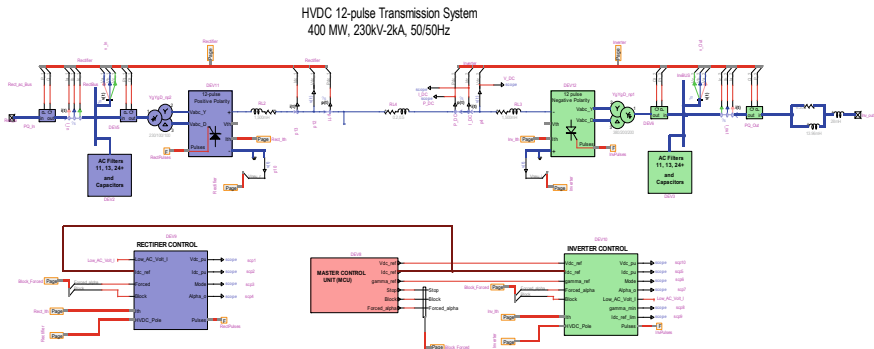
Fig. 4 LCC HVDC modelised in EMTP RV
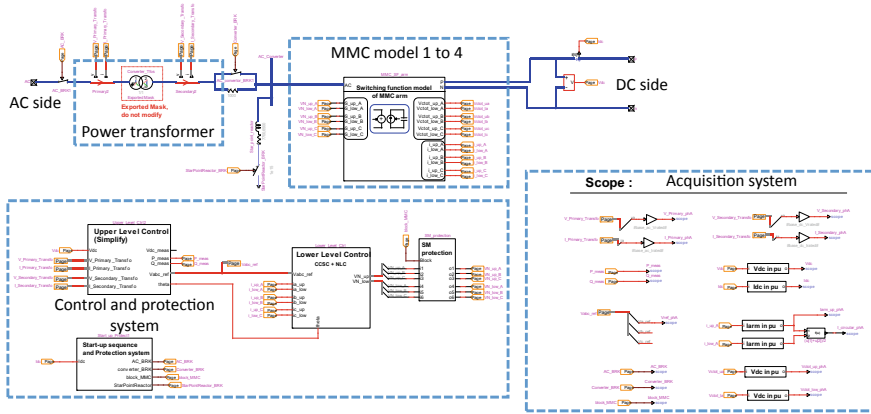
Fig. 5 VSC HVDC modelised in EMTP RV

Fig. 6 Control system of VSC HVDC
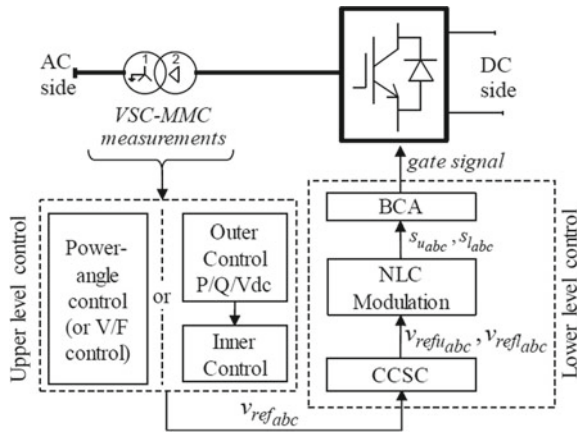
- 220 kV, 400 MW B2B HVDC between Xinqiao and Lào Cai;
- 220 kV, 300 MW B2B HVDC between Malutang and Hà Giang.

In each case, a LCC and VSC are simulated.

Different conversion technologies are considered in this study: Voltage Source Converter (VSC)—Modular Multilevel Converter (MMC) and Line Control Converter (LCC) Figs. 8 and 9.

Two issues are studied:

- voltage control,
- active and reactive control in normal operation and fault conditions.

Lào Cai and Hà Giang are weak points, therefore an analysis of voltage quality and harmonics are carried out.
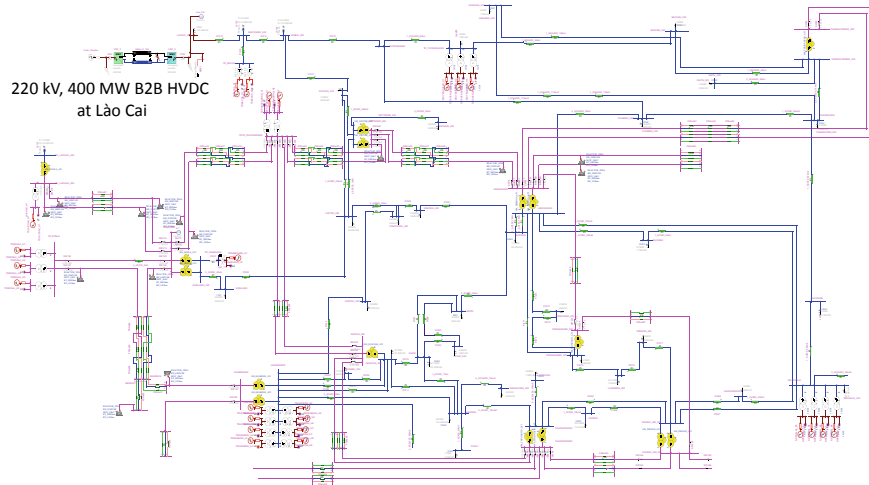
Fig. 7 500 and 220 kV Northern transmission network simulated in EMTP_RV
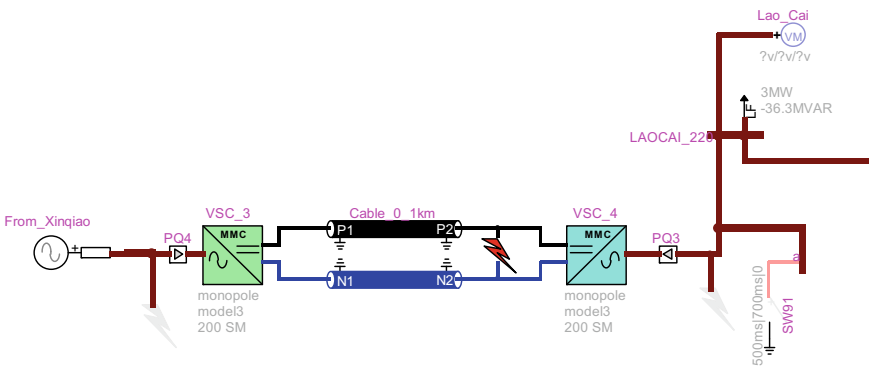
Fig. 8 VSC B2B system connected at Lào Cai

In each scenario, the comparison of the two technologies (VSC, LCC) is carried out and recommendations are provided.

For simplicity reasons, in this chapter only the simulation results for the B2B system connected at Lào Cai are given.

3.1 VSC B2B Installation

Figure 10 shows the response of VSC B2B system in case of the changes of active and reactive set points during normal operation as an effect of changes in power flows. It shows that:
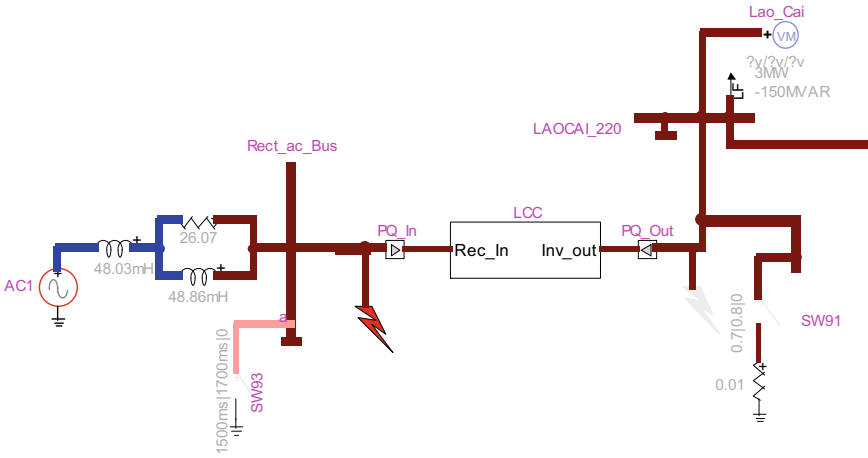
Fig. 9 LCC B2B system connected at Lào Cai
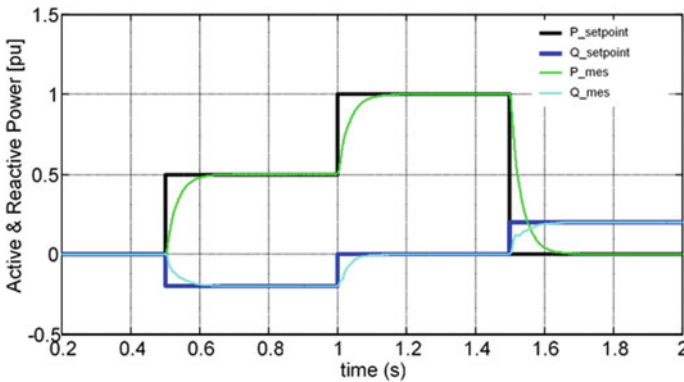
Fig. 10 Active and reactive power set point and active and reactive power obtained by simulation (VSC)

- The VSC B2B system has a fast response in normal operation mode;
- The active and reactive powers can be controlled independently: when the B2B system does not transmit active power ($P = 0$), it can contribute to control voltage by reactive power;
- Transients oscillations and harmonics are very limited (stable voltage when changing the set point).

Figures 11 and 12 show the voltage at Lào Cai and the active/reactive power flows through the VSC B2B in case of a single-phase to ground fault at Xinqiao (fault starting time: tfs = 0.5 s, fault clearing time, tfc = 0.7 s).

Figures 13 and 14 show the voltage at Lào Cai and the active/reactive power of VSC B2B in case of a three-phase fault at Xinqiao (tfc = 0.5s, tfc = 0.7s).
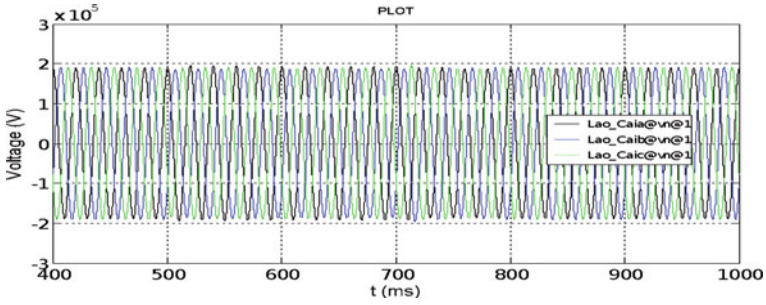
Fig. 11 Voltage at Lào Cai for a single-phase fault at Xinqiao (SVC)
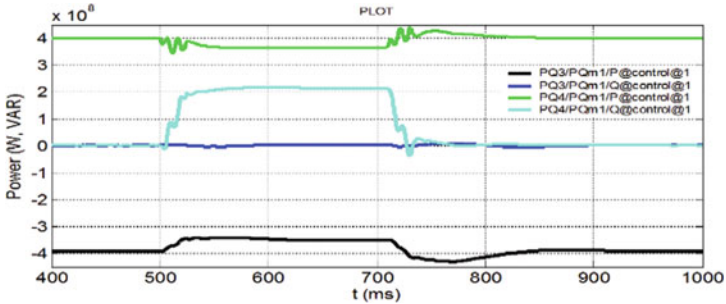
Fig. 12 Active and reactive power of B2B (VSC)
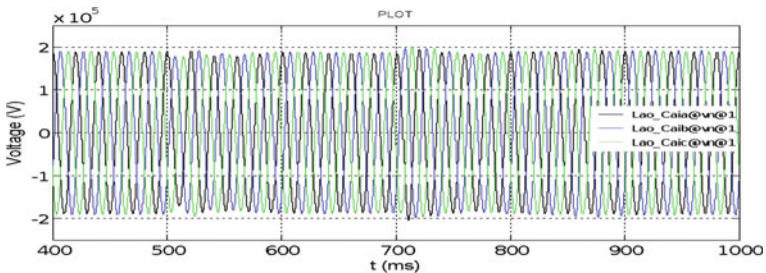
Fig. 13 Voltage at Lào Cai for a single-phase fault at Xinqiao (VSC)

The simulation results show that:

- The VSC B2B system has a fast response for a single-phase or three-phase fault.
- A fault in the China network does not influence the Vietnam network and viceversa Figs. 11 and 12: voltage at Lào Cai, Vietnam for a fault at Xinqiao, China).
- For a single-phase fault, the active power Fig. 11: black and green curves) is lightly reduced during the short-circuit (0.2 s from tfs = 0.5 s) and comes back to the initial value (400 MW) after the fault clearing (tfc = 0.7 s),the reactive power
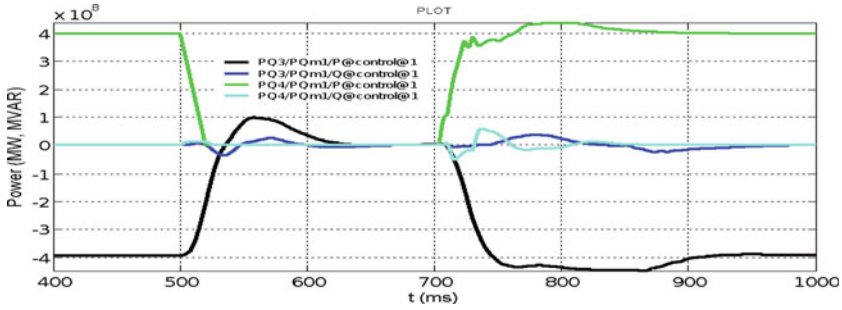
Fig. 14 Active and reactive power of B2B (VSC)

input (blue curve) is not changed but the reactive power output (aqua curve) is changed Fig. 11.

- For a three-phase fault, the active power Fig. 10: black and green curves) is reduced during the short-circuit (0.2 s from tfs = 0.5 s) and comes back to the initial value (400 MW) after the fault clearing (tfc = 0.7 s),the reactive power input (blue curve) and output (aqua curve) are not changed Fig. 10.

3.2   LCC B2B Installation

Here the same simulation as the proceeding section is provided. Figures 15 and 16 show the voltage at Lào Cai and the active/reactive power of LCC B2B in case of a starting and stopping of B2B (tstart = 0.3, tstop = 0.7 s). The simulation results show that:

- Starting process slower than B2B VSC, the active power increases at t = 0.3 s and reaches to 4 MW at t = 1.4 s.
- Active power decreases to zero when the B2B stops at t = 1.4 s.
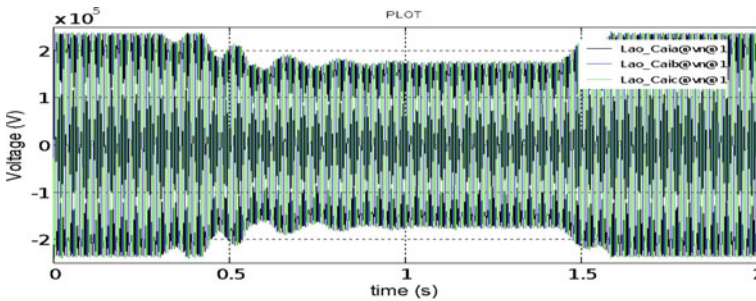
Fig. 15 Voltage at Lao Cai for a start/stop of LCC B2B
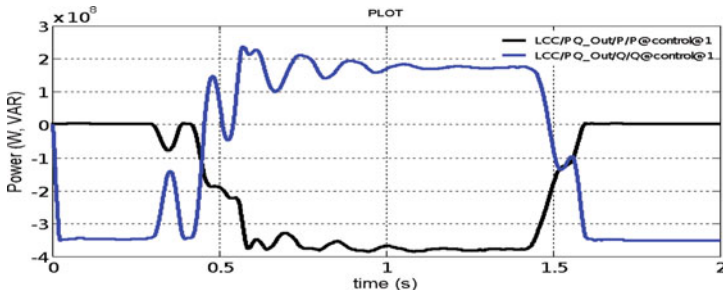
Fig. 16 Active and reactive power of B2B (LCC)

- Initially reactive power generated is more than +3 MVAR (reactive compensation: 2 MAVR + 2 × 0.5 MVAR of 11th and 13th harmonics filters) and this can cause an overvoltage at Lào Cai.
- When B2B is started, the reactive power is about −2 MVAR. The reactive power absorption of B2B can cause a voltage drop at at Lào Cai.
- When B2B is stopped, the reactive power is about + 3 MVAR. The reactive power production of B2B can cause an overvoltage at Lào Cai.

It appears that for the weak node (such as Lào Cai), the installation of a B2B LCC can cause large fluctuations in voltage. Therefore, a voltage control system such as a Static Voltage Compensator, SVC, or STATCOM must be installed at Lao Cai. This complicates the operation and increases the cost.

Figures 17 and 18 show the voltage at Lào Cai and the active/reactive power of LCC B2B in case of a three-phase fault at Xinqiao (tfc = 0.16 s, tfc = 0.17 s). For this case, when a fault occurs, the LCC B2B is stopped and it cannot restart.

Finally, a comparison of two technologies (VSC, LCC) and recommendations are carried out in Table 1. In this table, important performances of B2B are carried out such as power reversal, AC disturbances and dynamic, AC connection, AC voltage support and black start capacity. For power reversal, VSC present many advantages because it doesn't need mechanical switches. In case of disturbances, VSC presents many advantages such as no influence short circuit and high dynamic performance.
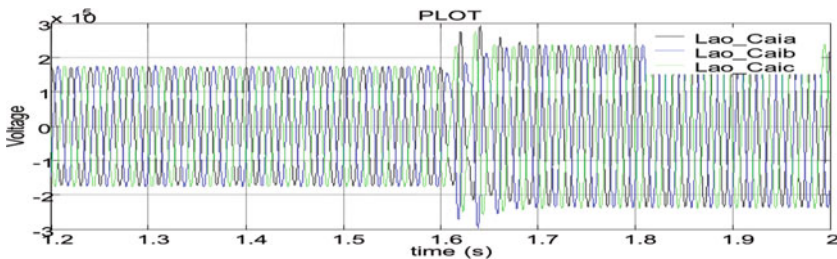
Fig. 17 Active and reactive power of B2B (LCC)
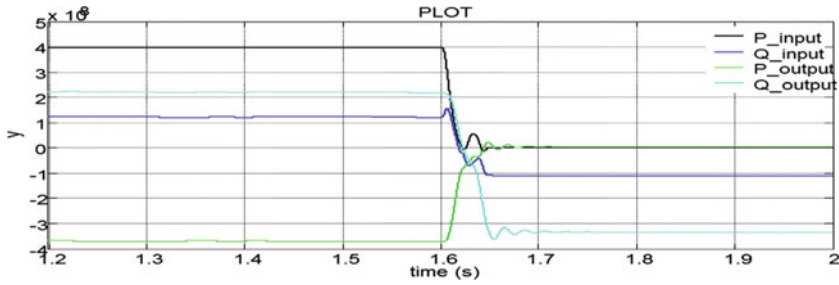
Fig. 18 Active and reactive power of B2B (LCC)

Table 1 Comparison of two technologies (VSC, LCC)

	LCC	VSC
Power reversal	Mechanical switches for voltage reversal as current cannot change direction	Change of current in the converter No need for mechanical switches
AC disturbances and Dynamic	Commutation failure: short circuit of the HVDC Grid for some time Low dynamic performance	No influence other than some loss of active power transfer. Fault ride through capability. High Dynamic performance
AC connection	Limited to medium and high short circuit capacity networks (difficulty at Lao Cai and Ha Giang); Non multiterminal	Connection to electrically weak even "black" AC-network such as Lao Cai and Ha Giang—Multiterminal
AC voltage support	Voltage quality is degraded Dependent P&Q	Provides reactive power support Independent control of P&Q
Black start Cap	No	Yes

VSC has a good reactive power support and voltage control, in particular in weak point as Lao Cai and Ha Giang and the independent control of active and reactive power. In case of grid restoration after a black out, LCC has not black start capacity.

4 Conclusions

The usage of new interconnection links with B2B on 220 kV Xinqiao–Lào Cai and Malutang–Hà Giang transmission lines to replace actual AC meet the following issues:

- Improving reliability, security, political issues and avoid the separations from the National Power System
- Control power flow;
- Allow changing new power purchase agreement.
- Enhanced flexibility for the operation
- Enhance system stability
- Permitting to isolate two power networks when a fault occurs
- Improve power quality for the system.

For a weak area, such as Lào Cai and Hà Giang, the solution of VSC B2B represents several advantages in comparison with LCC B2B and abilities to expand and increase capacity in the future.

References

Elliott D, Bell KRW, Finney SJ, Adapa R, Brozio C, Yu J, Hussain K (2016) A comparison of AC and HVDC options for the connection of offshore wind generation in Great Britain. IEEE Trans Power Deliv 31(2):798–809

Kirby NM (2012) HVDC system solutions. In: Transmission and distribution conference and exposition T&D. IEEE PES, pp 1–3

Kontos E, Pinto RT, Rodrigues S, Bauer P (2015) Impact of HVDC transmission system topology on multiterminal DC network faults. IEEE Trans Power Deliv 30:844–852

Mukhedkar RA, LCC & VSC—Comparison, Astom

Oni OE, Mbangula KI, Davidson IE (2016) A review of LCC-HVDC and VSC-HVDC technologies and applications. Trans Environ Electr Eng 1(3). ISSN 2450-5730

Rudervall R, Charpentier JP, Sharma R (2000) High voltage direct current (HVDC) transmission systems technology review paper

Sellick RL, Åkerberg M (2012) Comparison of HVDC Light (VSC) and HVDC Classic (LCC) Site Aspects, for a 500 MW 400 kV HVDC Transmission Scheme. In: 10th IET international conference on AC and DC power transmission, Dec 2012

Wikipedia, HVDC converters

Wikipedia, High-voltage direct current Siemens, Back to back HVDC

The Bioenergy-Fertilizer Nexus: A Challenge Achievable from Municipal Wastewater

Enrico Camelin, Giulio Cristina, Elisabeth Simelton, Debora Fino, and Tonia Tommasi

Abstract Several billion tons of liquid and solid waste are produced every year globally, contributing to soil and water pollution, greenhouse gas emissions and loss of important nutrients, like C, N and P in the environment. On the other hand, soils are facing a strong reduction of organic matter content while agricultural, food and productive processes mainly rely on external inputs, i.e. mineral fertilizers and chemical pesticides. Biological Wastewater Treatment (WWT) systems play an important role in improving water quality and human health. When properly treated, wastewater is a valuable resource to reduce pollution and combat water scarcity. This chapter aims to give a brief overview of the conventional biological WWT plant and its treatment steps, with a specific focus on the solid residues of WWT plant, i.e. sewage sludge (SS), which could be also valorised by anaerobic digestion (AD), converting it into valuable commercial products: biogas and fertilizers, applying hence the important concept of the Circular Economy. Moreover, we show the effects of application of AD residues (i.e. digestate) on soils with a high risk of desertification, such as with a prevalent sandy texture, with low soil organic matter and high depletion of macronutrients. Results demonstrate fertilizers' effect on poor soils, an issue that is nowadays a persistent environmental problem, with the advancing of desertification and the even more frequent presence of arid soils. The etiological agents of this phenomenon deal with deforestation, urban expansion, unsustainable soil management and climate change, issues that Vietnam and Europe try to solve with different approaches. As result, these soils, even distant, show common features in terms of structure (erosion, compaction), chemical composition (loss of organic carbon, nutrient imbalance) and biodiversity that need to be addressed.

E. Camelin · G. Cristina · D. Fino · T. Tommasi (✉)
Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
e-mail: tonia.tommasi@polito.it

E. Simelton
World Agroforestry (ICRAF), 249A Thuy Khue, Tay Ho, Hanoi, Vietnam

143

1   Introduction

Soil degradation is one of the major threats to global food security. Half of the world's agriculture land is affected by soil degradation. Specifically, 23 hectares per minute are lost to drought and desertification, equal to 12 million hectares per year. This affects the majority of the 2.6 billion people who directly depend on agriculture. In response, the target for Sustainable Development Goal 15.3 sets out to combat desertification and restore degraded land and soil by 2030 (United Nations 2019). Does modern technology allow us to scope for new, safe and circular ways to improve agriculture soils with urban sewage? In Vietnam, wastewater treatment systems are increasingly becoming built into new zones, while in rural areas new laws concern also the waste from a growing number of livestock (Nguyen 2015). In Italy and more generally in Europe, technology for urban wastewater treatments is at the forefront in removing water contaminants.

The overall scope of this chapter is to look for possible areas of knowledge exchange between the EU and Vietnam on the potentials to use sewage sludge for improving degraded agricultural soils.

Sewage sludge (SS) is a by-product from water cleaning processes. Urban wastewater is an effluent water that contains a mixture of residues from human metabolic waste, domestic uses, outdoor run-off and industries. Its main components are settleable solids, dissolved organic compounds, metals and microorganisms, whose concentrations vary considerably depending on season and location. Contaminants need to be removed before the water is returned to rivers, lakes and receiving superficial waters. Therefore, urban wastewater is commonly collected by sewerage systems connected to wastewater treatment plants (WWTP). Here, different cleaning processes take place yielding sewage sludge (SS) as main by-product. The SS is a semi-solid material classified as a putrescible waste, which contains desirable substances, e.g. added value compounds such as plant nutrients, and undesirable ones, e.g. heavy metals and organic pollutants. Principal destinations of SS are definitive disposals, such as in landfilling and incineration, or reuse as biosolids for agriculture. Biosolids derive from SS stabilisation, whereby putrescence and pathogens are reduced by the means of chemical or biological treatments. Within biological treatments, the two most exploited techniques are aerobic treatments and anaerobic digestion (AD). When comparing these two technologies, AD turned out to be more efficient in terms of energy recovery and in providing an effective stabilisation of SS. In fact, the two main products of AD are biogas, a sustainable and renewable source of energy, and anaerobic digestate from sewage sludge (SSAD). The fertilizing effects of SSAD are well known and its use as soil conditioner has been encouraged by many agencies. Nevertheless, direct soil application is strictly regulated to prevent harmful contaminations and to protect soil health (Council of the European Communities 1986). A recent innovative area of research is to test the feasibility of using sewage sludge for soil restoration, especially where desertification and nutrient depletion occur. These advancements could have a great impact on agriculture production and waste management.

2 Wastewater Treatment Process

2.1 Sewage Treatment in Europe and Vietnam

Technologies for disposal of industrial and urban wastewater received little attention until the negative effects of pollutants in aquifers, watercourses and seas were observed on human health and the environment, especially in western countries. Hence, stricter legislation for environment protection across Europe has resulted in more urban wastewater treatment plants (WWTPs) (Council of the European Communities 1991). Table 1 shows the proportion of the population connected to WWTPs, which is higher in northern and central Europe while it reaches only 70% in southern Europe (European Environmental Agency 2017).

The WWTPs in Europe are centralized or decentralized. Centralized WWTPs are more common and receive wastewater from highly developed sewerage systems that may be far away from the plant, and therefore common in densely populated areas. Usually, centralized WWTPs serve a quite large portion of population and exploit a synergic combination of techniques for water cleaning (such as physical and biological treatments). On the other hand, decentralized wastewater systems are more common in rural and remote areas with difficulties to connect to the sewerage network. Normally, they are smaller and less advanced under a technological point of view.

In Vietnam, with a population of nearly 100 million, there were 31 municipal WWTPs and, another 30 under construction or planned by 2018. Some estimates suggest that only 17% of urban wastewater is treated while 79% of wastewater is unsafely disposed of (ARCOWA 2018). The priorities are urban areas, but decentralized WWTP systems in the rural and semi-rural areas are also needed. Rapid urbanization without proper water treatment has resulted in a significant rise in water contamination issues, further aggravated by the monsoonal rains which cause serious flooding of urban sewage systems. To ameliorate the situation, Vietnam invested USD2.1 billion in water sanitation between 1995 and 2009, equal to 0.45% of GDP annually (World Bank 2013). Meanwhile, many efforts have been made to promulgate policies, legislations and regulations on environmental protection (National Assembly of the Socialist Republic of Viet Nam 2005), management, protection and exploitation of water resources (National Assembly of the Socialist Republic of Viet Nam 2012), as well as specific laws on domestic (Ministry of Natural Resources and Environment of Viet Nam 2008) and industrial (Ministry of Natural Resources and Environment of Viet Nam 2011) wastewater treatment.

Decentralised WWTPs are important as a large share of the population lives in rural and semi-rural areas, and a legislation on decentralized WWTP was approved in 2008 (Ministry of Natural Resources and Environment of Viet Nam 2008). Over time, decentralized systems may be included in expanding centralized networks to meet the needs of the rapid urbanization and to reduce the discharge of untreated wastewater. Although as much as 90% of the urban households have septic tanks for collecting domestic wastewater, most of this is disposed directly into drains. In fact,

Table 1 Population connected to wastewater treatment plants in selected European countries

Nation	Population		Population connected to wastewater treatment plants			
			Total connected to wastewater treatment		Urban wastewater collecting system	
	Average	Year	%	Year	%	Year
Albania	2,880,694	2015	19	2015	19	2015
Austria	8,546,356	2014	100	2014	95	2014
Belgium	11,159,407	2013	96	2013	91	2013
Bosnia Herzegovina	3,833,278	2013	n.a	–	35	2013
Bulgaria	7,177,991	2015	87	2015	75	2015
Croatia	4,207,993	2015	98	2015	55	2015
Czechia	10,546,059	2015	81	2015	85	2015
Denmark	5,643,475	2014	100	2014	91	2014
Estonia	1,314,545	2014	87	2014	82	2014
Finland	5,438,972	2013	100	2013	83	2013
France	66,312,067	2014	100	2014	82	2014
Germany	81,776,930	2010	100	2010	96	2013
Greece	10,892,413	2014	93	2014	93	2013
Hungary	9,843,028	2015	77	2015	79	2015
Iceland	318,041	2010	n.a	–	91	2010
Ireland	4,657,740	2014	97	2014	69	2014
Italy	59,277,417	2010	n.a	–	94	2009
Kosovo	1,788,274	2015	0.6	2015	54	2015
Latvia	2,012,647	2013	100	2013	71	2013
Lithuania	2,904,910	2015	77	2015	72	2015
Luxembourg	569,604	2015	98	2015	100	2015
Malta	445,053	2015	100	2015	99	2015
Netherlands	16,939,923	2015	100	2015	99	2015
Norway	5,188,607	2015	99	2015	86	2015
Poland	37,986,412	2015	94	2015	73	2015
Portugal	10,573,100	2010	n.a	–	81	2009
Romania	19,815,616	2015	48	2015	48	2015
Serbia	7,095,383	2015	n.a	–	59	2015
Slovakia	5,423,801	2015	n.a	–	65	2015
Slovenia	2,063,531	2015	93	2015	63	2015
Spain	46,480,882	2014	98	2014	97	2014
Sweden	9,696,110	2014	100	2014	87	2014
Switzerland	8,089,346	2013	100	2013	98	2013
Turkey	77,181,884	2014	64	2014	87	2014

(continued)

Table 1 (continued)

Nation	Population		Population connected to wastewater treatment plants			
			Total connected to wastewater treatment		Urban wastewater collecting system	
	Average	Year	%	Year	%	Year
United Kingdom	62,766,365	2010	99	2010	97	2010

Source Eurostat (2019a)

only some septic tanks accumulate sludge for subsequent treatments and in the early 2010s only in 4% of cases the septage was safely disposed (World Bank 2013).

2.2 Wastewater Treatment Technologies

Wastewater treatment plants are engineered to remove bulky and suspended solids and to effectively remove contaminants in the effluent, so that the water is clean enough to be returned to natural water bodies. Specifically, the organic load reduced in the wastewater is described by indicators such as chemical oxygen demand (COD), biological oxygen demand (BOD) and total suspended solid (TSS). The intensity of the wastewater treatment and the design of the plant depends on the contaminant load of the wastewater and the desired quality of the effluent water (Masotti 2012). The principal technologies adopted in the WWTP are described in Fig. 1.

Preliminary treatments

When urban wastewater reaches the WWTP, it first undergoes preliminary treatment. These operations remove solids, oil, soil and sand, to avoid damaging pipelines,
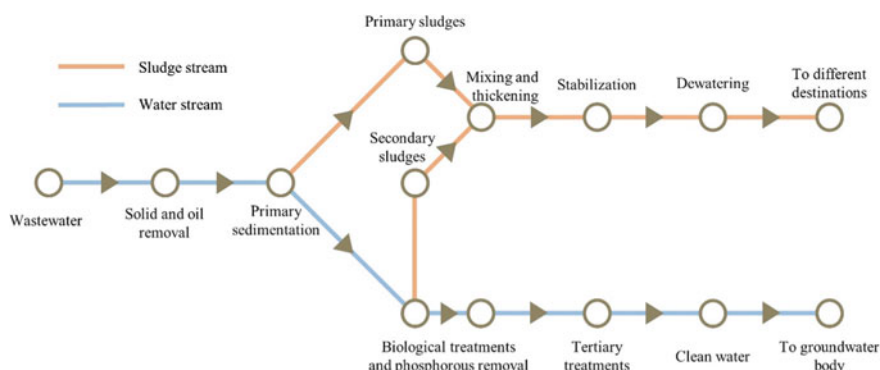

Fig. 1 Schematic of wastewater treatment plant. Water treatment and sludge treatment are depicted in blue and orange, respectively

pumps, and other WWTP equipment. Bulky solids are removed by racks, while smaller ones are retained with screens. Once solids are removed, larger particles can be grinded through different comminutors. Furthermore, grit chambers allow to remove sand and other inert solids by gravity, and oil and grease by flotation (Riffat 2012a).

Primary treatment

Primary treatment reduces 50–70% of the organic solids suspended in the wastewater through sedimentation. Sedimentation tanks are used for *flocculent sedimentation* of solids and flotation of residual oils and greases. Solids are scraped off the bottom of the tanks and recovered as primary sewage sludge, which is then sent to sludge treatment. Sedimentation can be enhanced by adding chemicals that agglomerate particles, which also improves the flocculation performance. This primary treatment reduces BOD by 40% (Riffat 2012b).

Secondary treatment

The secondary treatment further reduces the organic solids (BOD) through biological processes. This allows to get rid of a cocktail of organic molecules (e.g. carbohydrates, proteins, and fats), which are consumed by microorganisms. Here, the concept of *microbial consortium* refers to the symbiosis between a balanced combination of species that allow an efficient food chain in which each species feeds on another. The concerted microbial action results in an effective degradation of organic molecules. Technically, two types of treatment processes are used: *suspended growth* and *attached growth*. In suspended growth processes, the primary effluent is delivered to an aeration tank where organic matter is degraded by a biomass rich in microorganisms, so-called *activated sludge*. As the sludge volume increases, it is sent to a secondary clarifier where it is separated from the water by decantation. The sedimented sludge is then partially recirculated to the aeration tank to maintain the process (i.e. *return activated sludge*) and partially to sludge treatment as secondary sludge (i.e. *waste activated sludge*) (Riffat 2012c). For attached growth processes, trickling filters, biotowers, and rotating biological contactors are widely used. In these processes, microorganisms live in *biofilms* that are immobilised on apposite support to maximize the contact surface with the wastewater, e.g. rocks and gravel. In this case, sludge recirculation is not required, however less sludge is produced, and the BOD abatement may be less efficient. Whichever strategy is adopted, water can be further cleaned with advanced treatments before being returned to groundwater bodies (Riffat 2012d).

Tertiary treatments

Tertiary treatment, or advanced treatment, processes remove residual polluting components, such as nutrients and remaining solids.

Nutrients, such as nitrogen (N) and phosphorous (P), are removed to prevent algae eutrophication of groundwater. Nitrogen is typically removed by a *nitrification-denitrification* process, a biological treatment that often is integrated with the secondary treatment. Nitrogen is mainly present in wastewater as ammonium (NH_4^+),

which first is oxidized to nitrate (NO_3^-) by nitrifying bacteria, then converted by denitrifying bacteria into molecular nitrogen (N_2) and released as gas to the atmosphere.

Phosphorous can be removed through chemical and biological processes. In the chemical process, inorganic salts induce the phosphate precipitation as insoluble salts, for example from ferric chloride to ferric phosphate. On the other hand, biological phosphorous removal is conducted in bioreactors with *phosphorous accumulating organisms* (PAO), which eat and store P in their cellular structures. This bio-inspired approach allows a subsequent recovery of phosphorus, which can be cleverly recycled as fertilizer.

Finally, solids and non-degraded contaminants are removed. Solids are typically removed through filtration, which can be conducted using granular media or membranes, mainly for microfiltration and ultrafiltration. Residual volatile organic compounds and specific ions are removed through carbon activated adsorption and ion exchange resins, respectively (Riffat 2012e).

3 Sewage Sludge: The Main By-Product of WWTPs

Sewage sludge is the principal by-product of wastewater treatment and it originates from the primary and secondary treatments. It is a brownish and smelly slurry, with liquid to semisolid texture. Due to its richness in organic substances, it is a putrescible waste that can make SS even more malodorous as the degradation proceeds. Hence, the further handling of the SS is critical. This section briefly describes the principal techniques of SS dewatering and stabilisation in WWTPs, the main destinations of SS and its chemical profile, focusing on the components relevant for plant nutrition and growth.

3.1 Water Removal from the Sewage Sludge

Water can be removed from SS before or after stabilization. Specifically, it can be conducted as sludge thickening or sludge dewatering. Sludge thickening is used to increase the solid content in SS to make it suitable for further treatment (i.e. stabilization and/or dewatering) and can be done through gravitation, flotation or centrifugation (Gurjar and Tyagi 2017). Dewatering of SS ensures that humidity is reduced to the required levels for SS disposal or reuse. SS dewatering makes mobilisation easier and cheaper, allows SS thermal treatment and stops the decaying process (Campbell and Crescuolo 1982). Technically, SS dewatering can be done through decantation or mechanical processes. Decantation processes are usually slower and suitable for medium-sized WWTPs. Sedimentation takes place in lagoons, drying beds or constructed wetlands with sludge drying reed beds (Uggetti et al. 2009). Mechanical dewatering is often done with devices such as vacuum filters, pressure filter presses

and centrifuges. Prior to mechanical dewatering, two stages are required: thickening, as described above, and conditioning with polyelectrolyte, namely an ionic organic polymer with agglomerating properties. Mechanical dewatering is normally used in larger WWTPs, which handle larger amounts of SS. Regardless of the adopted process, decantation or mechanical one, the resulting dewatered SS is a shovellable solid with a dry matter content ranging between 15 and 30% (Gurjar and Tyagi 2017). Further water removal can be achieved through thermal drying to reach dry matter content up to 95%. This also functions as sterilisation and eliminates almost all viruses and pathogens. Unfortunately, SS drying is an energy consuming process and it is, therefore, not the first-choice dewatering strategy. Instead direct or indirect dryers or combined mode drying systems are more common. Indirect dryers, such as rotary tray dryers, operate at lower temperatures and produce less vapour and therefore are easier to manage. However, they might generate lower dry matter content than direct dryers. Depending on the drying process, dried sewage sludge can take different forms and shapes, such as granular, pelletized, powdery, and beads (Lowe 1995; Chen et al. 2002). All these forms are agronomically interesting since they can have application as soil amendment.

3.2 Sewage Sludge Stabilisation

To remove smell and pathogens from SS in one single step, chemical or biological stabilisation processes are required (Peirce et al. 2007).

Chemical stabilisation

In chemical stabilisation, chemical compounds are mixed with SS. Lime stabilisation is the most common technique, where hydrated lime ($Ca(OH)_2$), quicklime (CaO) or fly ashes are mixed in appropriate amounts (20–30%) to SS in order to reach a pH of 12 or more. The basic environment is favourable for accomplishing the stabilization purposes smoothly. Moreover, the process is exothermic and higher temperatures help to eliminate pathogens. However, the stabilization effect is only temporary and it lasts until the pH drops back to neutral values (Schanke Eikum 1983; Valderrama et al. 2013; Gurjar and Tyagi 2017).

Vermistabilisation

In vermistabilisation, earthworms such as *Eisenia fetida* (red wiggler worm), digest organic matter and have a remediation effect on sewage sludge, reducing the levels of organic and inorganic pollutants, such as polycyclic aromatic hydrocarbons and heavy metals, respectively. The resulting product, called *vermicompost*, has fertilizing properties, such as improved carbon-to-nitrogen ratio and phosphorous plant availability (Suthar 2010; Rorat et al. 2017).

Biological aerobic processes

Biological aerobic processes for SS stabilisation require oxygen and occur through aerobic digestion or composting.

Aerobic digestion is done in digesters, which are fed with SS and air or pure oxygen. During the first phase, biodegradable matter oxidises as microbial consortia, mainly bacteria, feed on organic matter and generate water and carbon dioxide as by-products. Once the "feed" is used up, the second phase (called *endogenous respiration*) starts. Here, microorganisms eat the cell structures of other microorganisms, generating water, carbon dioxide and nitrogen. When the digestion is completed, cell tissues have degraded by 80%, where the remainder is composed of inert compounds and recalcitrant cell components, such as cellulose, which require longer degradation times. The process efficiently reduces the SS volume and pathogens (Zhang et al. 2016; Demirbas et al. 2017; Gurjar and Tyagi 2017).

Composting is a process aimed to stabilise sewage sludge as well as waste from agriculture, food, or gardens. Microbial communities of bacteria, actinomycetes and fungi can reduce up to 30% of the volatile solids. Before composting, SS is dewatered to around 50% dry matter and mixed with a bulking agent, e.g. sawdust or fly ash, to improve aeration. The content is stocked in piles or containers and mixed to ensure proper ventilation. The composting process undergoes three phases: mesophilic (40 °C) and thermophilic (70 °C), where the organic matter is degraded, and the final curing phase, where degradation slow down until it stops. The degradation of organic matter properly happens during the first two phases, while the third is the conclusion of the process. The resulting product is odourless, rich in humic acids and has good soil conditioning properties (Stentiford and de Bertoldi 2010).

Anaerobic digestion

Anaerobic digestion (AD) is a fermentation process to reduce organic matter in absence of oxygen. The technology is often used to manage SS and other types of putrescible waste, such as biomasses from agriculture and food-processing industries as well as the organic fraction of municipal solid waste. AD is the most common SS stabilisation strategy in larger WWTPs. Despite the high initial investments and maintenance costs, this is currently the only technology that allows simultaneous stabilization and energy recovery. The main product is biogas, which is composed of about 60% methane and 30–40% carbon dioxide with minor amounts of nitrogen, hydrogen, siloxanes and hydrogen sulphides. The biogas is combusted in dedicated power stations (in WWTP) for combined generation of heat and electricity (CHP). The AD process can be summarised in three phases:

- *Liquefaction*: extracellular enzymes are synthesized and secreted by hydrogenotrophic bacteria which decompose the organic matter and makes it more soluble;
- *Acidogenic phase*: soluble organic compounds are catabolised by anaerobic microbial consortium composed of acid forming microorganisms, which drop the pH to 5–6.

- *Methanogenic phase*: methane forming bacteria convert organic acids to methane. These anaerobe microorganisms are sensitive to oxygen, pH and temperature variation.

In WWTPs, SS is exploited as pure or mixed (in the case of co-digestion with other kind of waste) fed for anaerobic digesters. Prior to this, the fermentation process is initiated by thickening and warming up the SS. Inside the digester, SS must be kept heated through the mesophilic or thermophilic phases to activate different microorganisms. Mesophilic digestion is performed by microorganisms living at moderate-warm temperatures (~35 °C), while thermophilic one involves microorganisms normally active at higher temperatures (~50 °C). The advantages of thermophilic digestion are higher digestion rates, biogas production, and pathogen removal, while the main drawback is a higher energy demand. Once biogas is produced, it is stored in gasometers which are water-sealed gas holders. Due to the low purity of the methane and presence of corrosive agents, biogas cannot be directly injected into the gas network. Recent efforts to purify biogas into biomethane have resulted in a renewable and high added-value fuel that can be directly injected in gas networks and motor vehicles (Peirce et al. 2007; Gurjar and Tyagi 2017).

The main by-product from this process is sewage sludge anaerobic digestate (SSAD). Once the AD process is over, SSAD is recovered from the digester, thickened and dewatered for further handling. Anaerobic digestates and SSAD are, in the broadest sense, a biomass rich in valuable macro- and micro-nutrients for plant growth. However, these products contain organic and inorganic contaminants, which strongly limit their use for agronomic purposes.

Despite being classified a "product" of sewage treatment in the Water Framework Directive (European Parliament and Council of the European Union 2000), SS is treated and managed as a waste. Therefore, SS must be properly disposed to encourage its recycling. Next, we discuss the final destination and chemical characterization of SSAD.

3.3 Fates of Sewage Sludge

As the WWTPs became more efficient in returning cleaner effluent water to the natural hydrological systems (Council of the European Communities 1991), a rise in SS production was observed as a consequence. Figure 2 gives a general overview of SS production in some selected European countries. The principal SS destinations include agriculture, composting, landfill, incineration and other treatments:

- *Agricultural use*: the use of SS for agricultural purposes has been promoted in Europe since the 1980s (e.g. European Union: Council of European Communities 1986). It is quite common in e.g. Ireland, Spain, Portugal, and regulated by national and international guidelines and policies. Other countries remain reluctant, mainly due to the potentially damaging effects of pollutants on soil health.
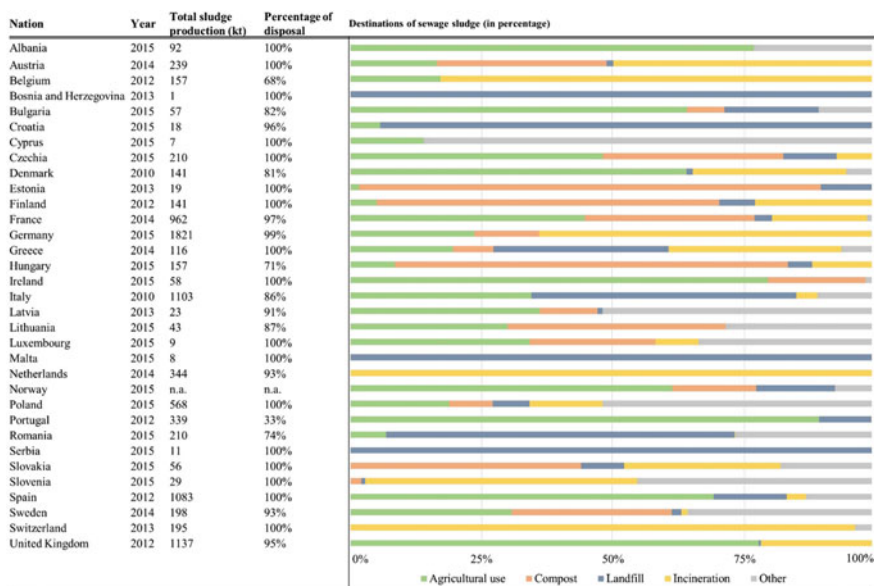
Nation	Year	Total sludge production (kt)	Percentage of disposal	Destinations of sewage sludge (in percentage)
Albania	2015	92	100%	
Austria	2014	239	100%	
Belgium	2012	157	68%	
Bosnia and Herzegovina	2013	1	100%	
Bulgaria	2015	57	82%	
Croatia	2015	18	96%	
Cyprus	2015	7	100%	
Czechia	2015	210	100%	
Denmark	2010	141	81%	
Estonia	2013	19	100%	
Finland	2012	141	100%	
France	2014	962	97%	
Germany	2015	1821	99%	
Greece	2014	116	100%	
Hungary	2015	157	71%	
Ireland	2015	58	100%	
Italy	2010	1103	86%	
Latvia	2013	23	91%	
Lithuania	2015	43	87%	
Luxembourg	2015	9	100%	
Malta	2015	8	100%	
Netherlands	2014	344	93%	
Norway	2015	n.a.	n.a.	
Poland	2015	568	100%	
Portugal	2012	339	33%	
Romania	2015	210	74%	
Serbia	2015	11	100%	
Slovakia	2015	56	100%	
Slovenia	2015	29	100%	
Spain	2012	1083	100%	
Sweden	2014	198	93%	
Switzerland	2013	195	100%	
United Kingdom	2012	1137	95%	

Agricultural use ■ Compost ■ Landfill ■ Incineration ■ Other

Fig. 2 Sewage sludge production and destinations in European countries. *Source* Eurostat (2019b)

- *Composting*: it can be performed together with other feedstocks, i.e.*co-composting*, primarily for agricultural purposes. While other types of stable SS (e.g. SSAD, lime stabilized SS) are used as fertilizers, compost is mainly exploited as soil conditioner (Kacprzak et al. 2017). Composting as SS destination is mostly common in e.g. Finland, Estonia and Hungary.

- *Incineration*: thermal treatments are primarily carried out to recover energy and reduce SS volume. Conventional incineration and co-incineration are heat treatment processes where SS is burnt alone or with e.g. coal, fuel oil or natural gas, returning by-products such as exhaust gases, slag and fly ash. The last ones may be furtherly recovered to produce cementitious materials. Incineration is the only possible SS destination in the Netherlands and Switzerland (Cieślik et al. 2015).

- *Landfilling*: this is an old waste management strategy to store waste on ground. Their negative environmental impacts have become more and more known over the last years, e.g. risk of soil and water contamination from leachate and carbon dioxide emissions (Kacprzak et al. 2017). Furthermore, since no recycling takes place, some countries are reaching a definite space limit (Council of the European Union 1999). Nevertheless, landfills remain the principal disposal solution in many poorer countries, including European countries such as Serbia, Croatia, Romania, and Bosnia and Herzegovina.

- *Other destinations*: other types of thermal treatments include vitrification (above 1000 °C in presence of silica) and pyrolysis (in absence of oxygen). As a matter of fact, these technologies are effective but quite expensive, thus limiting their application (Cieślik et al. 2015).

When it comes to Vietnam, landfill is the most common final destination of SS. The recovery of nutrients is sparsely practiced, while the reuse of untreated wastewater in agriculture and aquaculture is a common practice, especially in rural areas, where there are no WWTPs (ARCOWA 2018).

3.4   Chemical Characterisation and Correlation with Agronomic Features

Sewage sludge contains a range of chemical substances: some are nutrients for plant growth and others are phytotoxic, such as heavy metals. Nutrients in excess can become toxic to plants, while a deficiency usually damages their metabolism and physiological functions.

Table 2 gives an overview of key elements for plant growth that are present in SSADs, from 10 WWTPs in the Mediterranean area, i.e. Spain, Portugal, Italy, France and Greece (Fuentes et al. 2004; Walter et al. 2006; Alvarenga et al. 2007; Tarrasón et al. 2008; Carbonell et al. 2009; De Andres et al. 2010; Ferreiro-Domínguez et al. 2011, 2012; Koutroubas et al. 2014; Cristina et al., 2019).

The plant nutrients in SSADs include both macronutrients and micronutrients, depending on their presence in plant tissues and metabolic needs of the plant.

Macronutrients

The macronutrients present in SSAD include nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg) and sulphur (S) (Jones 2012), and their typical amounts are reported in Table 2. Nitrogen, in its ammonium (NH_4^+) and nitrate (NO_3^-) forms, is involved in essential plant processes and functions, such as amino acid synthesis and protein formation. It is also a fundamental compound of the deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), and present in chlorophyll and B vitamins. In SSADs, nitrogen can be present in high concentrations, which favors the use of SSAD as fertilizer.

Phosphorus is a basic compound of adenosine triphosphate (ATP), which is the most important metabolic "energetic carrier" in plants. Moreover, it is present in nucleic acids, and in enzymes and proteins as a result of post-translational modifications. Potassium frequently features in formulated chemical NPK-fertilizers (Jones 2012), where it regulates the water balance in plants by changing cell turgor pressure and by opening or closing stomata leaf. Lastly, it is essential for the accumulation and translocation of carbohydrates.

Calcium is important for the cell wall structure, cell membrane, cell permeability, and signal transduction. Moreover, it improves pollen germination and growth, and activates the enzymes required for cell mitosis, division, and elongation.

In plant cells, magnesium ions activate enzymes involved in respiration, photosynthesis, and DNA and RNA synthesis. Furthermore, magnesium is a fundamental element in the chlorophyll structure.

Table 2 Range (minimum, maximum and mean values) of SSAD chemical components that are essential for plant growth

Parameter	Unit	Minimum		Maximum		Mean value
		Value	References	Value	References	
O.M.	(% d.m.b.)	37.4	Koutroubas et al. (2014)	63.4	Cristina et al. (2019)	48.9
N	(% d.m.b.)	1.8	Koutroubas et al. (2014)	6.55	Cristina et al. (2019)	3.7
P	(% d.m.b.)	0.2	Koutroubas et al. (2014)	5.7	Cristina et al. (2019)	2.2
K	(% d.m.b.)	0.1	Alvarenga et al. (2007)	1.5	Koutroubas et al. (2014)	0.5
Ca	(g/kg d.m.b.)	3.2	Ferreiro-Domínguez et al. (2011)	82.7	Carbonell et al. (2009)	42.5
Mg	(g/kg d.m.b.)	3.4	Alvarenga et al. (2007)	14.8	Cristina et al. (2019)	8.7
Mn	(g/kg d.m.b.)	0.006	Ferreiro-Domínguez et al. (2012)	0.359	Ferreiro-Domínguez et al. (2011)	0.189
Fe	(g/kg d.m.b.)	13.9	Ferreiro-Domínguez et al. (2011)	33.05	Cristina et al. (2019)	23.7
Cd	(mg/kg d.m.b.)	0.4	Carbonell et al. (2009)	18.3	Fuentes et al. (2004)	4.3
Cr	(mg/kg d.m.b.)	25.5	Walter et al. (2006)	3809	Fuentes et al. (2004)	522.6
Cu	(mg/kg d.m.b.)	142.7	Ferreiro-Domínguez et al. (2012)	749	Tarrasón et al. (2008)	299.5
Hg	(mg/kg d.m.b.)	<0.1	Cristina et al. (2019)	2.6	Tarrasón et al. (2008)	1.2
Ni	(mg/kg d.m.b.)	14.7	Carbonell et al. (2009)	142.8	Ferreiro-Domínguez et al. (2011)	57.3
Pb	(mg/kg d.m.b.)	29	Carbonell et al. (2009)	167	Fuentes et al. (2004)	101
Zn	(mg/kg d.m.b.)	429.5	Carbonell et al. (2009)	7620	Alvarenga et al. (2007)	1480

O.M.: organic matter; d.m.b.: dry matter basis

Sulphur is present in two amino acids of plants, cysteine and methionine, and it is a constituent of several compounds essential for metabolism, such as Coenzyme A and Vitamin B1. Although none of the SSADs in Table 2 reported sulphur, this element is likely present in many SSAD compounds (Dewil et al. 2008).

Micronutrients

Micronutrients, such as boron (B), chlorine (Cl), molybdenum (Mo), iron (Fe), manganese (Mn), copper (Cu) and zinc (Zn), appear in SSADs and are required in relatively low concentrations in plants compared to the macronutrients.

Presence of boron, chlorine and molybdenum in SSAD has been reported, but their quantification is unusual in scientific studies (Epstein 2002). Boron deals with plant cell elongation and nucleic acid synthesis, and its presence in SSADs is certain (Chu and Poon 1999). Chlorine is necessary for photosynthesis reaction. Its presence in SSADs is not confirmed but, if present, it is in low concentrations. Molybdenum is a component of enzymes, such as nitrate reductase, which promotes cell assimilation through the reduction of nitrate to nitrite, and nitrogenase which is responsible of molecular nitrogen fixation.

Manganese ions are present in different enzyme families. It is required by decarboxylases and dehydrogenases taking part in the tricarboxylic acid cycle of the energetic metabolism, and superoxide dismutase, which protects from oxidative damage by reactive oxygen species. Last, manganese appears in a catalytic cluster in the oxygen-evolving complex, which allows the water-splitting activity, that is the first step of the photosynthesis (Nelson and Cox 2013).

Iron is a component of chemical groups, such as heme groups, associated with proteins like cytochromes. For instance, cytochrome $b_6 f$ complex is involved in electron transfer of reactions occurring during the photosynthesis (Willows 2007).

Copper and zinc are both considered micronutrients at low concentrations and contaminants when in excess. Similarly to iron, copper is associated with enzymes that are responsible of redox reactions, while zinc is as cofactor of other several enzymes. Due to the high amount of these elements in sewage sludge (Table 2), the European Union has limited their concentration in sewage sludge for agricultural use (Council of the European Communities 1986).

Organic and inorganic compounds

"Organic compounds" broadly include all substances with organic carbon (C_{Org}). In these compounds, the carbon atom creates chemical bonds especially with other C_{Org} atoms, hydrogen, oxygen, nitrogen, phosphorous and halogens. The two main types of organic compounds in SSAD are organic matter and organic pollutants, which have opposite agronomic effects.

SSAD dry matter usually consist of around 35–40% organic carbon. Its functional properties as soil amendment include acting as a depository for nutrients, reducing soil compaction, enhancing micro and macropores, increasing the microbial population and activity, rising cation exchange capacity (CEC) and soil water retention, which ameliorate the root environment and water uptake (Jones et al. 2005).

The downside of SSADs is the presence of toxic organic and inorganic pollutants. A study on sewage sludges in China found thirteen different categories of organic pollutants such as phthalate esters, antibiotics, polycyclic aromatic hydrocarbons, pharmaceuticals and hormones (Meng et al. 2016). Some of these compounds can be degraded with treatments, while others require a longer time, so that, when they

accumulate in soils, they may have adverse effects at various trophic levels (Kolpin et al. 2002; Harrison et al. 2006; Díaz-Cruz et al. 2009).

As regards inorganic pollutants, SSADs are strictly monitored in terms of heavy metals, such as cadmium (Cd), chromium (Cr), copper (Cu), mercury (Hg), nickel (Ni), lead (Pb) and zinc (Zn). The threshold levels for heavy metals are strongly regulated on SSAD for agriculture purposes (Council of the European Communities 1986). Example values are summarized in Table 2. The effects of heavy metals are various and complex. Whether they accumulate or become bioavailable depend on e.g. soil pH, oxidation number, and presence of other chemicals. When absorbed by plants at low levels, simple bioaccumulation occurs and often without any specific effect. As the concentration of heavy metals increases, phytotoxic effects emerge, such as growth retardation and inhibition of iron translocation (caused by excess of copper, nickel and zinc), and reduced root development (due to hexavalent chromium). Furthermore, cadmium may result in pulmonary diseases and kidney dysfunctions on humans and animals eventually eating the metal-exposed plants.

4 Soil Application of Sewage Sludge

4.1 Sewage Sludge as Fertilizer

The legislation on SS and SSAD for soil application is primarily targeted to prevent soil contamination e.g. in Europe the Council Directive 86/278 (1986) and in the USA 40 CFR Part 503 of the Environmental Protection Agency (2018). In agriculture the regulation on SS depends on: i) the quantity of nutrients, heavy metals and organic contaminants in the SS, ii) the characteristics of the soil in the disposal area, iii) the number of human pathogens in the SS, e.g. *Salmonella* spp. and *Escherichia coli*, and iv) the kind of treatment of sewage sludge.

Depending on the degree of dewatering, SS can be applied to soils in solid, semisolid or liquid form, through incorporation into the soil, injection below the soil surface, or spraying or spreading onto the soil surface (Zain et al. 2002). The former two are the preferred onces to reduce smells and odours.

4.2 The New Potential: Using Sewage Sludge for Soil Improvement

The benefits on plant growth from animal manure and human metabolic waste have been known for millennia and some properties were sought to reproduce in the emergence of synthetic fertilizers (Adjei and Rechcigl 2002). Many scientific works dealt with this topic, carrying out trials in greenhouse (Perez-Murcia et al. 2006) or in open field (Singh and Agrawal 2010). Among the pioneers, Bartow and Hatfield

in 1916 (Bartow and Hatfield 1916) proved that the biomass yields of lettuce and radish grown were higher on SS amended soils compared to an untreated one. Since then, research has established types and dosages of SS for different soils and plants species, e.g. barley (Antolín et al. 2005), kenaf (De Andres et al. 2010) and wheat (Koutroubas et al. 2014).

Further, toxic effects of heavy metals can be minimized on alkaline soils. For example, Cu, Cd, Mn, Ni and Zn are less bioavailable for plant absorption at basic soil pH compared to acid soil conditions (Navas et al. 1998; García-Gil et al. 2004; Antolín et al. 2005; Healy et al. 2016). Moreover, organic matter can absorb heavy metals minimizing their toxic effects. However, as the organic matter gradually mineralizes, a phenomenon called "sludge time bomb" may occur. Metals may be released into more bioavailable forms, which means that plants could uptake the metals and might suffer from phytotoxicity (McBride 1995; Chang et al. 1997; Parat et al. 2005; Mosquera-Losada et al. 2017). Moreover, SSAD application on clayey soils, composed by Na- or Ca-bentonite, may result in heavy metals absorption on to the clay materials, restricting their mobility (Usman et al. 2005).

Another example of the positive effect of SSAD application is in the reclamation and restoration of disturbed or contaminated land. Sopper (1993) showed that the application of SSAD on old mines improved the soil physical properties and released plant nutrients, which speeded up the revegetation in the heavily degraded landscapes. Similarly, the regrowth after the application of different doses of SSAD on a degraded semiarid soil in central Spain significantly increased within one year, as measured by total plant cover and total biomass yield (Walter et al. 2000). Moreover, a recent work demonstrated how SSAD can have a beneficial effect on plants grown on a soil really poor in organic matter and nutrients. Indeed, these plants significantly improved biomass production and physiological parameters (Cristina et al. 2019).

In summary, SS has benefits for soils and plants, especially on poor, disturbed, alkaline and clayey soils.

5 Soil Improvement for Sustainable Agriculture

5.1 Sustainable Agriculture Concepts

The call for 'sustainable agriculture' practices has emerged from evidence of the adverse effects on ecosystem functions associated with monoculture farming and agricultural malpractices, such as intensive use of chemical inputs (Pretty 2007). Sustainable agriculture aims to balance trade-offs between social, economic or environmental interests (Box 1), which has resulted in overlaps with e.g. agroecology (De Schutter and Vanloqueren 2011), climate-smart agriculture (FAO 2013), and nature-based solutions (Sonneveld et al. 2018), where similar principles are found. The use of nature-based solutions such as bioremediation and phytoremediation in agriculture has still limited use although these have important bearing on improving

the capacity of agriculture soils to buffer biotic and abiotic stress (see Principles 1 and 2 in Box 1).

Traditionally, such ameliorating farming practices include e.g. minimum tillage, catch crops, cover crops, green mulch and compost (Lal 2011). Furthermore, the potential roles of sewage sludge land application are heavily regulated, such as in Europe, while the technology is costly, especially in developing countries such as Vietnam.

> **Box 1. Five Principles of Sustainable Agriculture**
>
> *Principle 1.* Improving efficiency in the use of resources
> *Principle 2.* Direct action to conserve, protect and enhance natural resources
> *Principle 3.* Protect and improve rural livelihoods, equity and social well-being
> *Principle 4.* Enhanced resilience of people, communities and ecosystems
> *Principle 5.* Responsible and effective governance mechanisms
> *Source* FAO (2014).

5.2 Sustainable Agriculture and Policies on the Use of Agriculture Waste in Vietnam

A small share of urban waste in Vietnam is treated and most of agriculture originates from small rural farms. This section therefore briefly covers agricultural waste.

In Vietnamese policies, 'sustainable agriculture' has featured for decades. It first appeared after the economic opening policies 'doi moi' in the 1980s, when the country's ambition was to increase food productivity and ensure stable livelihood improvements for farmers (De Luu and Trường 2016).

As the country transitioned from low to lower-middle income country in 2011, agriculture development has ventured onto a path from food security to food safety issues, not the least for ensuring export quality to the EU (Nguyen-Viet et al. 2017). The Government has also aligned with international concepts and policies (for an overview see https://www.climatechange.vn/en/). One aspect of this can be seen in the increase in livestock. Between 2000 and 2017, the number of pigs increased from 20 to 27 million, and cattle from 4.1 to 5.6 million heads.[1] Although the 2018 Law on Livestock Waste Treatment stipulated the responsibility of farm owners to process waste, and the Ministry of Natural Resources and the Environment (2016) set thresholds for chemical substances introduced in agriculture, there is still no technical regulation or guideline from the Ministry of Agriculture and Rural Development (MARD) on how farmers can treat or use wastewater for irrigation or composting.

---

[1]Vietnam General Statistics Office 2019 https://www.gso.gov.vn/default_en.aspx?tabid=778.

Policies on the use of effluent of livestock for crop production are expected to be released by MARD in 2019.

Two common farm by-products, the slurry after on-farm biogas production and the use of microorganisms, could theoretically enable farmers to recycle farm waste with sufficiently high nutrient content, while minimizing the expenditures on inorganic fertilizers (Principle 1, Box 1). Farmers are demotivated to install wastewater treatment as it is too costly and if they follow the QCVN62-standards the treated water will lack value as fertilizer. Larger industrial farms could introduce bioremediation, filtration lakes and sedimentation systems (Nguyen 2015). Since these systems also contribute to greenhouse gas emission reductions, there are opportunities to support implementation through carbon credit mechanisms.

Some small-scale examples of improved manure treatment for soil improvement in Vietnam include compost and vermiculture (Le and Simelton 2018). The adoption of these practices is attributed to fast results, low input costs, increased resource use efficiency of on-farm agriculture waste, and the demand originating from a bottom-up community learning context rather than through top-down policies (Simelton et al. 2018). The perspective of these interventions as prototypes can encourage the next generation of farm residue management to develop, since more awareness is raised, more affordable technology becomes available, and enabling policies are implemented (Principle 3 and 4, Box 1).

In this regard, the agriculture research collaboration between Italy and Vietnam offers opportunities to exchange experiences on soil improvement and safe agriculture production technologies.

6 Next Steps for Advancing the Land Application of Sewage Sludge

This chapter has highlighted two important benefits of SSAD: the reuse of a waste and its effect as soil improver. In fact, the common final destination of SSAD as landfilling is not a sustainable solution (McDonald et al. 2010). Moreover, SS contains macro- and micro-nutrients and organic matter, which are useful substances for plant growth. Today, fertilizers are to a large extent derived from non-renewable resources, such as mines, at a high environmental cost. Thus, SSAD makes a perfect example of circular economy, a: "[…] concept which conceives of a production and consumption system with minimal losses of materials and energy through extensive reuse, recycling, and recovery" (Haupt et al. 2017).

Nevertheless, the possible presence of pathogens, and organic and inorganic contaminants in SS can be dangerous for humans, animals, plants and the environment. Several technologies can be adopted, for instance, ionic extraction, heat treatment or vermicomposting can reduce contaminants and heavy metal content (Camargo et al. 2016). Future research is needed to determine where AD is

the preferred SS stabilization technique, and to maximise the value of SSAD and the biogas yields.

To advance the use of SSAD as sustainable practice, advanced technologies and methods are needed to improve the extraction of compounds with high fertilizing qualities from SS and SSAD. For example, phosphorus can be extracted from SS and SSAD using acid washing and alkali extraction, as well as electrodialysis, which also can remove heavy metals (Ciešlik et al. 2015). Another promising technology is struvite precipitation. Struvite is a mineral composed by ammonia, phosphate and magnesium ($NH_4MgPO_4 \times 6H_2O$) with excellent fertilizing properties. Struvite can be obtained by mixing selected chemical compounds (e.g. magnesium chloride) in basic conditions into mixtures rich in ammonia and phosphate, such as SS and SSAD. Through struvite precipitation, two macronutrients are recovered simultaneously and directly formulated as fertilizer (Yu et al. 2017). Moreover, rare metals can be removed by incineration with plasma furnaces (e.g. silver, tellurium, thallium, bismuth, antimony, indium, gallium, tin, germanium, and lead) and used in industrial processes (Ciešlik et al. 2015). There are opportunities to further explore cost-share models for these processes.

Moreover, biogas productivity through AD can be advanced by reducing by-products, such SSAD. One of the most promising technologies is thermal hydrolysis, a pretreatment of sewage sludge that increases biogas production while reducing pathogens and some organic pollutants (Taboada-Santos et al. 2019).

These technologies have been designed and applied in some WWTPs. Two leading examples for the evolution of WWTP schemes include one in Athens (Greece), which uses thermal hydrolyzation of SS before anaerobic digestion, and a WWTP in Carbonera (Italy), which is equipped with a biological phosphorous removal system that allows P-recovery as struvite. The WWTP concept is gradually changing from treatment plant to biorefinery, "a network of facilities that integrates biomass conversion processes and equipment to produce biofuels, energy and chemicals from biomass" (Moncada et al. 2016). This means that the future of WWTPs is not only in wastewater purification and waste management, but also a model for circular economy and sustainable agriculture.

To set this path in motion, collaboration between Italy and Vietnam on sustainable waste management and agriculture practices could include: (i) developing standards and instructions for the use of agricultural and urban waste (SS, SSAD) as soil amendments and fertilizers in agriculture, (ii) collaboration on the policy process to evaluate national and transnational regulations and to implement controls for monitoring substances in soils and food, (iii) planning for a series of small and medium sized rural WWTPs closer to farms and rural towns, (iv) research collaboration on measures for frequent testing of soils to establish what nutrient status and to ensure correct soil improvement formulas for soil and crops, (v) developing a database of remediation functions of various plants and combinations of plants with economic and non-economic evaluations of using SSAD in the circular value chain of waste management-soil improvement-food/fodder production-waste management.

References

Adjei MB, Rechcigl JE (2002) Bahiagrass production and nutritive value as affected by domestic wastewater residuals. Agron J 94:1400–1410. https://doi.org/10.2134/agronj2002.1400

Alvarenga P, Palma P, Gonçalves AP et al (2007) Evaluation of chemical and ecotoxicological characteristics of biodegradable organic residues for application to agricultural land. Environ Int 33(4): 505–513. https://doi.org/10.1016/j.envint.2006.11.006

Antolín MC, Pascual I, García C et al (2005) Growth, yield and solute content of barley in soils treated with sewage sludge under semiarid Mediterranean conditions. Field Crops Res 94:224–237. https://doi.org/10.1016/j.fcr.2005.01.009

ARCOWA (2018) Wastewater management and resource recovery in Vietnam: current status and opportunities. http://seaknowledgebank.net/sites/default/files/wastewater-management-and-resource-recovery-in-VietNam.pdf

Bartow E, Hatfield WD (1916) Fertilizer value of activated sludge. Ind Eng Chem 8(1):17–20. https://doi.org/10.1021/i500001a004

Camargo FP, Sérgio Tonello P, Dos Santos ACA, Duarte ICS (2016) Removal of toxic metals from sewage sludge through chemical, physical, and biological treatments—a review. Water Air Soil Pollut 227(12):433. https://doi.org/10.1007/s11270-016-3141-3

Campbell HW, Crescuolo PJ (1982) The use of rheology for sludge characterization. Water Sci Technol 14(6-7):475–489. https://doi.org/10.2166/wst.1982.0120

Carbonell G, Pro J, Gómez N et al (2009) Sewage sludge applied to agricultural soil: ecotoxicological effects on representative soil organisms. Ecotoxicol Environ Saf 72(4):1309–1319. https://doi.org/10.1016/j.ecoenv.2009.01.007

Chang AC, Hyun H-N, Page AL (1997) Cadmium uptake for Swiss chard grown on composted sewage sludge treated field plots: plateau or time bomb? J Environ Qual 26(1):11–19. https://doi.org/10.2134/jeq1997.00472425002600010003x

Chen G, Yue PL, Mujumdar AS (2002) Sludge dewatering and drying. Drying Technol 20(4–5): 883–916. https://doi.org/10.1081/DRT-120003768

Chu CW, Poon CS (1999) The feasibility of planting on stabilized sludge-amended soil. Environ Int 25(4):465–477. https://doi.org/10.1016/S0160-4120(99)00009-4

Cieślik BM, Namieśnik J, Konieczka P (2015) Review of sewage sludge management: standards, regulations and analytical methods. J Cleaner Prod 90:1–15. https://doi.org/10.1016/j.jclepro.2014.11.031

Council of the European Communities (1986) Council Directive 86/278/EEC of 12 June 1986 on the protection of the environment, and in particular of the soil, when sewage sludge is used in agriculture. Official J Eur Communities L 181:6–12

Council of the European Communities (1991) Council Directive 91/271/EEC of 21 May 1991 concerning urban waste-water treatment. Official J Eur Communities L 135:40–52

Council of the European Union (1999) Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. Official J Eur Communities L 182:1–19

Cristina G, Camelin E, Pugliese M et al (2019) Evaluation of anaerobic digestates from sewage sludge as a potential solution for improvement of soil fertility. Waste Manage 99:122–134. https://doi.org/10.1016/j.wasman.2019.08.018

De Andres EF, Tenorio JL, Walter I (2010) Biomass production and nutrient concentration of kenaf grown on sewage sludge-amended soil. Span J Agric Res 8:472–480. https://doi.org/10.5424/sjar/2010082-1202

De Luu T, Trường D (2016) Phát triển nông nghiệp bền vững ở việt nam trong bối cảnh mới của hội nhập kinh tế quốc tế. In: Kỷ yếu hội thảΩo quốc tế Kinh tế Việt Nam thời kỳ hội nhập: Cơ hội và thách thức. Huế, Đại học Kinh tế Huế: NXB. Hồng Đức

Demirbas A, Coban V, Taylan O, Kabli M (2017) Aerobic digestion of sewage sludge for waste treatment. Energy Sources Part A Recovery Utilization Environ Eff. 39:1056–1062. https://doi.org/10.1080/15567036.2017.1289282

De Schutter O, Vanloqueren G (2011) The new green revolution: how twenty-first-century science can feed the world. Solutions 2(4):33–44. https://www.thesolutionsjournal.com/article/the-new-green-revolution-how-twenty-first-century-science-can-feed-the-world/

Dewil R, Baeyens J, Roels J, Van De SB (2008) Distribution of sulphur compounds in sewage sludge treatment 25(6):879–886. Environ Eng Sci https://doi.org/10.1089/ees.2007.0143

Díaz-Cruz MS, García-Galán MJ, Guerra P et al (2009) Analysis of selected emerging contaminants in sewage sludge. TrAC Trends Anal Chem 28(11):1263–1275. https://doi.org/10.1016/j.trac.2009.09.003

Environmental Protection Agency (2018) Code of Federal Regulations, Title 40 Part 503—Standards for the use or disposal of sewage sludge. https://www.govinfo.gov/content/pkg/CFR-2018-title40-vol32/xml/CFR-2018-title40-vol32-part503.xml

Epstein E (2002) Land application of sewage sludge and biosolids. CRC Press, Boca Raton. ISBN 9780367454746.

European Environmental Agency (2017) Urban wastewater treatment—assessment 4. https://www.eea.europa.eu/data-and-maps/indicators/urban-waste-water-treatment/urban-waste-water-treatment-assessment-4

European Parliament and Council of the European Union (2000) Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Official J Eur Communities L 327:1–73.

Eurostat (2019a) Population connected to wastewater treatment plants. Database update: 19th November 2018; Accessed on 30th July 2019. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_ww_con&lang=en

Eurostat (2019b) Sewage sludge production and disposal. Database update: 6th August 2018; Accessed on 10th August 2019. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_ww_spd&lang=en

FAO (2013) Climate-smart agriculture—sourcebook. Rome. ISBN 978-92-5-107720-7. http://www.fao.org/climate-smart-agriculture-sourcebook/about/en/

FAO (2014) Building a common vision for sustainable food and agriculture—principles and approaches. Rome. ISBN 978-92-5-108471-7. http://www.fao.org/policy-support/tools-and-publications/resources-details/en/c/418447/

Ferreiro-Domínguez N, Rigueiro-Rodríguez A, Mosquera-Losada MR (2011) Response to sewage sludge fertilisation in a *Quercus rubra* L. silvopastoral system: soil, plant biodiversity and tree and pasture production. Agric Ecosyst Environ 141(1–2):49–57. https://doi.org/10.1016/j.agee.2011.02.009

Ferreiro-Domínguez N, Rigueiro-Rodríguez A, Mosquera-Losada MR (2012) Sewage sludge fertiliser use: implications for soil and plant copper evolution in forest and agronomic soils. Sci Total Environ 424:39–47. https://doi.org/10.1016/j.scitotenv.2012.02.042

Fuentes A, Lloréns M, Sáez J et al (2004) Phytotoxicity and heavy metals speciation of stabilised sewage sludges. J Hazard Mater 108(3):161–169. https://doi.org/10.1016/j.jhazmat.2004.02.014

García-Gil JC, Plaza C, Senesi N et al (2004) Effects of sewage sludge amendment on humic acids and microbiological properties of a semiarid Mediterranean soil. Biol Fertil Soils 39(5):320–328. https://doi.org/10.1007/s00374-003-0709-z

Gurjar BR, Tyagi VK (2017) Sludge management. CRC Press, Boca Raton. https://doi.org/10.1201/9781315375137

Harrison EZ, Oakes SR, Hysell M, Hay A (2006) Organic chemicals in sewage sludges. Sci Total Environ 367(2–3):481–97. https://doi.org/10.1016/j.scitotenv.2006.04.002

Haupt M, Vadenbo C, Hellweg S (2017) Do we have the right performance indicators for the circular economy? Insight into the Swiss waste management system. J Ind Ecol 21(3):615–627. https://doi.org/10.1111/jiec.12506

Healy MG, Ryan PC, Fenton O et al (2016) Bioaccumulation of metals in ryegrass (*Lolium perenne* L.) following the application of lime stabilised, thermally dried and anaerobically digested sewage sludge. Ecotoxicol Environ Saf 130:303–309. https://doi.org/10.1016/j.ecoenv.2016.04.026

Jones JB Jr (2012) Plant nutrition and soil fertility manual - Second Edition. CRC Press. ISBN 9781439816097

Jones A, Montanarella L, Jones R (2005) Soil atlas of Europe. European Commission. Luxembourg. ISBN 928948120X.

Kacprzak M, Neczaj E, Fijałkowski K et al (2017) Sewage sludge disposal strategies for sustainable development. Environ Res 156:39–46. https://doi.org/10.1016/j.envres.2017.03.010

Kolpin DW, Furlong ET, Meyer MT et al (2002) Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000: a national reconnaissance. Environ Sci Technol 36(6):1202–1211. https://doi.org/10.1021/es011055j

Koutroubas SD, Antoniadis V, Fotiadis S, Damalas CA (2014) Growth, grain yield and nitrogen use efficiency of Mediterranean wheat in soils amended with municipal sewage sludge. Nutr Cycl Agroecosyst 100(2):227–243. https://doi.org/10.1007/s10705-014-9641-x

Lal R (2011) Sequestering carbon in soils of agro-ecosystems. Food Policy 36(1):S33–S39. https://doi.org/10.1016/j.foodpol.2010.12.001

Le TT, Simelton E (2018) Portfolio of climate-smart agriculture practices for scaling. No. 1. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Wageningen, The Netherlands. https://ccafs.cgiar.org/publications/portfolio-csa-practices-scaling#.X382VGgzbIU

Lowe P (1995) Developments in the thermal drying of sewage sludge. Water Environ J 9(3):306–316. https://doi.org/10.1111/j.1747-6593.1995.tb00944.x

Masotti L (2012) Depurazione delle acque. Tecniche ed impianti per il trattamento delle acque di rifiuto. Calderini (ed). ISBN 885065202X.

McBride MB (1995) Toxic metal accumulation from agricultural use of sludge: are USEPA regulations protective? J Environ Qual 24:5–18. https://doi.org/10.2134/jeq1995.00472425002400010002x

McDonald RI, Forman RTT, Kareiva P (2010) Open space loss and land inequality in United States' cities, 1990–2000. PLoS ONE 5(3):e9509. https://doi.org/10.1371/journal.pone.0009509

Meng XZ, Venkatesan AK, Ni YL et al (2016) Organic contaminants in chinese sewage sludge: a meta-analysis of the literature of the past 30 years. Environ Sci Technol 50(11):5454–5466. https://doi.org/10.1021/acs.est.5b05583

Ministry of Natural Resources and Environment of Viet Nam (2011) National technical regulation on industrial wastewater - QCVN 40:2011/BTNMT. https://circabc.europa.eu/sd/a/97d7bd92-e92e-4b42-a2b3-6a1d8d00d81d/National%20Technical%20Regulation%20%20on%20Industrial%20Wastewater.pdf

Ministry of Natural Resources and Environment of Viet Nam (2008) National technical regulation on domestic wastewater - QCVN 14 : 2008/BTNMT. https://qcvn.com.vn/en/sample-submission-information/laws-regulations-regulations/

Ministry of Natural Resources and Environment of Viet Nam (2016) National technical regulation on the effluent of livestock - QCVN62-MT:2016/BTNMT. http://www.gree-vn.com/pdf/QCVN-62-MT-2016-BTNMT-nuoc-thai-chan-nuoi.pdf

Moncada JB, Aristizábal VM, Cardona CA (2016) Design strategies for sustainable biorefineries. Biochem Eng J 116:122–134. https://doi.org/10.1016/j.bej.2016.06.009

Mosquera-Losada R, Amador-García A, Muñóz-Ferreiro N et al (2017) Sustainable use of sewage sludge in acid soils within a circular economy perspective. CATENA 149:341–348. https://doi.org/10.1016/j.catena.2016.10.007

National Assembly of the Socialist Republic of Viet Nam (2005) Law No. 52/2005/QH11 on environmental protection. http://www.ilo.org/dyn/natlex/docs/ELECTRONIC/73671/125789/F-1191853968/VNM73671%20Eng.pdf

National Assembly of the Socialist Republic of Viet Nam (2012) Law No. 17/2012/QH13 on water resources. https://vanbanphapluat.co/law-no-17-2012-qh13-on-water-resources

Navas A, Bermúdez F, MacHín J (1998) Influence of sewage sludge application on physical and chemical properties of Gypsisols. Geoderma 87(1–2):123–135. https://doi.org/10.1016/S0016-7061(98)00072-X

Nelson DL, Cox MM (2013) Oxidative phosphorylation and photophosphorylation. In: Lehninger principles of biochemistry, 6th edn. W. H. Freeman & Co Ltd, New York, pp 731–798.

Nguyen V-A (2015) Wastewater management and technology needs in Vietnam. World Intellectual Property Organization. https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_mnl_15/wipo_ip_mnl_15_t12.pdf

Nguyen-Viet H, Tuyet-Hanh TT, Unger F et al (2017) Food safety in Vietnam: where we are at and what we can learn from international experiences. Infect Dis Poverty 6:39. https://doi.org/10.1186/s40249-017-0249-7

Parat C, Chaussod R, Lévêque J, Andreux F (2005) Long-term effects of metal-containing farmyard manure and sewage sludge on soil organic matter in a fluvisol. Soil Biol Biochem 37(4):673–679. https://doi.org/10.1016/j.soilbio.2004.08.025

Peirce JJ, Weiner RF, Vesilind PA (1998) Chapter 9 - Sludge Treatment, Utilization, and Disposal. In: Environmental Pollution and Control (Fourth Edition), pp 125–135. https://doi.org/10.1016/B978-075069899-3/50010-9

Perez-Murcia MD, Moral R, Moreno-Caselles J et al (2006) Use of composted sewage sludge in growth media for broccoli. Biores Technol 97:123–130. https://doi.org/10.1016/j.biortech.2005.02.005

Pretty J (2007) Agricultural sustainability: concepts, principles and evidence. Philos Trans R Soc B Biol Sci 363:447–465. https://doi.org/10.1098/rstb.2007.2163

Riffat R (2012a) Preliminary treatment. In: Fundamentals of wastewater treatment and engineering, pp. 85–98. CRC Press, Boca Raton. https://doi.org/10.1201/b12746

Riffat R (2012b) Primary treatment. In: Fundamentals of wastewater treatment and engineering, pp. 99–118. CRC Press, Boca Raton. https://doi.org/10.1201/b12746

Riffat R (2012c) Secondary treatment: suspended growth processes. In: Fundamentals of wastewater treatment and engineering, pp 119–164. CRC Press, Boca Raton. https://doi.org/10.1201/b12746

Riffat R (2012d) Secondary treatment: attached growth and combined processes. In: Fundamentals of wastewater treatment and engineering, pp. 165–187. CRC Press, Boca Raton. https://doi.org/10.1201/b12746

Riffat R (2012e) Advanced treatment processes. In: Fundamentals of wastewater treatment and engineering, pp 287–322. CRC Press, Boca Raton. https://doi.org/10.1201/b12746

Rorat A, Wloka D, Grobelak A et al (2017) Vermiremediation of polycyclic aromatic hydrocarbons and heavy metals in sewage sludge composting process. J Environ Manage 187:347–353. https://doi.org/10.1016/j.jenvman.2016.10.062

Schanke Eikum A (1983) Lime stabilisation of sewage sludges. In: Carberry JB, Englande AJ (eds.) Sludge characteristics and behavior, pp 359–378. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-009-6860-8

Simelton E, Gammelgaard J, Le TT (2018) Guide for impact assessment of agro-climate information services. CCAFS working paper no. 242. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Wageningen, The Netherlands. https://ccafs.cgiar.org/publications/guide-impact-assessment-agro-climate-information-services#.X38tO2gzbIU

Singh RP, Agrawal M (2010) Biochemical and physiological responses of rice (Oryza sativa L.) grown on different sewage sludge amendments rates. Bull Environ Contam Toxicol 84(5):606–612. https://doi.org/10.1007/s00128-010-0007-z

Sonneveld BGJS, Merbis MD, Alfarra A et al (2018) Nature-based solutions for agricultural water management and food security. FAO Land and Water Discussion Paper no. 12. Food and Agriculture Organization of United Nations (FAO), Rome. http://www.fao.org/3/CA2525EN/ca2525en.pdf

Sopper WE (1993) Municipal sludge use in land reclamation. CRC Press, Boca Raton. ISBN 9780873719414.

Stentiford E, de Bertoldi M (2010) Composting: process. In: Solid waste technology & management, Christensen TH (ed.) pp 513–532, Blackwell Publishing Ltd. https://doi.org/10.1002/9780470666883.ch34

Suthar S (2010) Pilot-scale vermireactors for sewage sludge stabilization and metal remediation process: comparison with small-scale vermireactors. Ecol Eng 36:703–712. https://doi.org/10.1016/j.ecoleng.2009.12.016

Taboada-Santos A, Braz GHR, Fernandez-Gonzalez N et al (2019) Thermal hydrolysis of sewage sludge partially removes organic micropollutants but does not enhance their anaerobic biotransformation. Sci Total Environ 690:534–542. https://doi.org/10.1016/j.scitotenv.2019.06.492

Tarrasón D, Ojeda G, Ortiz O, Alcañiz JM (2008) Differences on nitrogen availability in a soil amended with fresh, composted and thermally-dried sewage sludge. Biores Technol 99(2):252–259. https://doi.org/10.1016/j.biortech.2006.12.023

Uggetti E, Llorens E, Pedescoll A et al (2009) Sludge dewatering and stabilization in drying reed beds: characterization of three full-scale systems in Catalonia, Spain. Biores Technol 100:3882–3890. https://doi.org/10.1016/j.biortech.2009.03.047

United Nations (2019) Sustainably manage forests, combat desertification, halt and reverse land degradation, halt biodiversity loss (15 - Life on Land). Sustainable Development Goals. https://www.un.org/sustainabledevelopment/biodiversity/

Usman A, Kuzyakov Y, Stahr K (2005) Effect of clay minerals on immobilization of heavy metals and microbial activity in a sewage sludge-contaminated soil. J Soils Sediments 5(4):245–252. https://doi.org/10.1065/jss2005.05.141

Valderrama C, Granados R, Cortina JL (2013) Stabilisation of dewatered domestic sewage sludge by lime addition as raw material for the cement industry: understanding process and reactor performance. Chem Eng J 232:458–467. https://doi.org/10.1016/j.cej.2013.07.104

Walter I, Cuevas G, García S, Martínez F (2000) Biosolid effects on soil and native plant production in a degraded semiarid ecosystem in central Spain. Waste Manage Res 18(3):259–263. https://doi.org/10.1034/j.1399-3070.2000.00111.x

Walter I, Martínez F, Cala V (2006) Heavy metal speciation and phytotoxic effects of three representative sewage sludges for agricultural uses. Environ Pollut 139(3):507–514. https://doi.org/10.1016/j.envpol.2005.05.020

Willows RD (2007) Chlorophyll Synthesis. In: Wise RR, Hoober JK (eds.) The Structure and Function of Plastids. Advances in Photosynthesis and Respiration, vol 23, pp 295-313. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-4061-0_15

World Bank (2013) Vietnam urban wastewater review—executive summary. Washington DC. https://www.worldbank.org/en/country/vietnam/publication/vietnam-urban-wastewater-review

Yu Y, Lei Z, Yuan T et al (2017) Simultaneous phosphorus and nitrogen recovery from anaerobically digested sludge using a hybrid system coupling hydrothermal pretreatment with MAP precipitation. Biores Technol 243:634–640. https://doi.org/10.1016/j.biortech2017.06.178

Zain SM, Basri H, Suja' F, Jaafar O (2002) Land application technique for the treatment and disposal of sewage sludge. Water Sci Technol 46(9):306–308. https://doi.org/10.2166/wst.2002.0265

Zhang Z, Zhou Y, Zhang J et al (2016) Effects of short-time aerobic digestion on extracellular polymeric substances and sludge features of waste activated sludge. Chem Eng J 299:177–183. https://doi.org/10.1016/j.cej.2016.04.047

Cities and Utilities in Vietnam

The Last 150 Years of Urban Mutations in Hanoi: An Investigation About Form and Morphology of the Vietnam's Capital City

Matteo Aimini, Ngyuen Dang Giang, and Do Binh Minh

Abstract The essay investigates urban evolution and architectural mutations in the territories of Hanoi city. The narrative is a continuous process, generated by a paradoxical event such as the colonial expo of 1931 that combines some key concepts to read the territories of the city. The various historical phases are recalled through the analogy of architectural, urban and economic design shapes. This type of narration allows to put on the same level, from the point of view of images and text, three apparently and very far away events, such as the French colonial, Soviet and Capitalist phases. In order to highlight a certain linearity in the mechanisms of the city's design. Times change, ideologies die, techniques evolve, materials allow new frontiers, but some basic principles remain unchanged and are recycled into different forms according to the age in which they lie.

1 Introduction

This process of re-reading urban phenomena began in 2005 thanks to an Asia Link exchange project and continued autonomously, with the support of IUAV and the fundamental help of the Vietnamese scholars of HAU who assisted the long and laborious research project.

The research project was formally concluded with the publication of the study in Vietnamese in 2017, thanks to the scientific contribution of Phu Duong University and the Italian Ministry of Foreign Affairs (MAECI). The study belongs to the category of treatises on architecture and history of the city covering a rich historical period unknown to the most. The study of the city and the dynamics of transformation is an important issue for the understanding and respect of the contexts in which we must operate. Regarding this topic, the study takes up a historical narrative tradition

M. Aimini (✉)
UNITN, Università degli Studi di Trento and IUAV, Istituto Universitario di Venezia, Venice, Italy
e-mail: matteoaimini@iuav.it; matteo.aimini@unitn.it

N. D. Giang · D. B. Minh
HAU, Hanoi Architectural University, Hanoi, Vietnam

typically,[1] that over time has set the basis for understanding and critical analysis about the complex systems that govern the future of cities. In an age in which Vietnam faces a great challenge such as the development and modernization of its metropolis, it was considered central to trace the historical and cultural moments that have designed the image of Hanoi today, to understand the richness of the built heritage versus the new architectural interventions that comes forward. An increasingly global "new" that sometimes does not reflect the cultural identities that have designed the city over the centuries, but rather works against the historical and cultural heritage that the city has built up over ten centuries.

2 Materials and Methods

2.1 Colonial Wonderland

In the international exhibition of colonies held in Paris from June 1931 to February 1932, appeared, for the first time, in the small pavilion of Tonkino, a meticulous work of anthropological and geographical investigation concerning the Red River Delta, developed by an unknown geographer named Pierre Gourou. The area under study covered 8000 villages scattered across an archipelago of small communities, with a total of six and a half million inhabitants.[2] *Les paysans du delta Tonkinois*, a pioneering study of rural sociology, it was carried out by means of demographic surveys and direct observation of the relational attitudes of about two thousand sample villages. Gourou did not hesitate to define this model as: "*A perfect relationship between man and nature […]. Outside this state of affairs, there can only be disorder and despair*"[3] (Fig. 1).

He was critical of the violent contamination of the territory by the French, in fact, he considered the appearance of Catholic churches to be an element totally alien to the rural landscape, as well as school buildings and hospitals built without taking into account the environment that housed them.[4]

In *Les Paysans*, Gourou underlined the millenary ecological capacity of adaptation that the farmers of North Vietnam had developed in symbiosis with their environment and how the villages served, articulated in a dense network, to maintain, control and govern the territory. "*The geographical explanation of the landscape must not consist in the relation between two terms, one consisting of the physical dimension and the*

---

[1]See authors as Rossi A., Aymonino C. and Benevolo L.

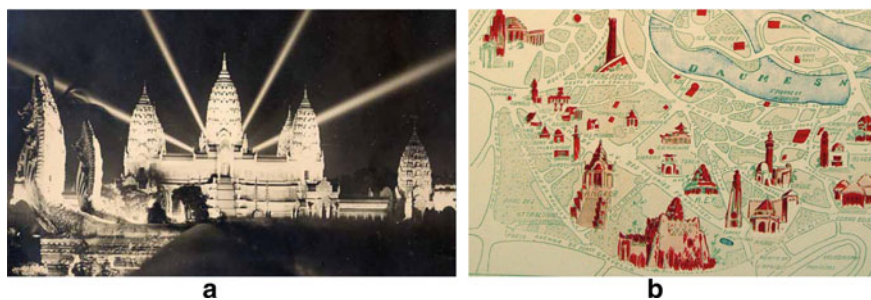[2]Gourou (1936).

[3]Ibidem, p. 575.

[4]Ibidem, p. 567.

Fig. 1 The fake Ankor-Wat. (**a**) The faithful reproduction in steel and concrete of the magnificent complex located in Cambodia during the expo of the colonies in 1931 in Paris (**b**). *Source* Les Cahiers de l'Ipraus (2001), Hanoi, Editions Recherches/Ipraus 2001

Fig. 2 The covered bridge in Son Tay (**a**) is a wonderful example of landscape architecture and the relationship with the environment (**b**). *Source* Fonds Pineau, Louis-Georges (1898–1987) SIAF/Cité de l'architecture et du patrimoine/Archives d'architecture du XXe siècle

other of human elements, three categories of data must be examined: the physical dimension, civilization and the interaction of human elements with the territory".[5]

The precognitive dimension of this statement is astonishing when compared to Article 1, paragraph a), of the European Convention on Landscape, drafted with some delay, almost seventy years later, which states that "*Landscape designates a certain part of the territory, as it is perceived by populations, whose character derives from the action of natural and/or human factors and their interrelationships*".

Gourou had keenly intuited the fragility of the Delta system, linked to new development and possible external contamination, and this is why his approach was conservative in relation to the colonial transformations underway (Fig. 2).

In fact, only by preserving the traditions of the villages and the techniques of the farmers, would the ecological system have remained intact. In the seventies in France, he was harshly accused by Marxist critics of taking "*third world positions*", encouraging with his words the evolutionary stagnation of countries struggling for economic independence. In truth, in 1936, on his return to Paris, Gourou

---

[5]Gourou (1948).

was included in the Guernut[6] commission, set up to evaluate the rules of government in the colonies. In this context, he strongly hoped for a cooperative development program for the Tonkino Delta, building a program based on the modernization of the agricultural process through the implementation of local cultivation techniques, the redesign of agricultural soils and the growth of local craftsmanship, organized in a small network of industries.

His vision, which today we would call sustainable development, was decisively against concessions, a colonial system of exploitation of landed estates that reasoned for large systems. However, all his proposals remained a dead letter and the commission was dissolved in March 1937.

Despite historical events, the question of the process of transition/evolution of rural situations is still a key topic to discuss the phenomena of change in the tropical territories of the Hanoian metropolis. In this paper, as will be seen below, we have tried to define a specific typological condition of the territories under strong pressure of urbanization, called *MetroRurale*, the battlefield where strident and paradoxical contrasts occur.

2.2 The Free Market and the Phenomenon of Informal Settlements

The final disintegration of the fragile balances outlined some time ago by Gourou took place after a tremendous period of war and famine. In 1986, three years before the total disintegration of the Soviet Union, Vietnam decided to open up to the economy of the international market, abolishing a communist system and favoring what could be defined as a socialism oriented towards the free market (*Doi Moi*).

An epoch-making change, comparable in impact to the French invasion towards the end of the nineteenth century.[7] From this moment on, growth appears unstoppable, the city registers increase of more than 3% every year.

From just over a million inhabitants at the end of the 1980s, it reached three and a half in 2007,[8] to reach six and three hundred thousand today. All the management and economic control systems have been reformed. The regime of soils changes and are experienced logics of private ownership and real estate investment. The first phase of the capitalist process is distinguished by three key factors: the reprogramming of the soil and the iconoclastic fury of the symbols of the past; the proliferation of the informal city, here called *Bonsai city* and finally the presentation of the sandwich masterplan, the first capitalist plan that finally turns out to be a strange hybrid "capital-colonial-socialist". In the early 90s the metropolis appears as frozen, its sky line is homogeneous and regular. The buildings do not exceed ten floors. What the B52

---

[6]Kleinen (2005).

[7]Logan (2000).

[8]Ledent (2002).

Americans and the Indochina war against the French cannot do, the global market takes care of it.

Under the motto "*fine and let it be*",[9] in 2000 the process of collective abuse reached its peak and 90% of the new houses in Hanoi were built illegally.[10] The phenomenon called *Bonsai City* or informal city, for the obvious reason of forced and miniaturized growth, is spreading rapidly. A new liberal-anarchist settlement creeps into the free spaces, besieges the colonial fabrics and corrodes the Soviet quarters. The orthophoto of the city appears as a low-resolution image, a dynamic set of multicolored pixels. The 36 streets district, the old commercial centre of the city, is organized in narrow lots, up to thirty meters deep, which housed single or two-family houses, with the shop on the street it has undergone a process of densification. The premises were marked by internal patios where the family's activities were concentrated. A fabric similar to a medieval village where the streets were organized for corporations of craftsmen. At the end of the 80s the city landscape of the district was entirely preserved but with the introduction of the free market, the patios were transformed into staircases and there was a volumetric extrusion of up to 6/7 floors. Sometimes the widths of the roadsides do not exceed three and a half meters.

The density of the neighborhood is skyrocketing. It has 1900 inhabitants per hectare.[11] This is impressive when you consider that Tokyo has 131, New York 112 and London 72.8. The living space is thus reduced to about 1.5 m^2 per person. If we set up this process with a very high income of the soils, even outside the historic center, we understand that the trend is the hyper subdivision of the building plot. The lower the attack on the ground, the lower the incidence of the cost of the land and the higher the revenues in case of high fragmentation. Combined with the almost total absence of building regulations, you will have the genesis of the "*tube house*". In this way we are witnessing the birth of an unusual model of territorial duplication. The influence of the late post-modern draws the styles and colors. Yellow tympanums, orders of Doric columns in the balconies and faux gold capitelli. Everyone is free to express themselves and the opulence of their home (Fig. 3).

The first floor of market socialism, on the other hand, opens with a tremendous effort. Between 1992 and 1997, 86 cities and 358 small villages had to adjust their development plans in line with the new design of the metropolis. The remaining 179 rural settlements were obliged to deliver their variants by 1999.[12] A scale never seen before. It is no longer just the single bulk built-up area of Hanoi that is involved, but the entire region, which is called upon to adapt to the new requirements.

The result is a hybrid plan, the fruit of previous colonial and communist visions, with the addition of a new element/ingredient that marks the entrance into the new capitalist era: the CBD, an acronym for Central Business District, the district of commercial affairs. It is here that it will be possible to introduce the typology that until now was missing in the city, the skyscraper. The high-rises will be located

---

[9]Hanh (2001).

[10]Mathes (1995).

[11]Japan Bank for International Cooperation (1999), op.cit., p. 8.

[12]"Mr Vo Van Kiet incontra le associazioni culturali ed artistiche", Kient *Truc*, no. 3, 1992, p. 37.
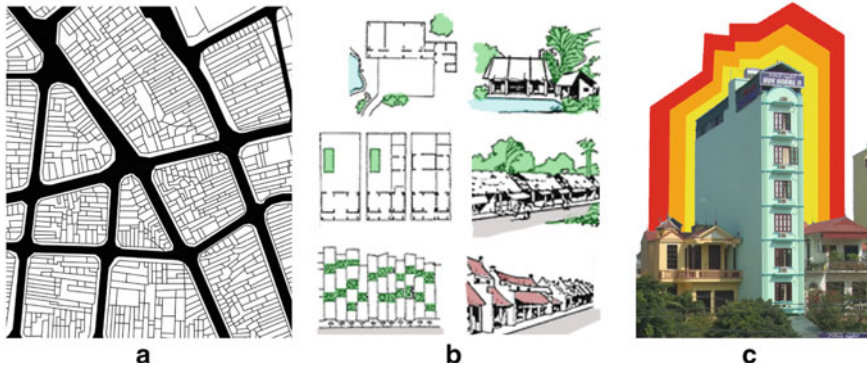
Fig. 3 The urban fabric of the 36 street district (**a**) and the genesis of the tube house: from the rural house to the urban settlement (**b**). The model repeats itself continuously imitating the mother type as in the case of this hotel that one day could easily become a condominium (**c**). *Source* Materials by the author

only on the right side of West Lake, while the historic center, identified within the perimeter of 36 streets, will be preserved. New commercial activities will be established respectively to the north and south. In a heartfelt motion in 1992, Vietnamese Prime Minister Vo Van Kiet issued a warning that today resonates like a sinister prophecy: "*We desperately need legislation for zeroing to protect our identity in architectural language. The zoning laws will help us to prevent that jumble of horrendous buildings and also regulate the types of buildings that will be built in the future, with the aim of protecting our culture and more generally our national identity in terms of architectural language*".[13]

This master plan allows to find a unique condition, which occurred in the mixing of three design instances conceptually belonging to three different development models. It is as if three life cycles met here their common ground, Colonial Forms (promenades, buildings and public spaces), Soviet Productivism (collective living and the system of industries), up to the Capitalist Model (the CBD and new infrastructures), recomposing themselves into a new single layer. If we were to bring this operation back to an equation, we could thus transcribe: Hanoi 1992 = Neo-Liberal Metropolis + Tropical Leningrad + Tonkino's Paris (Fig. 4).

2.3 Paris of the Tonkino

On February 6, 1874, the Annamita government granted an area of marshland south of the city, along the banks of the Red River.[14] The first bastion of western urbanity, located in a very small area and outside the city walls. The total area was about

---
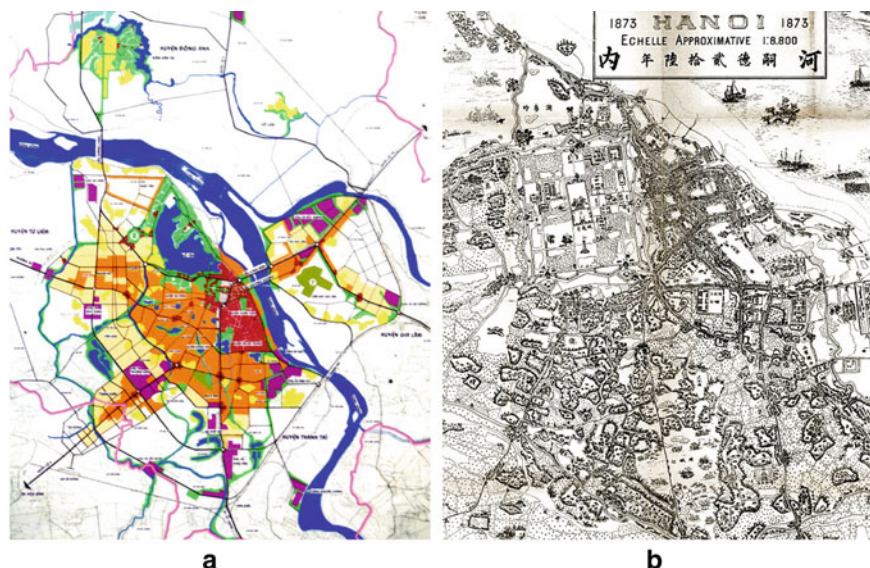
[13]Masson (1929).

[14]Ibidem, p. 98.

Fig. 4 The first post-Soviet floor: The Sandwich Masterplan (**a**) and Hanoi 1873 (**b**). *Source* Materials by the author; Biblioteque Nationale de Paris

eighteen and a half hectares.[15] The concession was the embryo of a new urbanity. "*France has the ability to do what the Annamites still do not have, the capital, the machine, the engineers and a long experience in business. The French will be your older brothers*".[16] The words of Paul Bert, who was proclaimed the first governor in January 1886, triggered the period of the heroic functionalism, whose material executor was August Henri Vildieu, architect in charge of the technical office for design. Hydraulic works, a clean slate of buildings symbolizing the local culture and the application of the colonial chessboard were the main reasons for their work. "*I arrived in Hanoi too late to save the interesting parts. The city gates deserved to be preserved. They could have made beautiful the future quarters of the city […] or the citadel itself built to remain over the centuries … the question is more embarrassing than ever*".[17]

Doumer was Paul Bert's successor in 1887. The new governor proclaimed himself the father of modern Indochina and felt the need to transform Hanoi into a great colonial capital. The governor's secret ambition was to create a miniaturized Paris. The volume was used as a symbolic weapon. The neoclassical French Beaux-Arts style was adopted as a new language and the functional program of the "buildings in the city" established new mechanisms of power.[18] The station and an enormous railway

---

[15]Noury (1992).

[16]Doumer (1905).

[17]Pèdelahore (1992).

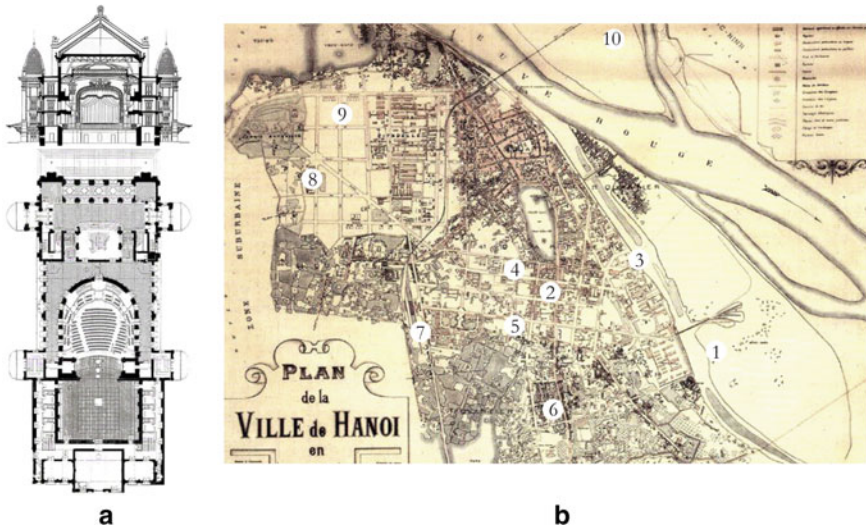[18]Les Cahiers de l'Ipraus (2001).

Fig. 5 The Hanoi Opera Hose (**a**) and Hanoi 1902 (**b**). (1) The concession; (2) The European Quarter; (3) The Opera House; (4) The Palace of Justice and the Central House; (5) The Palace of the Expo; (6) The European Residential Quarter; (7) The Central Station; (8) The Palace of the Governor; (9) The Botanical Garden; (10) The Doumer Bridge. *Source* Les Cahiers de l'Ipraus (2001), Hanoi, Editions Recherches/Ipraus 2001. Materials by the author

bridge over the Red River were built in a very short time, the reclamation of the ponds and the consolidation of the streets had the immediate effect of paving the way for many majestic buildings. Boyer and Harvey built the new Opera House,[19] Charles Lichtenfelder the Palace of the Governor of Indochina,[20] downsizing a megalomaniac project by H.A. Vildieu[21] that built later the Palace of Justice and the sadly famous Central House prison (Fig. 5).

Maurice Long, the third Governor General of Hanoi, established a new architecture and planning service for the city in 1923, hoping to introduce new development policies similar to those initiated by Governor Henri Prost and Grand Marshal Lyautey in Morocco. Long saw Hanoi as a laboratory for experimenting with the latest ideas and innovations in urban and legislative matters that had just been introduced in his mother country.[22] To this end, the governor personally invited E. Hebrard in Indochina to view the project for the holiday city of Dalat and to undertake a proposal for a plan for Hanoi: it was thus that the architect took up the post of director of the architecture and planning service in 1923.[23]

---

[19] Ibidem, p. 155.

[20] Logan (2000), op.cit., p. 87.

[21] Les Cahiers de l'Ipraus (2001), op.cit., p. 120 (in merito alla legge Cornudet del 1919).

[22] Logan (2000), op.cit., p. 99.

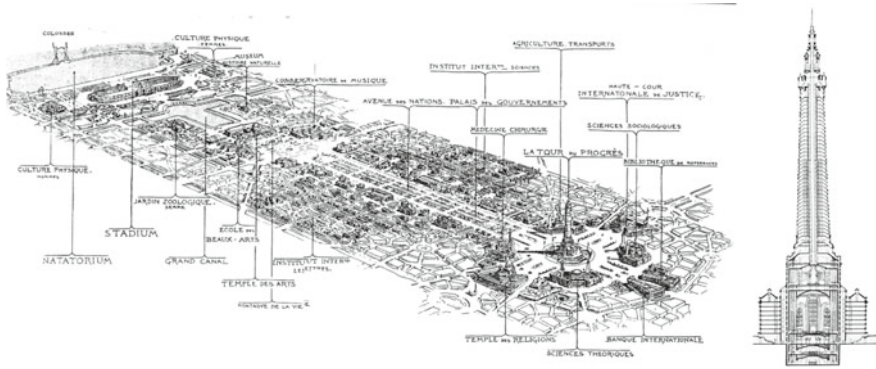[23] Gresleri G. e Matteoni D. (1982), op.cit., p. 66.

Fig. 6 The axonometric scheme of the axis of the nations of the World City and the Tower of Progress. *Source* Gresleri G. e Matteoni D. (1982), La città mondiale, Polis/marsilio Editori 1982

Ernest Hebrard, Grand Prix de Rome 1904 and a founding member of the first French Société des Architectes Urbanistes, before arriving in Indochina he had already signed some prestigious projects, such as the reuse of the Diocletian Palace in Split and the unfortunate enterprise of the World City,[24] the great utopia of a newly founded city, dedicated to the diplomacy and arts of the planet (Fig. 6).

Hebrard imagined the city as a hierarchy of infrastructures, predisposed to receive and distribute copious populations. A great hub in its embryonic state. Hebrard established in the project of the world city, a precise hierarchy between architectural, urban, picturesque, political and spiritual order. He encoded the individual elements into a series of spatially recognizable organizational principles: the administrative and cultural center, the heart of the system, the organization of places for recreation, in symbiosis with the masses of trees, residential districts with variable densities and finally the productive sectors outside the town.

Hebrard in Hanoi tackled the problem by mixing pragmatism with a dose of calibrated utopia, probably a legacy of the fantastic project of the world city. He worked for a lucid reorganization of the functional districts, the result of which was the appearance of a compact administrative center in the area of Ba Dinh (Hebrard 1928), where today stands the mausoleum of Ho Chi Minh.

The plan for the design and reorganization of the banks of the great Ho Tay (West Lake), transformed into a large promenade of public greenery. A sort of ecological infrastructure, which contained a series of activities for leisure and recreation. In doing so, Hebrard doubled Hanoi's size, thus preparing for a large expansion of the building in the areas to the west, towards West Lake, to the south around Lake Bau May, and to the east with the industrial district of Gia Lam.

The design of the plan was held together by the regulatory mechanisms of zoning: built to organize the functions of the new city, to ensure control over the indigenous peoples and activities, increasingly surrounded by the new urbanity. Hebrard thus

---

[24]Hebrard (1928).

extended the land-use zoning to the indigenous settlements, obsessed with the idea of achieving a perfect colonial society that rejected the monotonous order of the chessboard fabric but instead preferred the vocation of the places and the dynamics of hybridization. In addition to defining the layout of the future metropolis, Hebrard was responsible for a significant number of public buildings such as the Ministry of Finance, now the Ministry of Public Affairs, the Finot Museum, the highest expression of the concept of cultural relativism, the University of Hanoi and the church Cua Bac, in eclectic style with art deco influences. The economic crisis of 1929 slowed down the budget for public works and led Hebrard to abandon Hanoi definitively to its fate, leaving many of his projects in the middle and the program for public spaces at an embryonic stage (Fig. 7).

Fig. 7 Hanoi 1923. (1) The axis of nations; (2) The new industrial area; (3) The education sector; (4) The European commercial district; (5) The railway station (6) New residential districts; (7) The concession; (8) New railway bridges; (9) New residential promenade; (10) Sports sector; (11) The leisure area. *Source* Logan (2000)

2.4 Post-crisis

One year after the terrible Asian crisis of 1997, a second variant of the sandwich masterplan is presented.[25] The infrastructure is more marked, and on the other side of the Red River a new city appears, mirroring the first one, which connects the industrial district of Gia Lam with the airport of Noi Bai. The migratory flows towards the city are constantly increasing and the plan in response provides for a new redistribution of density. Public Officials predict from 2.13 to 4.70 million square meters of new residences.[26]

The arduous task of raising the living space from less than 3 m^2 per person to about 5.5 m^2[27] generates in the agricultural fragments on the edge of the suburbs, in close contact with agricultural territories, an army of towers of 30 and 40 floors. The territories turn into enormous patches of design anarchy, duly fueled by optimistic forecasts of urban development, which set for 2010 the variable target of 12/18 million square meters of recovery and new housing.[28] The equivalent of about 3 buildings of the same size as the Pentagon in Washington DC (synthetically the building with the largest surface area in the world)[29] (Fig. 8).

The change of perspective requires a revision of the concept of ownership. Until 1993, the land belonged to the Vietnamese people and therefore to the central government.[30] Only five years later the concept changes into a binary formula. Concession of land use with or without charges. Needless to say, in order to be able to build residential or commercial buildings, a sum equal to the cost of land use must be paid. Although not explicitly stated, this mechanism is considered to all intents and purposes a sale by the state to the private individual, who has the right to permanently use the land "rented".[31]

The plan foresees three macro areas of development: a central one, in the consolidated fabric, aimed at the development of new infrastructures and transports, a difficult and very expensive operation due to the dense and irregular fabric; a second one concerns the low-density areas, located in rural villages and a third one concerning virgin agricultural lands, where the main changes take place. The neo-liberal city is a product of our contemporary society. This phenomenon is evident in countries with strong economic development, where private investors, multinationals, banks from foreign countries and internationally renowned designers run and contribute to an unprecedented injection of urban botox. It rains in a few years in the capital

---

[25]Logan (2000), op.cit., p. 110.

[26]Japan Bank for International Cooperation (1999), op.cit., p. 84.

[27]Tuong Lai (1998).

[28]See Footnote 26.

[29]Vogel (2003).

[30]Japan Bank for International Cooperation (1999), op.cit., p. 2.

[31]Decree No.17/1999/ND-CP on Procedures for Exchange, Assignment, Lease, Sublease and inheritance of Land Use Rights and Mortgage of Land Use Rights and Contribution of Land Use Rights as Capital, Marzo 199 in *JBIC*, p. 41.
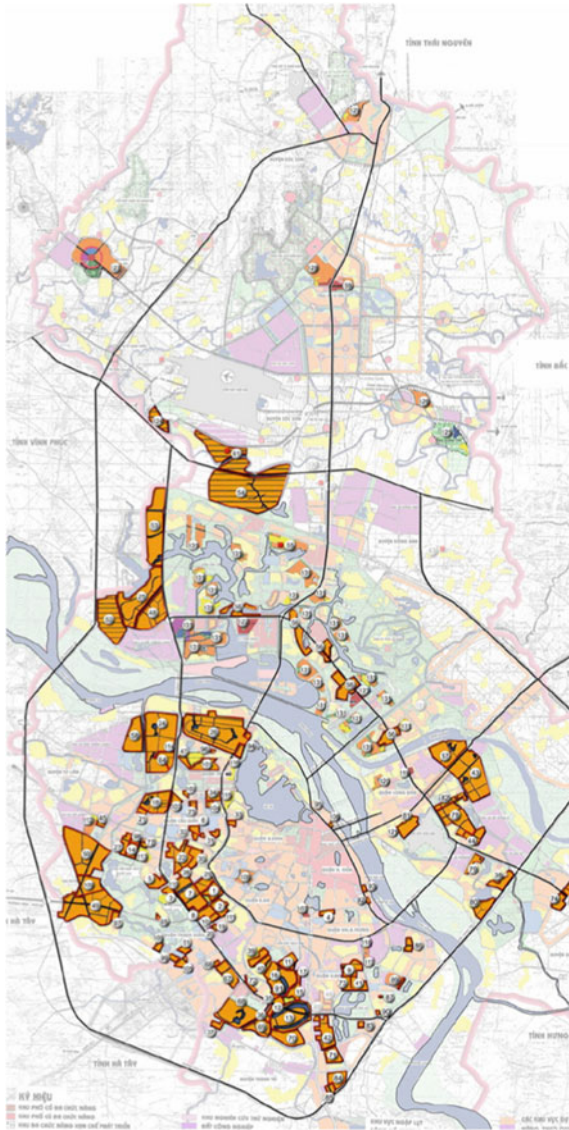
Fig. 8 Development areas already approved in the 1998 plan. *Source* Materials by the author

region, at least thirteen industrial parks,[32] for a total area of 7 times Central Park in Manhattan.

While the main source of production, agriculture, is temporarily shelved, despite constituting the primary armor of the settlements of the ancient city-region. The

---

[32]http://www.business-in-asia.com/vn_ip_north.htm#1.

balance traced by Pierre Gourou, in his study entitled "The civilization of the plant" is definitively crumbled. Japanese, Indonesian, Korean and Singaporean investments fell on the landscape of the city, in zones of special economic development. Golf courses, twin towns and new luxury districts mark the beginning of an unstoppable development. The most ambitious project was proposed in 1997 by Daewoo International. The objective is the design and management of 8000 ha for the settlement of one million inhabitants by 2040. For the first time after the Soviet teams, a number of major international architectural and engineering firms are invited to participate. Bechtel, a large American engineering company, prepares the feasibility study of the area and the Vietnamese government invites three internationally renowned studies for the development of the areas.

OMA will be responsible for the design of the Dong Anh[33] district, SOM[34] of the Van Tri lake area and the Nikken Sekkei of the Tu Liem site, on the left side of West Lake.[35] OMA and SOM are competing on the same areas. The two projects not only provide a portrait of a specific historical period, but also allow for some reflections on the concept of new urbanity. The two projects have obvious similarities because of the required program, but the directions of design and functional distribution are profoundly different. SOM, according to an iron logic of zoning, it act dividing and reorganizing the landscape according to a clear and well-defined scheme, using the element of water as an urban anesthetic for a new miniature of Chicago. While OMA shape the territory according to the scenarios they propose, balancing the inevitable artificiality and favoring new ecological possibilities. The landscape becomes an integral and inseparable part of the design. "*If there is a new urbanism will no longer be based on twin fantasies of order and omnipotence, it will be the staging of uncertainty and will no longer be interested in the arrangement of more or less permanent objects, but rather interested in the irrigation of territories with potential. To survive, urban planning will have to imagine something new […]*"[36] (Figs. 9 and 10).

The economic crisis that hit the countries of South Asia in 1997 dealt a severe blow to the brilliant Renaissance/experimental perspective. Daewoo is looking for an exit strategy and in July 1999 the project was transferred to Korean international cooperation. The American engineering company Bechtel and Daewoo International reworked another plan, which turned out to be the merger of the two previous projects of OMA and SOM. A boring synthesis generates a strange process of hybridization. It all looks like a 1930s garden city, where there is a miniature Chicago backed by a drop-shaped wedge-island. Despite everything, the plan is approved and definitively incorporated into the future urban development of the city. "*The certainty of failure must be our laughing/oxygen gas; modernization is our most powerful drug. Since we are not responsible, we must become irresponsible […]*".[37]

---

[33]Hanoi New Town (1997), *Documento di piano*, p. 1. Per gentile concessione di OMA.

[34]Hanoi New Town (1997), *Documento di piano*, p. 1. Per gentile concessione di SOM.

[35]Kwon (1997).

[36]Koolhaas (1995).

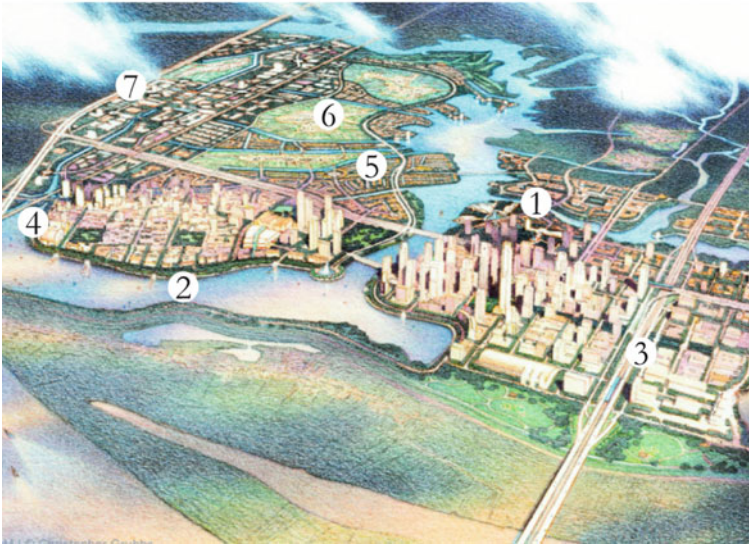[37]R. Koolhaas, op.cit., p. 959–971.

Fig. 9 SOM's design bird's eye view. (1) CBD; (2) Reception functions; (3) Government buildings; (4) Medium density neighborhoods; (5) Low density residences; (6) Existing Villages; (7) Production sector. *Source* Courtesy of Skidmore, Owings & Merril

2.5   Tropical Leningrad

The Vietnamese version of the MicroRayons materialized in the KTT, an acronym for Ku Tap The, districts of collective residences. Since the end of the 1950s, about 30 projects have been realized, for a total of 450 ha.[38] The area per person ranged from 4 to 6 m^2 and were designed to accommodate an average of 4000 to 10,000 inhabitants. The design of the open spaces was well organized and the distances between the buildings were more than ever satisfactory for a medium-density model. S.I. Sokolov, the future director of the Leningrad School of Urban Planning[39] arrived in Hanoi in the late 1960s. Already famous in his homeland for winning numerous architectural and urban planning competitions, he was called to work in Vietnam for his specific expertise in the field of new social districts. The orientation and dynamics of the residents of the Microrayons was in fact the title of his doctoral thesis.[40]

---

[38]Procacci F., Luong Thu Thao (2007), "Learning form Ktt", Hanoi Architectural University, Hanoi, p. 2. Paper presentato durante la conferenza *Internazionale di Architettura sostenibile e disegno urbano.*

[39]Ruble (1990).

[40]Ibidem, p. 206.

Fig. 10 OMA masterplan. (1) Research Island; (2) Business Island; (3) Sport Island; (4) Resort Island; (5) New Expansion; (6) Satellite Villages; (7) Existing Villages; (8) Cultural Corridor; (9) Embankment Recovery; (10) Cultural Activities; (11) West Lake. *Source* Courtesy of Office for Metropolitan Architecture

The Microrayons, the socialist version of neighborhood units, offered an opportunity to solve the problem of housing and density. Despite the defiance that the Soviets showed towards Howard's garden city as it was judged to be the bourgeois expression of the capitalist metropolises unable to resolve the perennial conflict between city and country because of the typological implications that the single-family house imposed, preventing the possibility of developing a collective consciousness compatible with the socialist spirit.[41] The new neighborhoods they designed, the so-called 'micro-districts', were, however, very similar if not identical in concept to the theories of neighborhood units designed by Clarence Perry in the late 1930s. In fact, in the center of the neighborhood were concentrated public buildings, such as the school, kindergarten, civic center or church, which were to serve a radius of about five hundred meters. On the edge of the neighborhood was a shopping center. The relationship between the buildings built and the public space was well balanced, to encourage open-air activities.[42]

The Soviet version was identical in the layout of the services and in the proportions of the public spaces, but differed in the greater density and in the commercial services,

[41]Kopp A. (1967), op.cit., p. 176.

[42]Perry (1929).

Fig. 11 The soviet vision for Hanoi city

which obviously distributed only the strictly necessary, standardized and possibly fruit of the work of the community that lived there. A sort of prototype for an autarkic neighborhood (Figs. 11, 12 and 13).

Beyond the Microrayon programme, the vision of the socialist city was composed of monumental public buildings and spaces of considerable importance, both in size and symbolic value, production hubs and infrastructure. The definitive glue for this new development was the city plan drawn up by I. S. Sokolov and his team, in collaboration with Huyen Tan Phat and his group.[43] The project was immediately called the 'Leningrad Plan', because the Soviet members belonged to the Institute for Urban Research in the soviet city. The plan was for 360-degree growth. The great Lake Tay Ho (West Lake) was immediately indicated as a new and important polarity for the city. Large boulevards with plates and skyscrapers, based on the Moscow model of the Kalinin Prospect, were traced throughout the city, with the intention of creating new polarities of services.

One of them razed to the ground half of the historical quarter of the 36 streets. In addition, a new railway link was planned to encircle the city and cross the Red River to the north, passing under the new gigantic bridge that connected the future Noi Ba airport, located 65 km away. A sinuous and calm green belt organized the containment of the new building, while on the other side of the river was established a large nature reserve for well-being and rest. Five thematic districts of an industrial nature were located at the edge of the town. However, the demographic forecasts for the year 2000 were totally unreal, as Sokolov himself later admitted. Behind the cryptic meaning of the statement lay the principle of colonial exploitation.

It was useless for a colony rich in raw materials to develop a heavy industry, first because the raw materials were needed by France, and second because a close

---

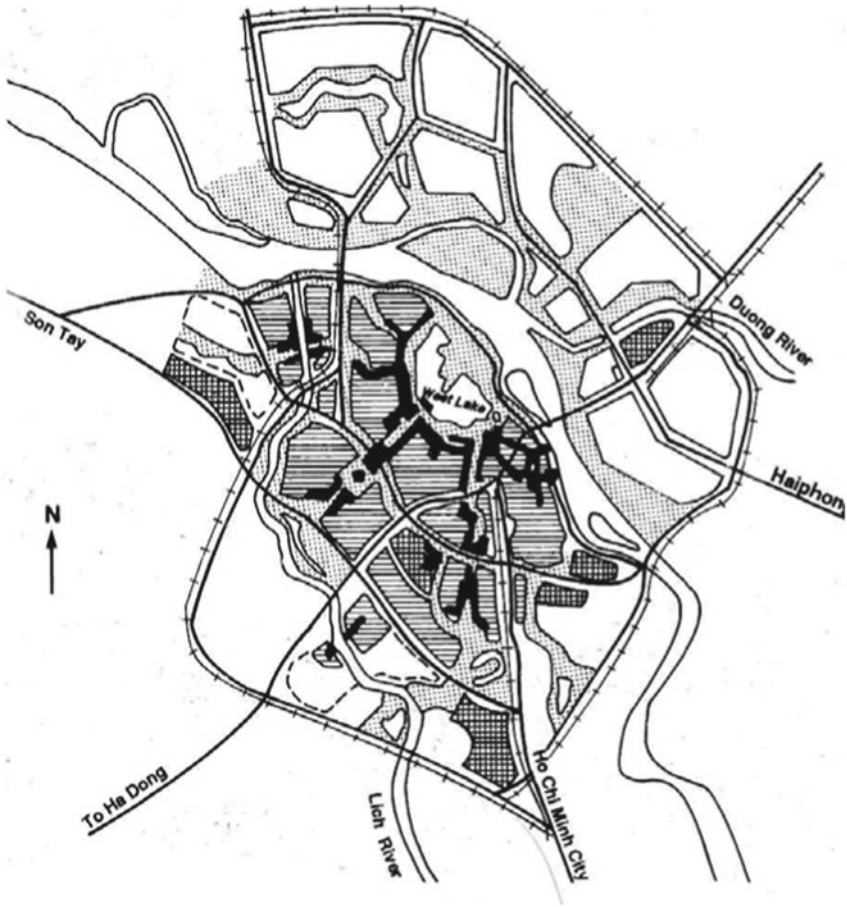[43]He Van Que (2000), op.cit., p. 89.

Fig. 12 The plan of the Leningrad school for Hanoi. *Source* Logan (2000)

bond of dependence had to be established for the benefit of the colonizing civilization. Independence and freedom encouraged by the Soviets, on the other hand, were based exactly on the opposite principle. Support for development and infrastructure, mass production on the socialist model, creation of a collective consciousness and enhancement of the urban and rural proletariat.

Despite the good intentions of economic independence, with the arrival of the Soviet architect G. G. Isakovich, the architectural language was a repetition of the system already consolidated during the French period. Isakovich studied at the Moscow school in the first half of the 50s,[44] having as mentor Professor B.S.

---

[44]Prokhorov (1970).

Fig. 13 Functional detail, note the public centers marked in black that correspond to the demolitions of much of the historic center. *Source* Logan (2000)

Mezentsev, already famous during the Stalinist period for having realized in collaboration with A. N. Dushkin one of the seven sisters of Moscow in Lermontov square.[45] The two professionals were real archistars of the communist regime. The area chosen to imprint the socialist sign in the city was the Ba Dinh area, exactly in the convergence of the axes drawn by E. Hebrard, the place of the declaration of independence of North Vietnam in 1954. Both Mao Tse Tsung and Ho Chi Minh were bound by the same destiny, both wanted to be cremated to disperse their ashes in the beloved land. The result was that they both ended up in a Mausoleum. Vuong Quoc My and Ngyuen Ngoc Chan were pushing for a modern, dignified and simple solution that would follow the Vietnamese version of the Communist Party.
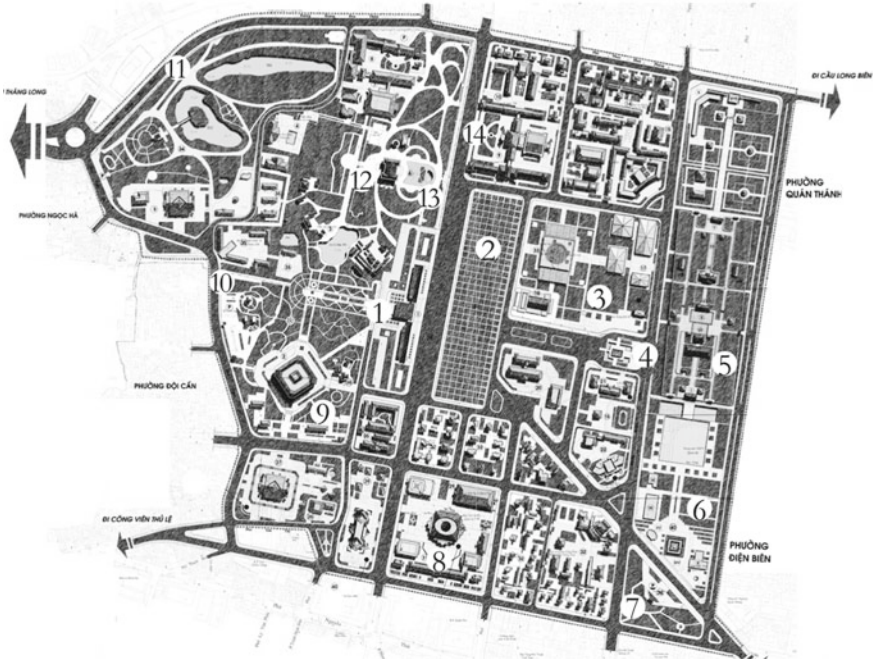
---

[45] Forte R. (2006), op.cit., p. 27.

Fig. 14 The project of the Vietnamese nation starts again from the construction of the monumental public space. Where the Russian level overlaps with the French one. (1) Mausoleum of Ho Chi Minh; (2) Grassy carpet; (3) The new palace of the popular assembly; (4) Monument to the unknown soldier; (5) The citadel; (6) The war museum; (7) The statue of Lenin; (8) Sports center; (9) The war of liberation museum; (10) The reconstruction of the Pagoda of a single pillar; (11) The botanical garden; (12) The palace of the governor of French Indochina; (13) The remains of the A. Serrault High School. *Source* Materials by the author

Obviously, the opinion of Mesentz and Isakovich prevailed, who claimed enormous experience in the construction of mausoleums, the result of which is still visible today. Isakovich's work was not limited to the Mausoleum but continued, in collaboration with Nguyen Truc Luyen, with the design of the building, symbol par excellence of the indissoluble bond with the Soviet Union: The Cultural Palace of Soviet-Vietnamese Friendship.[46] The similarities with the Moscow Youth Palace are quite evident. Built around the same time, they show how the mass production of Soviet architecture in the 1970 s, with rare exceptions, was highly standardized and aimed at creating a coordinated, univocal and recognizable image even in the most remote peripheries of the empire (Fig. 14).

Also, by Isakovich was the museum of Ho Chi Minh, located next to the Mausoleum in Ba Dinh Square. A three-store rhomboid building that housed the permanent exhibition of the revolution. In front of the old watchtower of the citadel, the only relic that survived the French iconoclastic fury, during the colonial period a

---

[46]Melnikov (1986).

memorial to the unknown soldier was placed, promptly scraped off after the victory of Dien Bien Phu and replaced by a bronze statue six meters high, representing one of the last remaining reproductions of V. I. Lenin in the world. The project was obviously designed by the specialist Isakovich and made by the sculptor A. A. Tyurentov.[47]

Between 1976 and 1985, gross domestic product grew by only 3.7% per year. Agriculture, despite efforts to industrialize the country, was still the key sector of the economy. Inflation in 1986 reached 744% on consumer goods.[48] Thus, opened the era of the Doi Moi (the renewal) with the consequent and timid reforms towards a market economy. Three years later the Berlin Wall fell. The Soviet galaxy was definitively shattered. In 1991, the former Soviet Union pulled the plug on all welfare structures in the socialist republics.

2.6   New Territories

November 2005, Vietnam's gross domestic product grew by 6.8% annually,[49] Hanoi city by 9.6% 55. In 2007, with the entry into the WTO, foreign investment in the territory increased dramatically. In 2010 there were more than 350 urban development projects in the city of Hanoi alone, for a total of one billion dollars.[50] Modernization proceeds in forced stages. The fifth capitalist period is just around the corner (2010–2015). Between 1994 and 1999, the population of rural municipalities in the capital's provincial city grew by 65%.[51]

The opening of the capitalist market, the definitive dismantling of the socialist agricultural cooperatives and the consequent redistribution of agricultural land to private subjects, has generated the process of informal urbanization of rural villages, but when in 2006, the state decides that the fate of agricultural land belongs to the thousands of municipalities and numerous provinces, decrees the beginning of the real estate boom in all the urban areas of the city, definitively putting an end to the pure agricultural landscape. "*The only existing distinction between urban and rural is gradually blurred into a scenario of porous spaces or irregular geographical development under the egenomonic command of the capital and the state*".[52]

The combination of a strong heterotopic condition of urban tissues, the extreme density conditions of agglomerations, the state of fragile city boundaries and the condition of 'liquid' territories define a new territorial typology: metrorural. A hybrid system, a new type of landscape that mixes artificial and natural, in a probable dichotomy with unexpected implications. *Metrorurale* is an ambiguous concept: ruralized metropolis or rural metropolisation?

---

[47]Logan (2000), op.cit., p. 198.

[48]Ipraus (2001), op.cit., p. 308.

[49]https://www.cia.gov/library/publications/the-world-factbook/geos/vm.html.

[50]http://en.vietrussia.com/bizcenter/0/news/1602/21608.

[51]Labbè D. (2010), op.cit., p. 40.

[52]Gregorio (2009).

The first manifestation of the powerful metropolitan condition takes place in a first strip of about 3 km, on the edge of the city, where there is a bizarre coexistence of situations that recompose themselves in an unbridled collage. "*A significant number of projects, 744 in total, are underway […]. Most of the development proposals concern the area to the west of the city, between the third and the future fourth ring road*".[53]

Leaving aside the *Bonsai City* phenomenon, which is always of lesser impact and much harmed by the competent authorities, the typological families of interventions in this area are reduced to three types of categories:

- Building/Megastructure
- Urban block
- Neighborhood/city.

The city is in turmoil, ten years have passed since the 1997 plan, very little has remained of the projects of OMA, SOM and Nikken Sekkei and the administration is concerned, in the face of a conspicuous demand for development, that it has not yet developed a suitable image for the future of the city. Unexpectedly, in 2007, the mayors of Hanoi and Seoul asked the Vietnamese Prime Minister to authorize a new visionary plan for the development of the shores of the Red River, signing an official agreement for technical assistance and development of the second phase of the disputed and famous $7 billion project of 1997.[54] The design is divided into four areas for a total of 2462 ha. The cost of the plan will be borne entirely by Korean private funds and development funding. The program includes hi-tech productions, luxury apartments, international exhibition areas, sports, tourism and leisure equipment.[55]

August 3, 2008 is a crucial day. Three international consultants are appointed by a board of experts to present their plans for the future of the city of Hanoi. Arata Isozaki (JP), OMA (NL), present a multi-polar and sustainable vision of new urban settlements. Perkins Eastman (US), Posco E&c (KO), Jica (KO) base their work on a system of green corridors along the urban development, guaranteeing the future metropolis more than 60% of unbuilt green spaces. RTKL (US), instead, presents a model of five satellite cities, each with a different functional program.[56] On September 24, Prime Minister Nguyen Tan Dung, in agreement with the Minister of Construction and the People Committee of Hanoi, announced that the choice fell on the project of the American-Korean hybrid consortium (PPJ). In December 2008, the agreement was signed.[57]

The Paris of Tonkino plan covered 500 km^2. In 1991, during the first capital phase, the boundaries of the provinces had contracted to 900 km^2 compared to the previous

[53]Per gentile concessione del Perkins Eastman office (2009), Documento di piano presentato nel 2008 e approvato nel luglio del 2011, p.40.

[54]http://www.skyscrapercity.com/showthread.php?t=624271.

[55]See Footnote 54.

[56]http://www.vietnamembassy-thailand.org/vnemb.vn/tinkhac/ns080804155854.

[57]Government according to Ordinance on Hanoi Capital No. 29/2000/PL-UBTVQH10 of December 28, 2000.
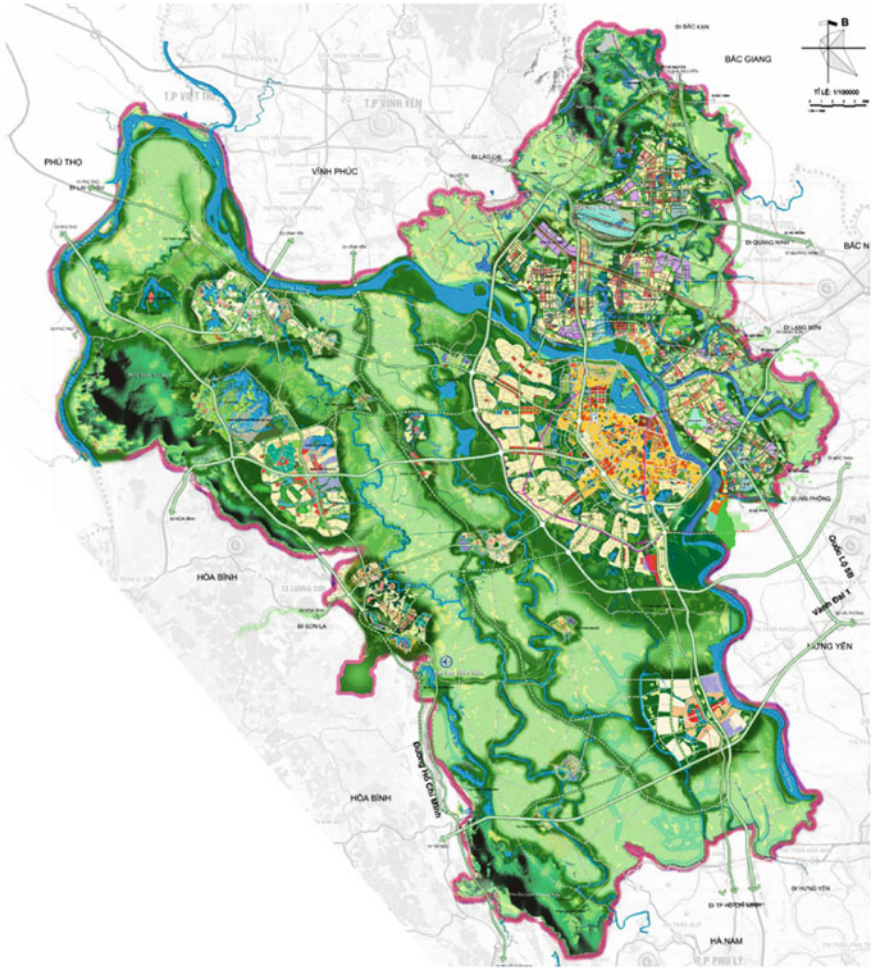
Fig. 15 The plan for Hanoi 2030/50. *Source* Courtesy of Perkins Eastman

3000 km^2 of the Soviet productivist period. A reduction due to the introduction of new rules for the soil market, probably made to exercise greater control over the development of peripheral rural districts, to which the central government had changed the status quo, from rural to urban, to allow the inflow of capital.[58] With the new vision of 2008, the administrative boundaries of the city return to expand dramatically, including the province of Ha Tay and some districts and municipalities of Vin Phuc and Hoa Binh, thus reaching 3300 km^2. Data comparable to the total area of the Aosta Valley region (Fig. 15).

---

[58]Quertamp (2003).

After the NeoLiberal city, which only fifteen years earlier timidly took its first steps in the world economic process, the strategic design becomes more aggressive, extending to a scale no longer purely urban but territorial, in size, environmental changes and management of territorial resources. Expanding Hanoi involves 6.5 million inhabitants of which almost 4 are classified as rural populations.[59]

MetroRurale is no longer limited to the edges of urbanized areas, but becomes a determining factor on a large scale. The region is the key to attracting new investment, relocating inhabitants and raising living standards.

The sixteen points contained in the introduction to the PPJ consortium's plan document exhaustively outline the intentions and ambitions of this project. Some aspects such as environmental conservation and maintenance, the development and redesign of primary and secondary infrastructures seem to be decisive.[60] Practically everything that was possible to predict has been included and can be summarized in the following seven categories: Environmental Strategy, Infrastructure, Consolidated City, Expansion Bags, High Density Garden City, Eco-Sustainable Cities, Multi-purpose Satellite Cities and Productive Districts.

> Our plan has been built around the concept of sustainability, we are trying to push Hanoi to recognize the wonderful and unique opportunity it has to accomplish something that the Chinese have failed to do….[61]

Compared to the 1997 plan and the 2007 Haidep report, the expansion of Hanoi's territory is unbalanced inwards, extending as far as the mountain ranges on the border with the province of Hoa Binh. The region is covered by a dense network of rivers, canals and water reserves. There are more than a thousand rural villages in the area, dedicated to crafts and rice culture in one of the most fertile areas of the region.

The plan states that 70% of the surface area of the metropolitan area is devoted to green areas or interventions aimed at preserving the specificity of the collective space. The areas of the Day and Tich rivers, which are subject to frequent seasonal flooding, will be the cornerstone of the ecological corridor.[62] A large strip that crosses the interior of the region from north to south, blocking on the one hand the development of the city and ensuring within it a targeted development and content of rural villages, encouraging the maintenance in terms of production and local traditions. The ecological corridor, a direct evolution of the greenbelt, is an enormous hinge between the consolidated built-up area and the widespread settlement that will develop according to the indications of the plan.

The water regime is a fundamental aspect of the regional spatial planning, since the territories are located below the level of the Red River. This is demonstrated by the various plans for flood containment infrastructure and the consolidation of decompression reservoirs.

Given the disastrous state of public transport in the city, the plan directly addresses the issue, reiterating the need, already proposed in the document Haidep 2007, for a

---

[59]Hanoi Statistical Yearbook (2008), Hanoi Statistical Office (2009), p. 40.

[60]Perkins Eastman (2009), op.cit., p. 3.

[61]Mark Lamster (2011).

[62]Perkins Eastman (2009), op.cit., p. 24.

series of light and underground metro lines that connect various points of the center and the future urbanity of the region. Japan, China and France are the main promoters of these pharaonic infrastructures.[63]

The conceptual image of the project unequivocally evokes Howard and the garden city.[64] More a visual assonance than a real semantic affinity. A compact orange heart, to which a semi-moon is attached to the west. Cities and new expansions are surrounded by a soft green shape. At 360° appear various beige balloons of changing size, all connected by a radius of about 50 km, with axes converging towards the center. The direct codification of this scheme translates into a partial freezing of the consolidated city, a buffer expansion, three ecotowns, five satellite cities, three production districts.

The new expansion area, already partly covered in the definition of Metrorurale, is one of the souls of the project and represents the strip most adjacent to the consolidated city. The area is between the third and fourth ring road and the small river Nhue. In this area the most evident MetroRural condition is manifested. In fact, the dossier estimates that in the future you will be accommodated from 1.2 to 1.3 million inhabitants.[65] The approach used is of a restraining type, fifteen bags with irregular shapes and vaguely amoeboid, perhaps inspired by the functioning of the immune system, contain the majority of the 744 projects, from 5 to 300 ha, which over time have already been partly approved.

An enormous urbanized continuum, in which architectural typologies alternate, already evolved in comparison to the first capitalistic phase, whose main typologies are the following:

- Asian-style super-block, from 5 to 15 ha
- Condovillas, from 20 to 80 ha
- Villas + Service Centre, from 80 to 200 ha
- The last typology concerns the road infrastructures coming out of the third ring road, which connects the third and fourth rings, and then transforms once you meet the ecological corridor into park-ways.

Beyond the fourth ring road, where the development for expansion pockets ends, meeting the first major ecological corridor, the development plan of the consortium PPJ provides for the construction of three Eco-city controlled development and low-density, about 65,000 inhabitants each.[66] Phuc Pho will be used for advanced agricultural and food processing processes. Chuc Son will be dedicated to social and leisure activities. Quac Oai is a very special case: it represents an anomaly with respect to the above mentioned dimensional standards. Designed directly by the authors of the

[63] Ibidem, p. 30.

[64] Ibidem, p. 19.

[65] Perkins Eastman (2009), op.cit., p. 21.

[66] Perkins Eastman (2009), op. cit., p. 27.

masterplan, Posco Perkins and Jina, the small Eco-City will eventually occupy an area of 2324 ha, for a total of 240,000 people.[67]

New Town terminology was introduced in Vietnam in 1994 with the first experiment at Linh Dam of 184 ha. Since then, New Towns have developed rapidly in terms of quantity and quality and have been a positive factor in meeting capitalist and socio-economic development. Since 2005, 131 new New Town projects have been approved and built.[68] Until 2008, the New Cities were mostly neighborhoods, expansions at the edge of the city, included in the current plan for 2020. Today they are truly new cities, or rather satellite urbanity, 20–50 km from Hanoi. The plan foresees five of them. New settlements designed to be self-sustainable, with different functions depending on their vocation, ready to accommodate a total of 1.47 million inhabitants.[69]

Soc Son, logistics hub. SonTay, tourism, health, wellness and recreation, Hoa Lac, the university centre, Xuan Mai, the old Russian satellite city, will host industry and intensive agriculture. Phu Xuyen will be the second logistical hub towards the red river. On the other side of the river, already the subject of much attention in previous years, intensive development is expected for the district of Dong Anh (former Oma project), of about 500,000 inhabitants, while Me Linh (former Som project) represents the most successful industrial area of Hanoi, which over time has managed to preserve its intensive production of flowers and its agricultural areas. It is estimated that in ten years it will reach a population of about 450,000 inhabitants.[70] Gia Lam, the old industrial district of Hebrard, exactly under Dong Anh, the plan document foresees a doubling of the population by 2030, thus reaching the figure of 700,000 inhabitants becoming the service centre of the eastern area of the Red River.

3   Results and Conclusion

Hanoi, still image of a lightning-fast expansion, of a driven urban desire. Rain of buildings in a territory under constant pressure. The imagery and reflections produced on the future of the metropolis in such a short period of time are disturbing. An immense number of projects of all shapes, genres, sizes and types.

Behind the inexorable ferment hides a creeping sense of nausea in the continuous recognition of the same concept of urbanity, repeated many times and in various possible variations. The pyramid of the urban form is the skyscraper as the vertex of a geometric figure whose base is supported by thousands of swarming terraced houses.

---

[67]Dong Truc—Ngoc Liep Ecological City, Jina Architect Seoul, 16/12/2010, http://jina.co.kr/?p=118.

[68]Nguyen To Lang (2007).

[69]*Urban Design Guidelines* (2010), op.cit., p. 22.

[70]*Urban Design Guidelines* (2010), op.cit., p. 21.

Each new project shows the Bio/Eco label. Annihilating the very meaning of the concept of sustainability by oversupplying it. Cities will never be sustainable, eco-compatible or bio something else. Cities are artificial machines, human products and therefore imperfect artifacts. With the passing of time, we will come closer and closer to a level of perfection, as a function whose limit tends to infinity without ever reaching it.

For now, we are satisfied with a sea of concreteness without a real manifesto, perhaps, the absence of a manifesto is the manifesto itself. The new tropical metropolis of the twenty-first century should probably think about reinventing the concept of city, rather than importing models that are already widely tested and not really sustainable, even if they are sold as such. There are empty satellite cities in China. Entire lots for millions of deserted inhabitants, which are likely to be filled with difficulty. Sisters of the same real estate bubble that is currently passing through Hanoi. Urbanity is not just about building roads and buildings. Urbanity means shaping densities creatively, preparing territories for a type of colonization that is light, ever-changing, linear and compact, within which forms are recomposed according to need. Reversible and restraining, it should work with the ground in a mutual symbiotic relationship. How sustainable is a plan that would cost about 160 billion dollars to implement really? Where is our ability to judge the metropolis? Soils, densities and the malleable program are our only salvation. Territories explode and others shrink. One day even the expanding ones will stop and afterwards what will happen?

References

Doumer P (1905) L'indochine française (souvenir), Viubert et Nony Editeurs, Paris, p 123

Gourou P (1936) Les Paysans du delta tonkinois. Etudes de gèographie humaine. Les èdition d'art et d'histoire, Paris, p 570

Gourou P (1948) La civilisation du vegetal. Indonesie, 1, 5:389

Gregorio M (2009) On the edge: facing the urban transition in Hanoi's Western Suburbs. Preliminary Report, Hanoi, The Ford Foundation, p 10

Hanh LTH (2001) Improving privatisation of housing stock in Vietnam: a case study of Hanoi, Lund, Sweden (inedito), Retrieved May 6th 2005 from www.hdm.lth.se

Hebrard E (1928) L'urbanisme en Indochine. L'architecture XLI(2):40–41

Kleinen J (2005) Tropicality and topicality: Pierre Gourou and the genealogy of French colonial scholarship on rural Vietnam. Singap J Trop Geogr III(26)

Koolhaas R (1995) What ever happened to urbanism? In: S, M, L, XL, OMA with B. Mau, The Monicelli Press, New York, pp 959–971

Kwon YH (1997) Development plan and feasibility study on Hanoi New Town, Daewoo Construction Technology Report, p 191

Lai T (1998) Urban housing development and urban management—a sociological approach. Soc Sci p 1

Lamster M (2011) The grass Isn't always greener. Architect http://www.architectmagazine.com/planning/the-grass-isnt-always-greener-a-master-plan-for-h.aspx

Ledent J (2002) La population: évolution passée et développement futur. Hanoi: enjeux modernes d'une ville millénaire In: Charbonneau F, Hau D (eds) Montréal, Trames, pp 64–86

Logan SW (2000) Hanoi biography of a city. University of Washington Press, Seattle, p 169

Masson A (1929) Hanoi pendant la période héroique (1873–1888). Paris Libraire Orientaliste Paul Geunther, Paris, nota 1, p 20

Mathes M (1995) Drawn and quartered—Hanoi's ancient centre is feeling the pressure from outside and in. So, what can architects and officials do to stop the rot? In: Vietnam Economic Times, pp 24–25 citato da M. Weibel (2004), *Asien*, 92:39–40

Melnikov E (1986) Il palazzo dello scambio culturale di Hanoi. Arkhitektura SSSR, (2):104–107

Noury J (1992) L' Indochine en cartes postales. Avant l'hourigan 1900–1920. Publi-fusion Editeur, Paris, p 7

Pèdelahore C (1992) Hanoi miroir de l'architecture française outre-mer. Mardaga, Liege, Belgium, p 300

Perry C (1929) The Neighborhood Unit, Regional Survey of New York and its Environs. In: Neighborhood and community planning, vol VII, New York, p 186

Prokhorov A (1970) The Great Soviet Encyclopedia, 3rd edn (1970-1979), The Gale Group, Moscow

Quertamp F (2003) Ha Noi: Une péri-urbanisation paradoxale. Université de Bordeaux III—Michel de Montaigne, Bordeaux, pp 86–90

Ruble BA (1990) Leningrad: shaping a soviet city. Berkley Press, pp 89–91

Water Resources Planning and Management in a Changing Climate and Society

Andrea Castelletti, Matteo Giuliani, and Rodolfo Soncini-Sessa

Abstract Large water storage systems play a key role for securing water, energy, and food to the Vietnamese society now and in the future. These systems require robust and adaptive operations capable of adapting to hydroclimatic variability and managing multi-sector demands across multiple time scales, from daily operation to strategic river basin development. This is especially true for river basins with high intra-annual and inter-annual variability, such as monsoonal systems that need to cope with seasonal droughts while also preventing extremely damaging floods. In this chapter, we summarize ten years of research on the integrated and robust management of the Red River basin in Vietnam, where a multireservoir network serves as important source of hydropower production, multi-sectoral water supply, and flood protection for the capital city of Hanoi. We illustrate a decision-analytic framework to support dam operators and policy makers in operating the Red River basin reservoir network under present and future climate forcing and multi-sector water demand. Results show that conflicts across sectors are very likely to increase in the future, in particular exacerbating the food-flood tradeoff, and that dam re-operation and coordination are fundamental tools to mitigate the projected negative impacts of global change.

1 Introduction

Climate change and socio-economic development are projected to severely challenge water systems management worldwide (Trenberth 2011; Rodell et al. 2018), calling for flexible and timely adaptation of water systems operations to such uncertain and evolving conditions (Benson 2016). Traditionally, scenario-led approaches have been used as a basis for developing adaptation strategies, by describing the performance of water resource systems under a discrete set of global projections

A. Castelletti (✉) · M. Giuliani · R. Soncini-Sessa
Politecnico di Milano, Milan, Italy
e-mail: andrea.castelletti@polimi.it

(Wilby and Dessai 2010). However, the discrete nature of these projections, which focus mostly on climate drivers, challenges the use of top-down approaches for decision-making purposes, and, indeed, most research studies stop at the impact assessment stage (Brown et al. 2012). To better characterize the risks in the operations of water systems, recent works have shifted from scenario-led toward scenario-neutral approaches, which determine the vulnerability of the system when exposed to a wide range of plausible, uncertain scenarios whose probabilities are not known or widely agreed upon (Lempert 2002; Herman et al. 2015). Several studies have assessed the vulnerability of water reservoir operations by using synthetically generated inflow scenarios only (Whateley et al. 2014; Culley et al. 2016; Herman et al. 2016; Quinn et al. 2017). Yet, assessing the impacts of inflow uncertainty on water reservoir operations represents only one part of the equation.

In this chapter, we adopt a scenario-neutral approach to identify the major vulnerabilities for the management of the Red River basin in Vietnam with respect to interdependent climate and socio-economic drivers. In this basin, four major reservoirs are operated to protect the capital city Hanoi from flooding, while also providing the surrounding region with electric power and meeting multisectoral water demands for the agricultural and urban economies (Castelletti et al. 2012a). The core of our decision-analytic framework (Fig. 1) is the generation of climate scenarios by means
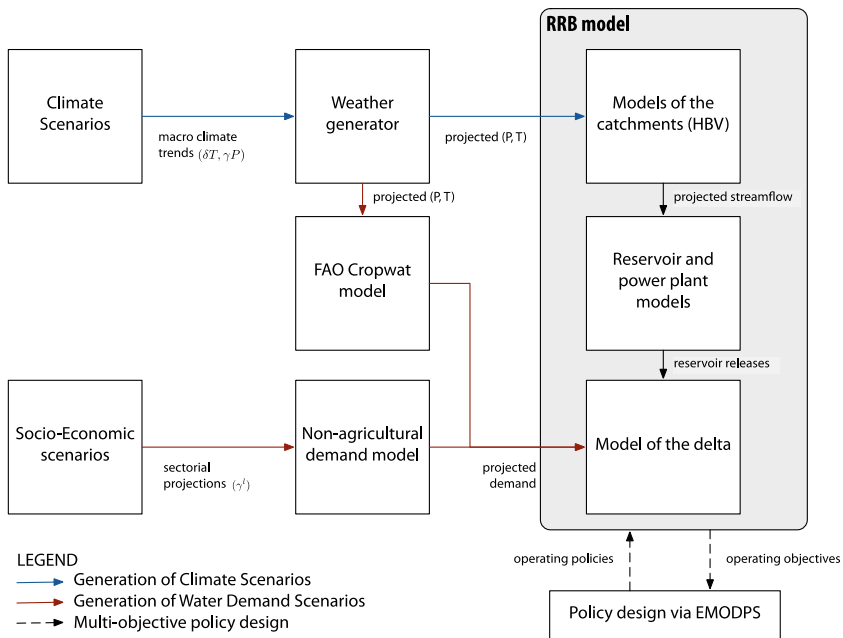

Fig. 1 Red River Basin decision-analytic framework. The blue and red solid lines shows the procedure used for the generation of climate and water demand scenarios, respectively; the dashed lines show the multi-objective policy design. Details about each component of the framework are provided in Sects. 2 and 3

of a semi-parametric weather generator combined with additive and multiplicative scaling factors. Given the uncertainty in how climate change will affect East Asian monsoon (Hijioka et al. 2014), this synthetic generation of climate exposures allows investigating a wide range of plausible changes in climate (Culley et al. 2016). Moreover, changes in intraannual variability of monsoonal hydrologic regimes may also challenge existing management strategies that strive to balance the competing intersts of flood protection in the wet season and water supply in the dry one (Quinn et al. 2018). The projected climate variables are used both as input to hydrological models for generating future streamflow and for projecting the agricultural water demand. In addition, the non-agricultural component of the water demand is synthetically generated using scaling factors to explore alternative plausible evolutions of the Vietnamese society. We then test the performance of a potential compromise solution, which successfully balances the competing interests in the basin under historical conditions, under a large ensemble of 5600 future scenarios in order to identify the main sources of vulnerability for the basin.

2 The Red River Basin

The Red River basin (RRB) is a transnational river basin located in South-East Asia, covering an area of 169,000 km^2 shared by China, Laos, and Vietnam. The Vietnamese part covers 25 provinces and cities with the total area of 86,700 km^2, accounting for 51.3% of the basin area. The RRB is formed by three main tributaries, namely Da River, Thao River, and Lo River as illustrated in Fig. 2. Among the tributaries, the Da River is the largest, contributing about 49% to the total flow entering in the delta region. The delta covers 11 provinces and cities, including the capital city of Hanoi. With the total population living in the delta exceeding 20 million people (Devienne 2006), this region has the highest population density in Vietnam. More than half of the delta is less than 2 m above mean sea level and is protected from flooding and storm surges by a system of river and sea dykes (Vinh Hung et al. 2010). The region is mainly cultivated with rice and the irrigation relies on a combination of gravity and pumping systems. Agriculture is the most important water user accounting for the 58% of the total water demand and involves around 50% of the local workers. The river delta is the second largest rice production area of Vietnam, which is, in turn, the third largest rice exporter in the world (USDA 2019). The other consumption sectors include domestic use, aquaculture, industries, and livestocks. For these reasons, the delta has an important role in terms of food security and socio-economic development.

The RRB is characterized by a sub-tropical monsoon climate composed of a wet and a dry season. The wet season lasts from May to October, cumulating 85–95% of the total annual rainfall (Le et al. 2007). The available water in the basin is quite abundant but, unfortunately, it is not evenly distributed in terms of space and time: the average river flow entering the delta region measured at Son Tay station varies between 8000 m^3/s during the monsoon peak and 1500 m^3/s in the dry months. This
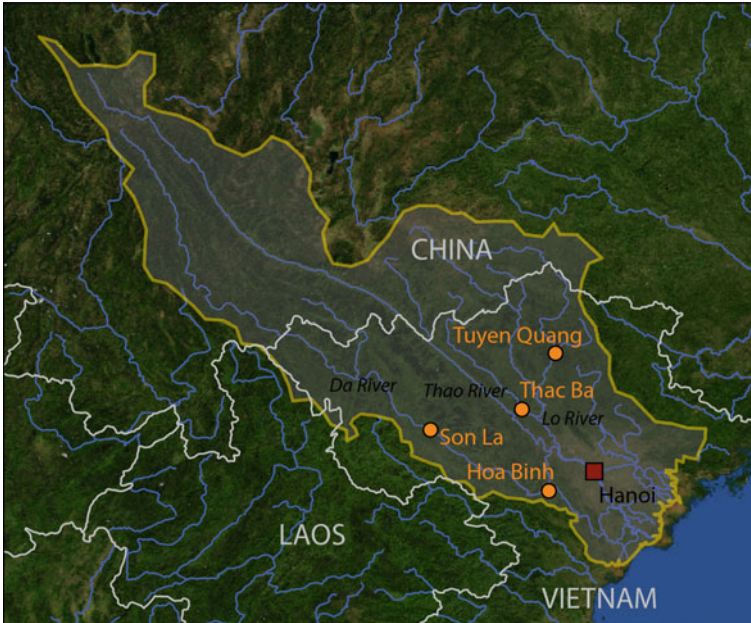
Fig. 2 Map of the Red River basin

Table 1 Storage and power capacity of reservoirs in Red River basin

Reservoir	Storage [Mm^3]	Power capacity [MW]	River
Son La	10.23	2400	Da River
Hoa Binh	7.68	1920	Da River
Tuyen Quang	3.24	342	Gam River
Tach Ba	2.08	120	Chay River

uneven distribution causes serious effects on livelihoods of local people such as water deficit for irrigation and domestic uses. Since the '70s, several water reservoirs have been built in the basin. The four largest dams have a total installed power capacity of 4782 MW (see Table 1), but they also contribute in terms of flood control and multi-sectoral water supply to the delta region (Giuliani et al. 2016a).

2.1 Model and Data

This section provides an overview of the RRB model developed within the IMRR Project (Integrated and sustainable water Management of Red-Thai Binh River

System in a changing climate)[1] funded by the Italian Cooperation under the Italian Ministry for Foreign Affairs. The model is composed of a combination of conceptual and data-driven models assuming a time-step of 24 h for modeling and decision-making processes.

The five river basins (i.e., Da, Lo, Gam, Chay, and Thao) are described with conceptual HBV models. The simulated daily discharges are obtained as a function of daily temperature and precipitation related to each sub-catchment. The precipitation data used for the calibration of the sub-catchment models are produced within the APHRODITE project (Research Institute for Humanity and Nature, Japan) (Yatagai et al. 2012).

The dynamics of the four main reservoirs (i.e., Son La, Hoa Binh, Tach Ba, and Tuyen Quang) is described by the mass balance equations of the water volumes stored in each reservoir (j = SL, HB, TB, TQ), i.e.

$$s_{t+1}^{SL} = s_t^{SL} + q_{t+1}^{D} - r_{t+1}^{SL} - E_t^{SL}$$

$$s_{t+1}^{HB} = s_t^{HB} + q_{t+1}^{D,lat} + r_{t+1}^{SL} - r_{t+1}^{HB} - E_t^{HB}$$

$$s_{t+1}^{TB} = s_t^{TB} + q_{t+1}^{C} - r_{t+1}^{TB} - E_t^{TB}$$

$$s_{t+1}^{TQ} = s_t^{TQ} + q_{t+1}^{G} - r_{t+1}^{TQ} - E_t^{TQ}$$

where t is the time, q_{t+1}^{i} with $i = D, C, G$ are the water volumes of the Da, Chay, and Gam rivers flowing into Son La, Tach Ba, and Tuyen Quang reservoirs, respectively, in the time interval $[t, t+1)$; $q_{t+1}^{D,lat}$ is the lateral contribution of the Da River catchment, which is not captured by Son La and flows into the Hoa Binh reservoir; E_t^{j} represents the mean daily losses for evaporation for each reservoir. The daily release of the j-th reservoir is defined as $r_{t+1}^{j} = f(u_t^{j}, \cdot)$ to describe the nonlinear, stochastic relation between the decisions u_t^{j}, and the actual released volume r_{t+1}^{j}. Indeed, between the time t at which the decision u_t^{j} is taken and the time $t+1$ at which the release r_{t+1}^{j} it determines is completed, the inflow is affecting the reservoir storage, and the actual release may not be equal to the decision, for instance because of the activation of the spillways (Soncini-Sessa et al. 2007). In the adopted notation, the time subscript of a variable indicates the instant when its value is deterministically known.

Each reservoir is connected to a multi-turbines hydropower plant. In these plants, each turbine can be independently activated with the desired flow because each of them is fed by an independent penstock controlled by a valve. The daily maximum energy production of each plant can be estimated as a function of the daily water volume released from the reservoir, the reservoir level, and the specific features of the installed turbines such as the hydraulic capacity or the head loss functions. This

---

[1]https://xake.elet.polimi.it/imrr/.

relationship is estimated assuming that the operator chooses the optimal allocation of the releases among the turbine in different hours of the day.

The reservoir releases and the natural streamflow of Thao and Lo rivers flow downstream and reach the city of Hanoi and the irrigation districts in the river delta. The river delta was originally described by a 1D hydrodynamic model (MIKE11) that considered the water flow in 907 rivers and canals forming the delta river network. It includes also the description of structures within the Delta, such as bridges and pumping stations. This model is based on the observations measured in 5000 cross sections on 4200 km of canals, registered during two monitoring campaigns in 1999–2000 and 2009–2012. MIKE11 provides a very accurate and spatial distributed description of the delta dynamics but is highly computational expensive (i.e., it takes 2 days for simulating 16 years). Due to the impossibility of using MIKE11 for the policy design process, that requires a high number of simulations, dynamic emulators were developed to simplify the canals dynamics and consequently reduce the computational time (Castelletti et al. 2011,2012b). The emulation model is constituted by recursive Artificial Neural Networks (ANNs) trained over the results of the De Saint Venant equations and it reproduces the outputs needed for the objective calculation. The resulting ANN models provide approximate values of the spatially distributed volume of water available in the irrigation canals and the water level in Hanoi.

2.2 Operating Objectives and Policy Design

According to a direct interaction with the stakeholders, three main water-related interests were identified and associated to the following operating objectives' formulations (Giuliani et al. 2016a):

- Hydropower production (J^{HP}): the daily average energy production in the system, to be maximized;
- Flood control (J^F): the daily average flood damage in Hanoi, estimated by the dimensionless non-linear cost function reported in Fig. 3 defined by direct stakeholders and local expert's consultation, which depends on the corresponding water level in Hanoi estimated by the ANN model;

- Water supply (J^S): the daily average squared water deficit between the estimated volumes of water available in the irrigation canals, estimated by the ANN model, with respect to the total water demand of the Red River delta (Fig. 4). The adopted quadratic formulation aims to penalize severe deficits in a single time step, while allowing for more frequent, small shortages (Hashimoto et al. 1982).

The operating policies determining the coordinated operations of the four modeled reservoirs are formulated as Gaussian radial basis functions (RBFs) mapping a vector of policy inputs I_t to reservoir releases. The operating policies depend on the storage of the four reservoirs, along with the day of the year (d_t) to represent the seasonality
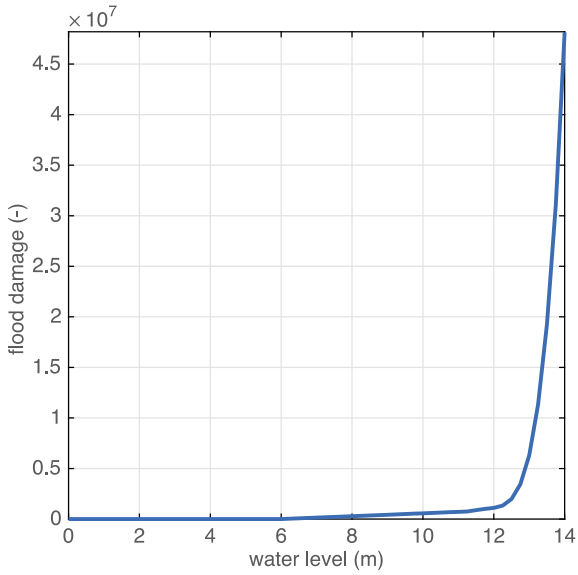
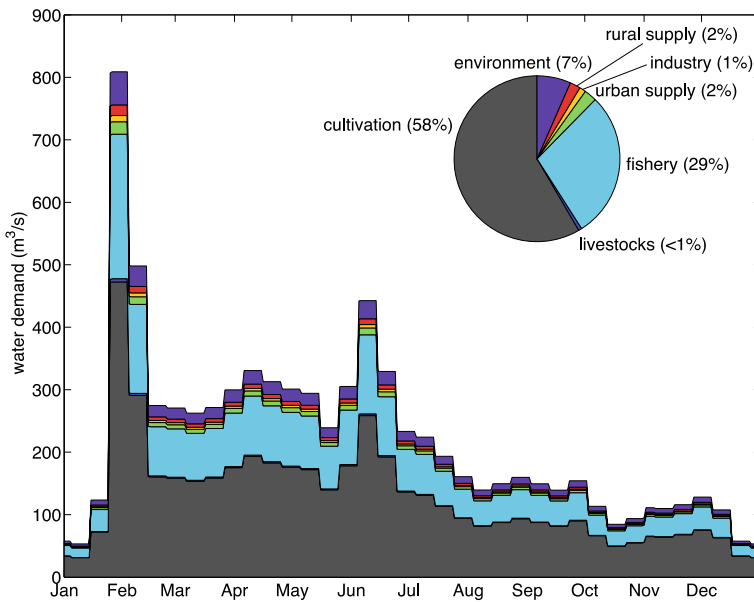**Fig. 3** Non-linear cost function estimating the flood damage in Hanoi

Fig. 4 Multisectoral water demand in the Red River Delta

of both inflows and water demands, and the total previous day inflow (Q_t) to inform about the current hydrologic conditions. Formally, the RBF representation of daily release policies for the j-th reservoir is defined as

$$u_t^j = \sum_{i=1}^{N} w_i \varphi_i(I_t)$$

where $N = 11$ is the number of RBFs $\varphi(\cdot)$ and w_i is the weight of the i-th RBF used in the convex combination (i.e., weights are formulated such that they sum to one and are nonnegative). The single RBF is defined as follows:

$$\varphi_i(I_t) = \exp\left[-\sum_{k=1}^{M} \frac{((I_t)_k - c_{k,i})^2}{b_{k,i}^2} \right]$$

where M is the number of policy inputs and $\boldsymbol{c}_i, \boldsymbol{b}_i$ are the M-dimensional center and radius vectors of the i-th RBF, respectively. The total number of policy parameters is equal to $N(2M+1) = 176$.

The operating policies are then designed via Evolutionary Multi-Objective Direct Policy Search (Giuliani et al. 2016b), a Reinforcement Learning approach that combines direct policy search, nonlinear approximating networks, and hierarchically parallelized multi-objective evolutionary algorithms (Giuliani et al. 2018). The advantage of using EMODPS against other optimal control methods (e.g., Castelletti et al. 2013) is the possibility of designing the coordinated control across multiple reservoirs alleviating the well-known curse of dimensionality (e.g., Biglarbeigi et al. 2018) and of estimating an approximation of the Pareto front in a single run of the algorithm, which supports the exploration of the multidimensional tradeoffs between the conflicting objectives (e.g., Giuliani et al. 2014). Moreover, EMODPS allows enlarging the information used for conditioning operational decisions (Giuliani et al. 2015; Giuliani and Castelletti 2019). Specifically, the set of Pareto approximate operating policies p^* are generated by solving the following multi-objective optimal control problem:

$$p_\theta^* = arg \min_{p_\theta} \boldsymbol{J} = | - J^{HP}, J^F, J^S |$$

where $\theta \in \Theta$ and the problem is constrained by the dynamics of the system. To perform the optimization, we use the Multi-Master Borg MOEA (Hadka and Reed 2015), a hierarchically parallelized self-adaptive version of the Borg MOEA that transforms the search dynamics of the algorithm through communication across multiple parallel optimizations, increasing the algorithmic reliability for finding favorable solution sets (Reed and Hadka 2014). Each optimization was run for 1 million function evaluations. To improve solution diversity and avoid dependence on randomness, the final set of Pareto optimal policies is obtained as the set of nondominated solutions identified from the results of 50 random optimization trials. In total,

we run 50 million simulations requiring around 350,000 computing hours. Each optimization run was parallelized over 512 processing cores of the Texas Advanced Computing Center Stampede Cluster (https://www.tacc.utexas.edu/stampede).

3 Future Scenario Generation

3.1 Climate Scenarios

Climate scenarios were generated by implementing a two-step approach: first, synthetic time series of daily precipitation and temperature for each sub-basin are generated using a hybrid weather generator; then, macro climate trends are introduced by applying scaling factors to the synthetically generated scenarios. The weather generator is composed of a Markov chain modeling the daily average precipitation occurrence over the entire RRB, coupled with a k-Nearest Neighbor bootstrap to simulate precipitation and temperature in each sub-basin (Da, Gam, Lo, Chay, and Thao) preserving the spatial correlation among the multiple sites. In the Markov model, three possible states are considered: dry (D), wet (W), and extremely wet (EW). A dry day occurs when the precipitation is less than 0.33 mm, an extremely wet day when the precipitation is greater than the 80th percentile of daily records for the considered month, and a wet day when the amount of rain is between the dry and extremely wet thresholds. This model was calibrated using historical observations by estimating unconditional and transition probabilities for the three states (see Table 2).

A 40-years Markov chain was generated from the estimated monthly probabilities. The kNN bootstrap is subsequently applied to generate temperature and precipitation sequences in each basin: given the states of two consecutive days (d_1, d_2) in the Markov model (e.g., D,W), k neighbors of d_1 are selected as the historical observations in a 7-day window centered in the calendar day of d_1 that are similar to (d_1, d_2) (i.e., have the same state transition from dry to wet); a single neighbor is finally sampled according to the probability determined by the kernel estimator proposed in Lall and Sharma (1996), where the probability of selecting one neighbor decreases according to its distance from d_1.

The synthetic scenarios are then altered using additive factors for temperature and multiplicative factors for precipitation. Specifically, we applied the following scaling factors: $\delta T = [-2, -1, …, +4, +5]$ °C and $\gamma P = [0.8, 0.9, …, 1.5, 1.6]$, which introduce large variability in both positive and negative directions. The combination of these factors yields an ensemble of 56 possible climate futures, where each scenario is composed of a 40-years long, daily trajectory of temperature and precipitation for each sub-basin. It is worth mentioning that the combined application of the weather generator and the scaling factors produced climate scenarios that differ from the historical conditions not only in the mean values, but also in terms of intra- and inter-annual variability.

Table 2 Unconditional and transition probabilities of dry (D), wet (W), and extremely wet (EW) days

Month	P(D)	P(W)	P(EW)	P(D|D)	P(D|W)	P(D|EW)	P(W|D)	P(W|W)	P(W|EW)	P(EW|D)	P(EW|W)	P(EW|EW)
Jan	0.692	0.170	0.138	0.816	0.128	0.057	0.315	0.662	0.023	0.532	0.460	0.008
Feb	0.692	0.170	0.138	0.475	0.201	0.324	0.094	0.600	0.306	0.067	0.659	0.274
Mar	0.558	0.303	0.138	0.530	0.000	0.470	0.649	0.000	0.351	0.685	0.000	0.315
Apr	0.285	0.572	0.143	0.777	0.196	0.027	0.420	0.580	0.011	0.620	0.372	0.007
May	0.108	0.747	0.145	0.413	0.153	0.434	0.028	0.535	0.437	0.144	0.690	0.166
Jun	0.043	0.823	0.134	0.545	0.000	0.455	0.774	0.000	0.226	0.536	0.000	0.464
Jul	0.012	0.839	0.149	0.728	0.227	0.045	0.000	1.000	0.000	0.734	0.259	0.007
Aug	0.031	0.823	0.146	0.409	0.223	0.368	0.012	0.829	0.159	0.311	0.299	0.390
Sep	0.107	0.751	0.142	0.548	0.000	0.452	0.670	0.000	0.330	0.500	0.000	0.500
Oct	0.228	0.621	0.151	0.532	0.450	0.018	0.324	0.676	0.000	0.860	0.115	0.024
Nov	0.489	0.413	0.098	0.205	0.363	0.431	0.023	0.720	0.257	0.477	0.207	0.316
Dec	0.722	0.161	0.118	0.669	0.000	0.331	0.606	0.000	0.394	0.525	0.000	0.475

3.2   Water Demand Scenarios

The water demand from the agricultural sector (w^A) represents the largest contribution to the total demand of the Red River delta. Differently from other works that generally generate water demand scenarios independently from the climate conditions (e.g., Herman et al. 2014; Quinn et al. 2018), we estimated future agricultural demands by simulating the FAO CROPWAT model fed by climate data (i.e., monthly total precipitation, average temperature, humidity, wind speed, and solar radiation), cropping pattern, and cultivated area. The model was run considering 13 distinct regions of the delta, each associated to a different meteorological station, using the historical crop distribution that includes mainly rice, maize, potatoes, and oranges. Future agricultural demands were generated by running the model under each climate scenario, thus producing 56 different scenarios of irrigation requirements.

The non-agricultural component of the water demand (w^{NA}) is estimated as the product of a given water consumption sector (e.g., number of people) and its corresponding standard water use (e.g., liters/person, see Tables 3 and 4). For each of the 13 sub-regions of the delta, 8 different sectors were modeled (see Table 5). An ensemble of future demands was then generated for each sector by sampling 100 scaling factors γ^l ($l = 1, …, 8$) generated via Latin Hypercube Sampling. To preserve the spatial correlation across the districts, the same scaling factor for each sector is applied to all the districts, thus assuming a homogeneous pattern of socio-economic changes.

The combination of 56 climate-dependent scenarios of irrigation requirements with 100 non-agricultural demands yields a total of 5600 water demand scenarios.

Table 3 Standard of water use for different sectors (except aquaculture)

Water consumption sector	Standard of water use
Industries (*Ind*)	128 m^3/day/ha
Urban population (*UrbP*)	343 l/day/person
Town population (*TownP*)	257 l/day/person
Rural population (*RurP*)	115 l/day/person
Cows and Buffalos (*Cow*)	65 l/day/animal
Pigs (*Pig*)	15 l/day/animal
Poultry (*Poul*)	1 l/day/animal

Table 4 Standard of water use for aquaculture (aquac) [m^3/month/ha]

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0	1333	1713	1713	1714	1714	0	0	0	1714	1714	1714

Table 5 Historical water consumption sectors across the 13 regions of the Red River delta

Region	Aquaculture [ha]	Industries [ha]	Urban pop [people]	Town pop [people]	Rural pop [people]	Cows [animals]	Pigs [animals]	Poultry [animals]
Tich Thanh Ha	3336	2426	0	129,500	1,226,800	125,155	690,968	6,395,804
Nhue	5908	1574	2,062,700	89,843	1,626,750	44,411	591,790	5,787,805
Righ-day	288	30	0	2232	43,577	2772	25,242	277,873
6 Pumping Station	5987	1003	246,125	40,745	958,638	57,172	529,948	3,920,507
North of Ninh Binh	442	148	0	6374	178,677	21,965	84,446	647,620
South of Ninh Binh	5021	732	87,002	65,547	626,315	43,903	287,851	2,746,907
Central Nam Dinh	4174	70	0	60,483	564,279	11,846	240,309	1,777,679
South of Nam Dinh	8325	0	0	58,178	733,574	7207	288,558	2,134,604
North of Thai Binh	5794	292	20,800	48,000	976,800	45,360	601,400	5,079,400
South of Thai Binh	4004	994	83,200	20,000	712,200	24,240	425,600	2,882,600
Downstream of Thai Binh	21,594	2965	559,800	342,306	1,846,021	50,747	793,580	8,311,083
Bac Hung Hai	2886	2451	91,541	100,919	848,998	57,684	377,910	3,943,000
North of Dug River	14,254	2811	415,357	191,218	2,320,015	96,469	1,104,534	11,605,567

4 Results and Discussion

4.1 Policy Performance Over History

The aim of this section is the tradeoff analysis of the system performance, measured in terms of hydropower production J^{HP}, flood control J^F, and water supply J^S under historical conditions. The performance of each solution is estimated via simulation of the RRB model over the historical horizon 1990–2010 using observed streamflows and water demands. Figure 5 illustrates the performance of the designed Pareto optimal operating policies, where hydropower production and flood damages are plotted on the primary axes, while the water supply deficit is represented by the color of the circles. The black arrows indicate the directions of increasing preference, with the ideal solution that would be represented by a blue circle in the top-left corner of the figure, while each marker represents a different tradeoff between the three objectives.

Results show that a clear conflict exists between the maximization of hydropower production and the prevention of flood damages. The best performance in terms of J^F can be indeed obtained by maintaining the reservoirs' level as low as possible to buffer the monsoon peak of the inflows, but this strategy negatively impacts on the energy production that would require constant releases equal to the turbine capacity and high hydraulic head. This opposite strategy that maximizes hydropower production
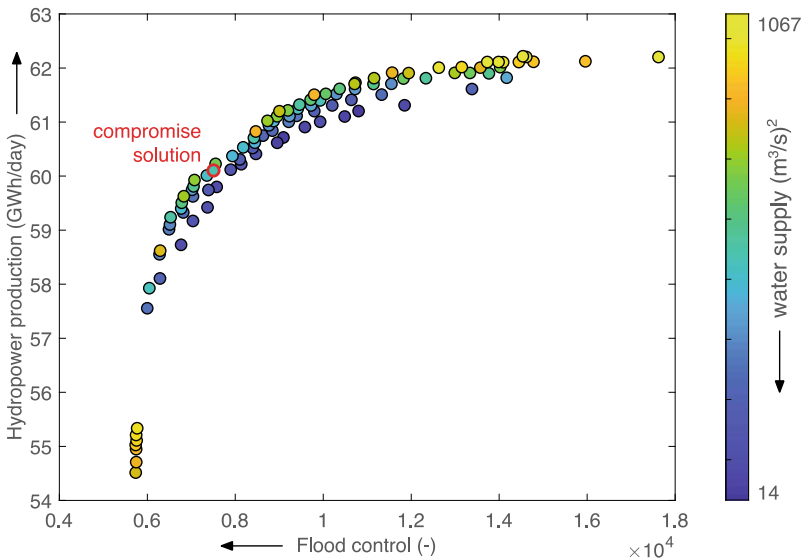
Fig. 5 Performance of the Pareto optimal policies over history (1990–2010), where hydropower production and flood damages are plotted on the primary axes, while the circles' color represents water supply deficit. The circle marked with a red edge represents the selected compromise solution

also increases the vulnerability of the water supply in the delta and attains very high values of water supply deficit (i.e., yellow circles). The conflict between flooding and water supply is instead weaker except for the best flood solutions in the bottom-left corner of the figure, because the drawdown of the reservoirs' level during the monsoon season is indirectly making available large water volumes in the Red River delta. Overall, the best solutions in terms of water supply requires an ad hoc operation of the reservoirs to satisfy the water demand peak in February, resulting in solutions (i.e., dark blue circles) that appear dominated in the J^F, J^{HP} projection of the objective space.

Overall, when evaluated over historical hydroclimatic and water demand conditions, the set of Pareto optimal operating policies illustrated in Fig. 5 provides a rich context for understanding the complex management tradeoffs and dynamics in the RRB and represents a valuable tool for supporting the identification of candidate compromise solutions. A potentially fair compromise solution for balancing the three objectives is the one identified by the circle marked with the red edge in Fig. 5. This solution has been selected according to the criterion of the minimum distance from the Utopia point (Eschenauer et al. 1990), which identifies the absolute optima of the three objectives; its performance is equal to $J^{HP} = 60.1$ (GWh/day), $J^F = 7505$ $(-)$, $J^S = 25$ (m^3/s)^2, i.e., better than 70% of the solutions in the Pareto optimal set in terms of hydropower production, 72% in terms of flood damages, and 49% in terms of water supply.

4.2   Future Scenarios

The ensemble of future scenarios generated according to the methodology described in Sect. 3 is illustrated in Fig. 6, including the projected streamflow (panel a), the projected agricultural demand (panel b), and the projected non-agricultural demand (panel c). This ensemble is composed of 5600 members, obtained as the Cartesian product of 56 climate scenarios associated to the projected streamflows and agricultural demands, and 100 socio-economic scenarios determining the non-agricultural component of the water demand.

Streamflows and agricultural demands are illustrated in heat-maps reporting the temperature and precipitation changes (δT, γP) on the x- and y-axes, respectively. Figure 6a resents a clear vertical gradient with streamflows that increase primarily moving from the bottom to the top part of the figure following the trend of increasing precipitation. Temperature, instead, appears less relevant even though we can observe that the maximum future streamflows (dark blue color) are expected in the top-left corner of the figure (i.e., high precipitation-low temperature); vice versa, the minimum streamflows (light green color) are expected in the bottom-right corner of the figure for low precipitation and high temperature. This pattern is inverted for the projected agricultural demand reported in panel (b): the lowest demand values (white color) are obtained for low temperature and high precipitation in the top-left
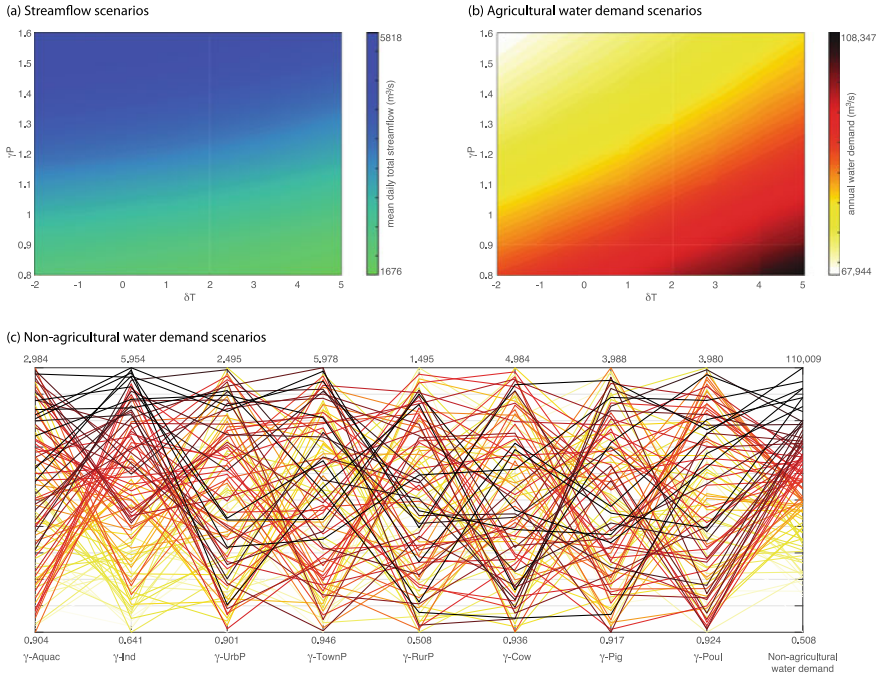
Fig. 6 Future scenarios of streamflow (panel a), agricultural demand (panel b), and non-agricultural demand (panel c). Streamflow and agricultural demand are illustrated through heat maps depending on the projections of temperature (x-axis) and precipitation (y-axis). Non-agricultural demand is illustrated in a parallel axes plot, where the first 8 vertical axes are the scaling factors for the different sectors, and the last axis on the right is the resulting demand; each line is one scenario shaded according to the demand value

corner of the figure, and increases following an oblique gradient moving towards high temperature and low precipitation scenarios (dark red) in the bottom-right corner.

The non-agricultural demand scenarios depend on the projections of 8 sectorial demands and are illustrated in Fig. 6c: this parallel axis representation shows each scenario as a line crossing the first eight vertical axes at the values of the scaling factors adopted for each water demand sector, with the last axis on the right reporting the resulting demand. Each line (scenario) is shaded according to the demand value. Figure 6c allows recognizing aquaculture and industries as the two sectors that are responsible for the highest projected demands, with most dark red lines that intersect the first two vertical axes in their upper part. Despite the maximum sampled values of γ–*TownP* and γ–*Cow* (reported on the top of the axes) are two out of the three maximum scaling factors, town population and number of cows are instead less influencing the final demand. We can indeed observe mixed patterns of scenarios covering the corresponding axes, meaning that high values of demands (dark red lines) are obtained for either high or low scaling factors.

4.3 Vulnerability Analysis

To evaluate the potential vulnerability of the compromise solution identified in Fig. 5 with respect to changes in climate and socio-economic conditions, we run a Monte Carlo simulation of the RRB model under this operating policy over the future scenarios discussed in the previous section. Figure 7 provides a synthesis of the outputs of these 5600 simulations by illustrating through a parallel axis plot the variability of the system performance in terms of hydropower production (first vertical axis), flood control (second vertical axis), and water supply (seventh vertical axis) across the scenarios. The axes are oriented so that the direction of preference is upward. To help the interpretation of the results, the figure also shows the main underlying climate and socio-economic drivers: temperature and precipitation changes (δT, γP), mean daily total streamflow (Q^{tot}), agricultural (w^A) and non-agricultural (w^{NA}) water demands. Each line hence represents a single simulation of the compromise policy that crosses the vertical axes at the values of its corresponding performance or at the values of the climate and socio-economic drivers determining the scenario used for the simulation; all lines are shaded according to their performance on the water supply objective, with blue lines representing small deficit values (i.e., good performance).

Figure 7 shows that, not surprisingly, hydropower production and flood control performance are strongly impacted by water availability, with the best (worst) performance in terms of J^{HP} (J^F) attained in the scenarios characterized by the highest values of Q^{tot} (and hence high γP as illustrated also in Fig. 6a). In particular, the performance of the selected compromise solution degrades in terms of hydropower production in 50% of the generated scenarios, with a worst-case performance equal to 24.4 GWh/day (i.e., 40% of the production over historical conditions) and a best-case
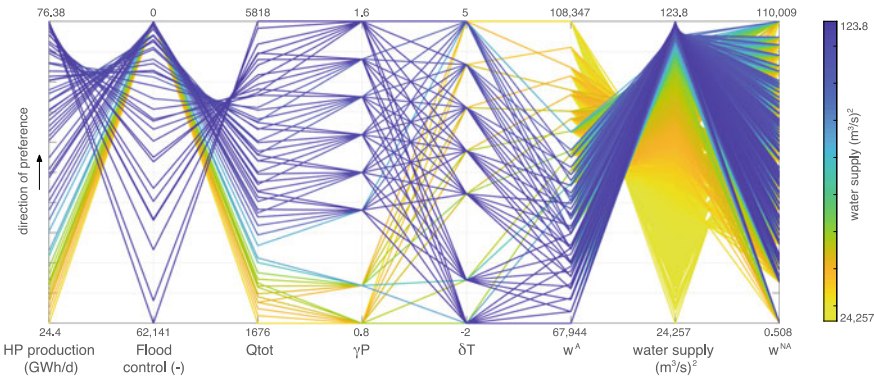
Fig. 7 Performance of the compromise solution selected in Fig. 5 across the generated 5600 scenarios. Hydropower production, flood damages, and water supply deficit are plotted on the first, second, and seventh vertical axes of the parallel axes plot, where each line is shaded according to the water supply performance. The other axes illustrate the main climate and socio-economic drivers for each scenario

performance equal to 76.4 GWh/day. In terms of flood control, the selected solution degrades its performance in 37.5% of the scenarios, including an 800% higher flood damage in the worst-case scenario.

Water supply is instead impacted by both water availability and water demands. Small water supply deficits (blue lines) are indeed obtained for scenarios with low agricultural demands and high-water availability represented by both Q^{tot} and γP axes. At the same time, good performance is obtained for all the considered temperature changes as well as most of the non-agricultural water demands. Yet, it is important observing that none of the 5600 scenarios allows meeting the performance of the selected compromise solution over historical conditions. This large vulnerability of the water supply sector can be explained by the projected growth of the non-agricultural demands which, with the exception of aquaculture, generates an increase in the total demand during the entire year; this component of the water demand is instead less important under historical conditions (see Fig. 4), when the irrigation demands peaks drive the allocation of the available water over the year. This challenge might be mitigated by re-optimizing the operating policy to better match the new pattern of water demand, even though this is likely negatively impacting on hydropower production and flood control, ultimately exacerbating the historical tradeoffs discussed in Fig. 5.

5   Conclusions

The aim of this chapter is assessing the vulnerability of the Red River basin under changing climate and society. To address this problem, we synthetically generated an ensemble of 5600 scenarios including changes in precipitation and temperature as well as in water demands. A semi-parametric weather generator combined with additive and multiplicative scaling factors was used to generate synthetic time series of temperature and precipitation including variations in the mean and in the inter-annual pattern while keeping the seasonal characteristics of the historical records. Moreover, the spatial correlation has been preserved among the Red River tributaries. These climate variables were then used as inputs to both the HBV models that produce projections of streamflows and the FAO CROPWAT model that estimates future agricultural water demands. The non-agricultural component of the water demand was generated by applying scaling factors to the historical water consumption of aquaculture, industries, population, and livestock to reflect plausible evolution of the Vietnamese society. After identifying a potential compromise solution able to balance the competing interests of hydropower production, flood control, and water supply under historical conditions, we tested its performance under projected conditions. Results show that in most of the generated scenarios the selected solution suffers a degradation of performance compared to the historical one, particularly in terms of water supply reliability. This result is due to the significant increase of the projected water demand that is largely driven by the fast-developing Vietnamese

society, suggesting the need of accounting for future demographic and economic changes when planning robust adaptation strategies.

References

Benson R (2016) Reviewing reservoir operations: can federal water projects adapt to change. Columbia J Environ Law 42:353

Biglarbeigi P, Giuliani M, Castelletti A (2018) Partitioning the impacts of streamflow and evaporation uncertainty on the operations of multipurpose reservoirs in Arid Regions. J Water Resour Plan Manag 140

Brown C, Ghile Y, Laverty M, Li K (2012) Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour Res 48

Castelletti A, Antenucci J, Limosani D, Quach Thi X, Soncini-Sessa R (2011) Interactive response surface approaches using computationally intensive models for multiobjective planning of lake water quality remediation. Water Resour Res 47

Castelletti A, Galelli S, Ratto M, Soncini-Sessa R, Young P (2012a) A general framework for dynamic emulation modelling in environmental problems. Environ Model Softw 34:5–18

Castelletti A, Pianosi F, Quach X, Soncini-Sessa R (2012b) Assessing water reservoirs management and development in Northern Vietnam. Hydrol. Earth Syst. Sci. 16:189–199

Castelletti A, Pianosi F, Restelli M (2013) A multiobjective reinforcement learning approach to water resources systems operation: Pareto frontier approximation in a single run. Water Resour Res 49

Culley S, Noble S, Yates A, Timbs M, Westra S, Maier H, Giuliani M, Castelletti A (2016) A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate. Water Resour Res 52:6751–6768

Devienne S (2006) Red river delta: Fifty years of change. Moussons, Recherche en sciences humaines sur l'Asie du Sud-Est (pp. 255–280)

Eschenauer H, Koski J, Osyczka A (1990) Multicriteria design optimization: procedures and applications. Springer, Berlin, Heidelberg, NewYork

Giuliani M, Anghileri D, Vu P, Castelletti A, Soncini-Sessa R (2016a) Large storage operations under climate change: expanding uncertainties and evolving tradeoffs. Environ Res Lett 11

Giuliani M, Castelletti A (2019) Data-driven control of water reservoirs using El Nino Southern Oscillation indexes. In: Proceedings of the 2019 ieee international conference on environment and electrical engineering and 2019 IEEE industrial and commercial power systems Europe (EEEIC/I CPS Europe), pp 1–5

Giuliani M, Castelletti A, Pianosi F, Mason E, Reed P (2016b) Curses, tradeoffs, and scalable management: advancing evolutionary multi-objective direct policy search to improve water reservoir operations. J Water Resour Plan Manag 142.

Giuliani M, Herman J, Castelletti A, Reed P (2014) Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. Water Resour Res 50:3355–3377

Giuliani M, Pianosi F, Castelletti A (2015) Making the most of data: an information selection and assessment framework to improve water systems operations. Water Resour Res 51:9073–9093

Giuliani M, Quinn JD, Herman JD, Castelletti A, Reed PM (2018) Scalable multiobjective control for large-scale water resources systems under uncertainty. IEEE Trans Control Syst Technol 26:1492–1499

Hadka D, Reed PM (2015) Large-scale parallelization of the Borg multi-objective evolutionary algorithm to enhance the management of complex environmental systems. Environmental Modelling & Software 69:353–369

Hashimoto T, Stedinger J, Loucks D (1982) Reliability, resilience, and vulnerability criteria for water resource system performance evaluation. Water Resour Res 18:14–20

Herman JD, Reed PM, Zeff HB, Characklis GW (2015) How should robustness be defined for water systems planning under change? J Water Resour Plan Manag 141

Herman JD, Zeff HB, Lamontagne J, Reed PM, Characklis GW (2016) Synthetic drought scenario generation to support bottom-up water supply vulnerability assessments. J Water Resour Plan Manag 142

Herman JD, Zeff HB, Reed PM, Characklis GW (2014) Beyond optimality: multistakeholder robustness tradeoffs for regional water portfolio planning under deep uncertainty. Water Resour Res 50:7692–7713

Hijioka Y, Lin E, Pereira JJ, Corlett R, Cui X, Insarov G et al (2014) Chapter 24: Asia, Working Group II contribution to the IPCC Fifth Assessment Report Climate Change

Lall U, Sharma A (1996) A nearest neighbor bootstrap for resampling hydrologic time series. Water Resour Res 32:679–693

Le TPQ, Garnier J, Gilles B, Sylvain T, Van Minh C (2007) The changing flow regime and sediment load of the Red River, Viet Nam. J Hydrol 334:199–214

Lempert RJ (2002) A new decision sciences for complex systems. Proc Natl Acad Sci 99:7309–7313

Quinn JD, Reed PM, Giuliani M, Castelletti A, Oyler J, Nicholas R (2018) Exploring how changing monsoonal dynamics and human pressures challenge multireservoir management for flood protection, hydropower production, and agricultural water supply. Water Resour Res 54:4638–4662

Quinn JD, Reed PM, Giuliani M, Castelletti A (2017) Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems. Water Resour Res 53:7208–7233

Reed PM, Hadka D (2014) Evolving many-objective water management to exploit exascale computing. Water Resour Res 50:8367–8373

Rodell M, Famiglietti J, Wiese D, Reager J, Beaudoing H, Landerer F, Lo M (2018) Emerging trends in global freshwater availability. Nature 557(7707):651

Soncini-Sessa R, Castelletti A, Weber E (2007) Integrated and participatory water resources management: theory. Elsevier, Amsterdam, NL

Trenberth K (2011) Changes in precipitation with climate change. Climate Res 47:123–138

USDA (2019) Grain: World markets and trade. Technical Report United States Department of Agriculture (USDA), Foreign Agricultural Service

Vinh Hung H, Shaw R, Kobayashi M (2010) Flood risk management for the riverside urban areas of Hanoi: the need for synergy in urban development and risk management policies. Disaster Prev Manag: An Int J 19:103–118

Whateley S, Steinschneider S, Brown C (2014) A climate change range-based method for estimating robustness for water resources supply. Water Resour Res 50:8944–8961

Wilby RL, Dessai S (2010) Robust adaptation to climate change. Weather 65:180–185

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A (2012) APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bull Am Meteorol Soc

GNSS-Based Solutions for Road Applications in Vietnam

Hiep Van Hoang, Thuan Dinh Nguyen, Tung Hai Ta, Gabriella Povero, and Micaela Troglia Gamba

Abstract Navigation Satellite Systems is a cross cutting technology that is used in many fields of modern life. Precise determination of position, velocity and time is particularly important in road applications and related technological solutions, which enable considerable improvements in traffic management and monitoring. In Vietnam, public transport vehicles and lorries must be equipped with black boxes to enable traffic management and control. This requirement and other safety needs, along with the increase of malicious tampering of black boxes to prevent their correct functioning, have motivated the development of some technological solutions and research lines. Some of the achieved results have been already applied in the real Vietnamese environment.

1 Introduction

In Vietnam, the road domain is a strategic sector in which solutions based on Global Navigation Satellite Systems (GNSS) are constantly expanding. Ministries with their Governmental Agencies and private actors are increasingly using GNSS-based systems to track and monitor vehicles and fleets with the main goal of: improving safety, increasing efficiency, reducing costs, and preventing law infringements and incorrect behaviors. However, there are several issues that need to be solved to properly reach the desired goals. Monitoring systems should be robust and able to mitigate possible disruptions of service. Their performance should be assessed carefully and comply with the technical specifications, and they should be as user friendly as possible reducing human errors and optimizing their efficiency. According to the

H. Van Hoang · T. D. Nguyen · T. H. Ta
Navis Centre, Hanoi University of Science and Technology, Hanoi, Vietnam

G. Povero (✉) · M. Troglia Gamba
Fondazione LINKS, Torino, Italy
e-mail: gabriella.povero@linksfoundation.com

report of the Ministry of Transportation of Vietnam, currently 1.2 million of commercial vehicles (trucks, busses, taxis) are equipped with monitoring black boxes, and according to the law, they must send Position Velocity and Time (PVT) information to the data center of the Directorate of Road Administration (Thuy 2016). This huge amount of data (about 2 terabytes per day) is used to monitor the behavior of drivers and transport companies identifying traffic law infringements (in particular for what concerns speed limits). This traffic monitoring system is a key element in the policy on traffic management and control implemented by the Ministry to increase safety and efficiency. Actually, after the system started its operations in April 2014, the number of casualties in traffic accidents has dropped significantly throughout the country. However, not all vehicles are properly sending their PVT data to the Data Center of the Directorate of Road Administration, thus impairing efficiency and performance of the whole monitoring system. This is related to both technical and non-technical reasons.

An important source of problems is related to the correct functioning of the equipment and of the communication channel between black boxes and the Data Center. According to the law, black boxes must be certified as compliant with their requested specifications by the Vietnam National Metrology Institute, but the testing procedure adopted so far consists in mounting the black box on a vehicle that is then driven on different roads at different speeds in order to verify the proper behavior of the equipment. This solution is costly and time consuming and allows officers to check only samples of the equipment, which are currently brought to the market by about 50 different providers.

Another important factor that can prevent PVT data to be sent to the Data Center is related to possible problems in GNSS signal tracking at the black box level. This problem can be caused by different factors. While in some areas, environmental conditions (limited sky view, dense tree canopy, etc.) can block the correct reception of GNSS signals, the most common reason for which black boxes cannot properly receive GNSS satellite signals is related to interference, either unintentional or intentional. Indeed, other communication signals can generate interference in the GNSS bands and this can be particularly severe in some locations. However intentional interference is probably the most important source of troubles. As a matter of fact, several drivers do not want to be tracked and want to avoid penalties when they infringe traffic regulation laws. For this reason, they install illegal jammers in their vehicles disrupting the functioning of their black box. Such kind of equipment, which can be easily bought through Internet or on the black market, disrupts the functioning of any receiver in a radius of several tenths or even hundreds of meters and is becoming a very serious problem that needs proper solutions.

Last but not least ionospheric scintillation, which is particularly strong at low latitudes, can impact considerably on radio signals. Vietnam is located under the Southeastern Asian Equatorial Ionospheric Anomaly and it is therefore particularly affected by this phenomenon which can disrupt radio signals traveling through the atmosphere (Spogli et al. 2016; Povero et al. 2017). For this reason, a proper monitoring and forecast of ionospheric conditions can contribute to improve the robustness of provided services.

In light of all these facts, considering the social interest of GNSS road applications and the importance that they have for public administrators, the Navis Centre, in collaboration with the NavSAS group of Fondazione LINKS and Politecnico di Torino in Italy, has devoted a considerable share of its Research and Technology Development (RTD) activities to studies and applications in this sector.

2 Technological Solutions for Traffic Monitoring Developed at the Navis Centre in Vietnam

On the technology development side, different tools have been developed. Some of them are already available on the market while others have just passed the proto-type stage and will reach the market in the next future. Considering the importance of tracking devices, a black box compliant with Vietnamese regulations was developed in 2015 and made available on the market. However, instead of focusing on ordinary tracking black boxes that are already commercialized by a large number of service providers, technology transfer efforts at the Navis have been oriented to satisfy the needs of some particular niche users. Indeed, for their fleet they require more sophisticated solutions in terms of safety, security, monitoring, and possibly need ancillary services integrated in the boxes (such as video and communication channels) or on the central monitoring server (such as route planning or very precise and robust localization thanks to ionospheric scintillation monitoring).

In Fig. 1, the general structure of this kind of system is represented. At user level, represented on the left-hand side, the vehicle is equipped with the black box which
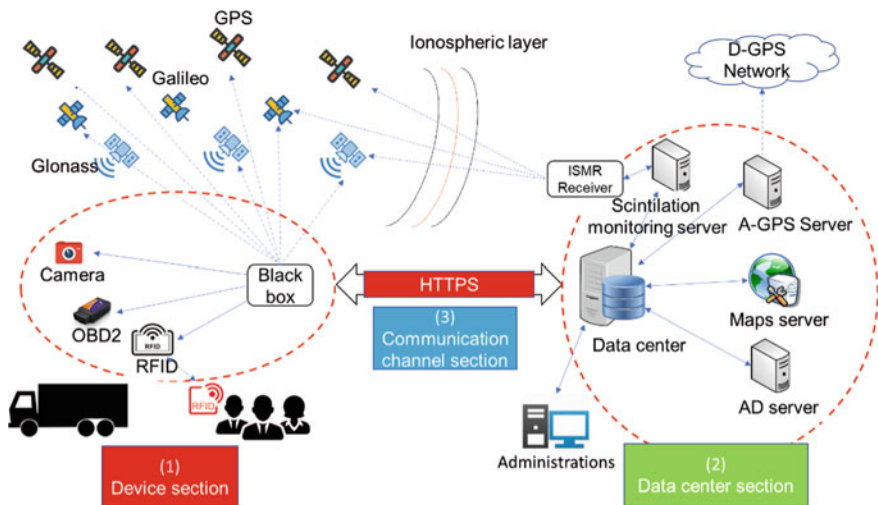
Fig. 1 General structure of the tracking system developed at Navis with extra features and services

embeds the GNSS receiver delivering the user's position. Additional sensors such as camera, RFID device, etc. can be installed onboard the vehicle depending on users' needs. On the right-hand side, the data center is shown. On top of the data storage section, additional services can support the user device by sending it information on ionospheric conditions, errors, routes, etc. so as to improve its performance. Users and data center are connected by a secure radio communication channel. Obviously, all the additional features, both at black box and at server level, are modular and can be selected to fit the needs of each specific user. Two important applications of this kind have been already commercialized: a monitoring system for cash delivery vehicles and a monitoring system for school busses and pupils.

In Vietnam the use of cash is still wide spread and collection and delivery of cash to bank branches as well as to ATM machines is an important part of the bank's logistics. An efficient control and management of cash delivery vehicles is therefore a quite important issue that requires reliable and robust solutions able to ensure a good level of reliability, security, and safety that ordinary black boxes cannot provide. The product designed at Navis, besides providing an interference-resilient real time monitoring of the PVT, supervises also other pieces of information such as: discrepancies from the scheduled route, opening and closing of doors, stops, personnel onboard etc. that need to be reliably monitored, promptly releasing alarms in case of need.

As for the monitoring of school busses and pupils, it should be noted that most primary and secondary private schools in Vietnam have their own bus service for pickup and delivery to transfer children between home and school. Parents are increasingly willing to know where their children are during the transfer and be sure that they reach safely their destination. A tracking system has therefore been designed and provided to a private school, which was willing to offer parents the possibility to monitor the transfers of their children.

Another equipment that already passed the prototype phase is the NAVISIM, a low-cost GNSS signal simulator that has been developed both for research purposes and for the testing of GNSS-based black boxes. To this extent, the NAVISIM has been designed with a specific interface that allows it to communicate with the black box collecting its output data (those that should be transmitted to the data center) while simulated GNSS signals are fed to its antenna through a cable. In this way, the closed-loop test setting depicted in Fig. 2 is obtained.

It allows users to verify whether output data of the black boxes are consistent with the GNSS signal fed to the antenna. In addition, with this simulator several black boxes can be tested at the same time in the lab, without mounting them on cars. Furthermore, tests can be performed using uniform and repeatable testing conditions. All these aspects turn out in a much higher efficiency of the testing process with respect to current practices. The simulator consists of three main processing blocks, namely: computing propagation delay, encoding navigation message, and generating digitalized signal. The first block computes the propagation delay between the visible satellites and the receiver, taking into account ionospheric and tropospheric delays. The second block encodes the navigation messages. The last block synthesizes the
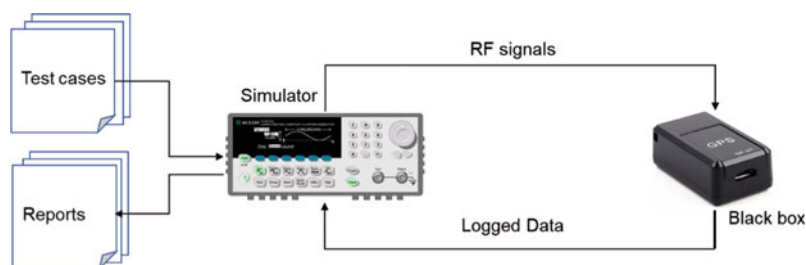
Fig. 2 Closed-loop setting for testing performances of tracking devices

information provided by the previous blocks and generates the line-of-sight and non-line-of-sight signals, including interferences, and noises.

3 Detecting Interference for Robust GNSS-Based Solutions

One of the hottest topics in GNSS research worldwide is how to protect GNSS-based services and make them more robust. In this field, a very important role is played by signal processing techniques able to detect and mitigate the effects of interference on GNSS receivers. In Vietnam, the problem of disruption of location-based services has become more and more relevant after the entry into force of the regulation for traffic monitoring, which imposes the installation of black boxes on public vehicles and lorries (Thuy 2016). In the last few years, the number of jammers has significantly grown in the country. For this reason, researchers of the Navis Centre in Vietnam and of Fondazione LINKS in Torino, Italy, have joined their efforts to study this issue and develop a low-cost portable prototype to detect interfering signals, which can impair the functionalities of a GNSS receiver.

3.1 *Jamming*

Interfering signals are one of the most known (and not solved) nuisance for GNSS receivers. According to (Pullen and Gao 2012), GNSS Radio Frequency Interference (RFI) can be divided into three categories:

- Malicious interference, defined as 'RFI intentionally transmitted to prevent the use of GNSS (or make its use hazardous) for as many users as possible'. Reference (Pullen and Gao 2012) specifies that, in case of such type of RFI, 'it makes sense to provide backup services to support transportation and other critical infrastructures';
- Uninformed interference, that is the 'intentional transmission of signals near GNSS frequencies, but without the desire to cause harm';

- Accidental interference, defined as the 'unintentional transmissions appearing near GNSS frequencies (typically from malfunctions of equipment design to transmit at non-GNSS bands)'. An example of this type of interference is that induced by Digital Video Broadcasting—Terrestrial (DVB-T) on the GPS L1 signals (Motella et al. 2008).

Although many variants of malicious interference exist, scientists generally group intentional interfering signals in *jamming*, *meaconing* and *spoofing*.

As for *jamming*, it is the act of intentionally directing powerful electromagnetic waves toward a victim receiver with the ultimate goal of denying its operations (Borio et al. 2016). Most jammers used in a civil context broadcast frequency-modulated periodic signals with a central frequency, which varies according to a periodic pattern. Jammers are a growing concern for GNSS applications, since they can be easily purchased on the web and used illegally. Depending on the properties of the jammer's frequency, the jammers can be divided into different classes, according to the characteristics of their transmitted signals. In most cases, the goal of malicious jammers is to totally deny GNSS-based services in a certain geographical area. Despite the clear threat posed by a jammer, strong interfering signals can be detected, and in turn the GNSS-based system can switch to backup positioning means or raise a warning for the users. On the contrary, intermediate power values turn out to be the most dangerous cases. Indeed, they might be severe enough to significantly decrease the receiver performance, but not strong enough to be detected and/or force the receiver to lose the lock of the tracked signals.

3.2 Development of a Low-Cost Portable Interference Detector

GNSS receivers with fully-capable RFI detection modules have been so far quite complex and, as a consequence, expensive and with high power consumption. Thus, these types of receivers are usually limited to specific professional or military applications. However, the growing number of GNSS-based applications and the need for high positioning accuracy and reliability are more and more demanding for consumer-grade receivers able to be aware of the level of RFI in their surroundings, so as to properly react to possible threats in real-time. In addition, application service providers can benefit from availability of real-time maps of RFI over wide areas to take prevention actions.

In literature, several works report on the creation of maps of interference sources in GNSS bands through various ad-hoc data collections to test specific detection, mitigation, or localization algorithms. However, those data collection campaigns can only offer a non-real-time sample of the average interference scenario in a certain environment, while new applications need a real-time picture of presence of RFI nearby a certain position.
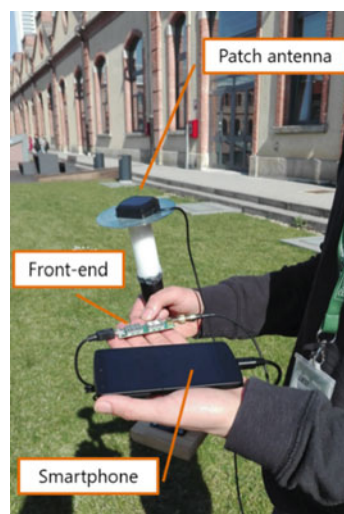
For this reason, researchers of the Navis Centre in Vietnam and of Fondazione LINKS in Italy have developed a prototype of a Portable Interference Detector (PID), which can be used to easily depict the interference scenario in mobility. Such a device can provide a useful tool to understand the RFI local environment during the deployment of the Regulation of the Department of Transport related to the tracking of all commercial vehicles.

The need of a simple, low cost and portable device has driven the design towards a prototype implemented on a smartphone. The whole system consists of three hardware components: a patch antenna, easy to be used in mobility; a front-end, to convert the analogue signal acquired by the antenna to a digital signal; a smartphone with suitable available computing power, as shown in Fig. 3.

The smartphone hosts the core part of the PID, which has been conceived as an App running under the Android Operative System. The high-level structure of the App is shown in Fig. 4. The App consists of a Graphical User Interface (GUI) and a receiver. The receiver is made up of several modules, namely: the data grabber, which collects the digitalized data produced by the Front End; the GNSS receiver, which processes the data and computes the position; the interference detector, which processes the data together with information received by the GNSS receiver and detects the presence of interfering signals. The communication module allows the connection of the device with a control server. As for the GNSS receiver, eNGene (Troglia Gamba et al. 2015), which is the GNSS software receiver developed by the NavSAS Group now running in real time on personal computers and embedded systems, has been adapted to run onto the microprocessor of a smartphone. A careful selection of the smartphone was needed to identify the model with the most suitable characteristics.

The GUI, shown in Fig. 5, allows users to interact with the receiver, to easily modify the configuration, to change the options and run it. The basic status of the
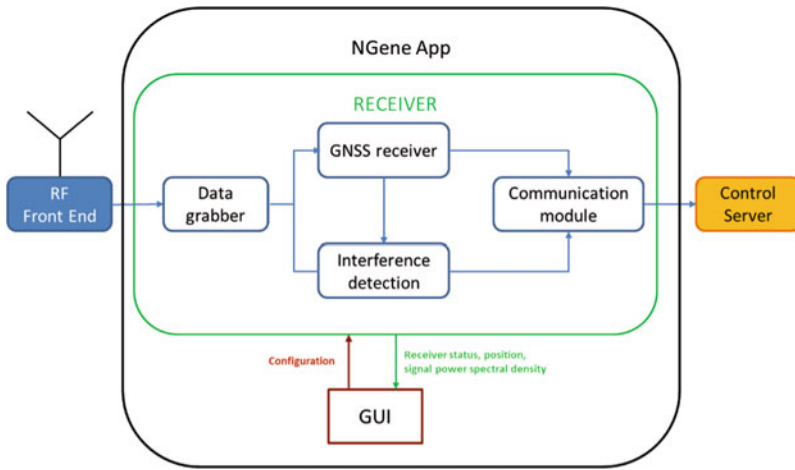
Fig. 4 High-level architecture of the prototype for interference detection
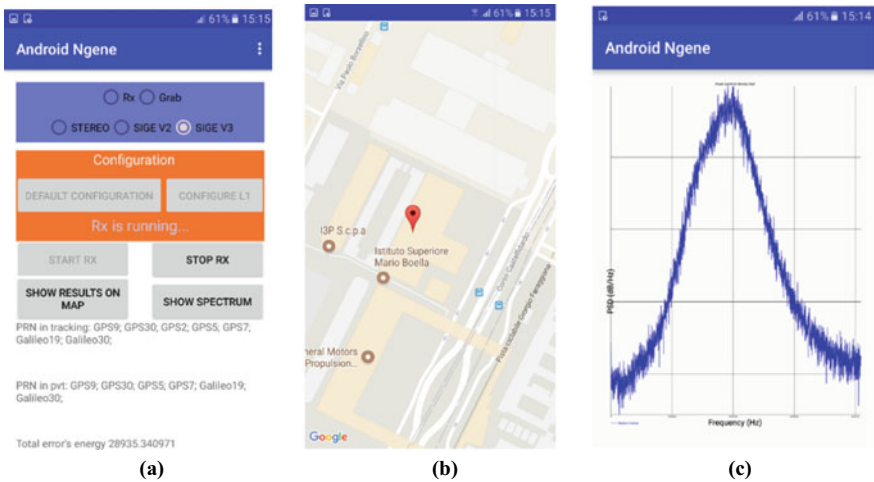
Fig. 5 GUI of NGene App: **a** Control panel; **b** device position shown on a map; **c** power spectral density of the received signal

receiver is displayed on the screen. The real-time position of the receiver is shown on the maps. In addition, the GUI plots the Power Spectral Density (PSD) of the received signals to allow users to monitor the presence of interfering signals in real-time.

The prototype has been tested in both in-lab tests and on-field measurement campaigns. The in-lab tests aimed to determine the sensitivity of the device to wide-band noise and continuous waves. As for the on-field campaigns, they have been conducted in an urban scenario with the goal to test the capability of the prototype
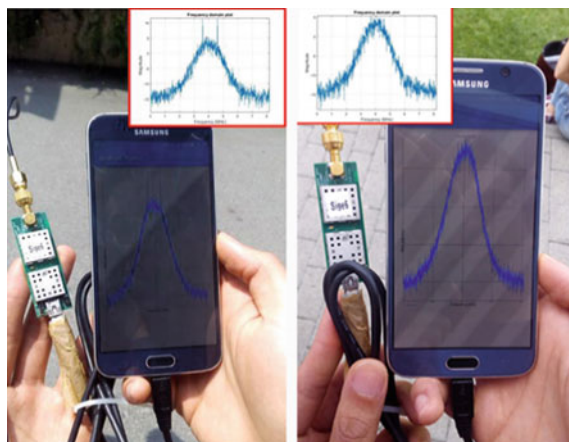
Fig. 6 Detection of interfering signals from space: interference detected in open sky conditions (left); no interference when the western portion of the sky was obscured by nearby buildings (right)

in detecting interfering signals in real-time and in actual conditions (Nguyen et al. 2018).

During the conducted measurements campaigns, several disturbances were detected, most of them likely generated by unintentional sources. However, the prototype was also successfully used to identify an anomalous signal broadcast by a non-operational GPS satellite (Dovis 2018), as shown in Fig. 6.

4 Conclusions

The RTD activities presented in this work have been conducted by researchers of the NAVIS Centre in Hanoi, Vietnam in collaboration with the Navigation Technologies Research Area of Fondazione LINKS (former ISMB) in Torino, Italy. Focus of presented activities is the use of GNSS-based solutions for applications in the road domain targeting peculiar characteristics of the Vietnamese environment. In particular, tracking devices and systems specifically tailored on needs of Vietnamese governmental agencies and customers have been developed and implemented. In addition, to make GNSS-based applications more and more robust and reliable, a prototype of a low-cost portable detector of interfering signals has been designed, produced and tested. This kind of device can be used to have a real-time picture of the interfering scenario in a determined area so as to ease the implementation of the Regulation on tracking of public vehicles in Vietnam.

Science and Technology in the framework of the project GILD (Progetto di Grande Rilevanza-Protocollo Esecutivo di Cooperazione Scientifica e Tecnologica bilaterale Italia-Vietnam 2017–2019, N_T.38.ITA/18).

References

Thuy DC (2016) Application of tracking devices for ensuring traffic safety of transport business activities. In: International workshop on research and applications of EGNSS on ITS Hanoi, Vietnam

Spogli L et al (2016) Formation of ionospheric irregularities over Southeast Asia during the 2015 St. Patrick's Day storm. J Geophys Res: Space Phys 121(12):211, 233. https://doi.org/10.1002/2016JA023222

Povero G et al (2017) Ionospheric Monitoring in South East Asia in the ERICA study. NAVIGATION 64(2):273–287. ISSN: 2161–4296. https://doi.org/10.1002/navi.194

Pullen S, Gao GX (2012) GNSS jamming in the name of privacy. Inside GNSS, 7(2)

Motella B, Pini M, Dovis F (2008) Investigation on the effect of strong out-of-band signals on global navigation satellite systems receivers. GPS Solutions 12(2)

Borio D, Dovis F, Kuusniemi H, Lo Presti L (2016) Impact and detection of GNSS jammers on consumer grade satellite navigation receivers. Proc IEEE 104(6).

Troglia Gamba M, Nicola M, Falletti E (2015) eNGene: an ARM based embedded real-time software GNSS receiver. In: Proceedings of the international conference institute of navigation (ION) GNSS+ 2015, Institute of Navigation, Tampa, FL, USA, 14–18 Sept 2015; pp 3178–3187.

Nguyen HL, Troglia Gamba M, Falletti E, Ta HT (2018) Situational awareness: mapping interference sources in real-time using a smartphone app. Sensors 18:4130. https://doi.org/10.3390/s18124130.

Dovis F et al (2018) Anomalous GPS signals reported from SVN49. GPS World. Available online: https://gpsworld.com/anomalous-gps-signals-reported-from-svn49/. Accessed 14 Sept 2018.

Blockchain for Smart Cities: Applications for IoT and Video Surveillance Systems

Pierluigi Gallo , **Uy Quoc Nguyen, Suporn Pongnumkul, and Giorgia Barone**

Abstract The recent revolution of the Internet of Things introduces new engaging operating scenarios. The IoT paradigm enables the intertwined use of physical and software components through the interconnection of devices that exchange data with each other without direct human interaction in several fields, especially in industrial and home environments. In the framework of the Italian-Vietnamese cooperation on the topics of smart cities and Blockchain, we present two applications of the blockchain technology, which can be applied, respectively in indoor, for monitoring and controlling smart homes, and in outdoor, for visual monitoring through video surveillance systems. In both cases, we propose decentralised architectures that aim at solving common IoT problems and vulnerabilities, with a specific focus on privacy issues. The introduced flexibility, extensibility and security in indoor IoT environments and outdoor video surveillance systems permit to foresee smarter cities that assist citizens with innovative services both in indoor and outdoor environments. We propose a reinterpretation of the role of the IoT gateway, making it blockchain-aware and handling IoT transactions without a central trusted authority. As for video surveillance, we focus on two different directions. First, we provide the capability to analyse multiple video-flows from different cameras deployed around the city by heterogeneous owners, guaranteeing the integrity of the timestamps and camera settings. Second, the system ensures that the positions of cameras, their orientation and their mechanical parameters are not physically manipulated; otherwise, the platform generates a warning. We guarantee citizens' privacy monitoring and

P. Gallo (✉) · G. Barone
Dipartimento di Ingegneria, Università di Palermo, Palermo, Italy
e-mail: pierluigi.gallo@unipa.it

G. Barone
e-mail: giorgiamaria.barone@gmail.com

U. Q. Nguyen
Posts and Telecommunications Institute of Technology, Hanoi, Vietnam
e-mail: uynq@ptit.edu.vn

S. Pongnumkul
NECTEC, Khlong Luang, Thailand
e-mail: Suporn.Pongnumkul@nectec.or.th

tracking camera settings and storing video flows. The absence of digital and physical manipulations guarantees citizens against possible violations of their privacy. To face these risks, we introduce a blockchain-based video surveillance system that jointly provides validation and immutability to camera settings and surveillance videos, making them readily available to authorised users in case of events and paving the way to new distributed city-wide monitoring systems.

1 Introduction

'Smart City' is a development goal of many cities around the world, in response to the increasing complexity of urban areas and the potential of connected and intelligent technologies, especially Internet of Things (IoT), in improving urban experiences. We focus on two relevant aspects of smart cities: the capability of using IoT for a wide range of data types (temperature, humidity) and the video-surveillance, city-wide sensing augmented by computer vision. As crime and accident rates usually are high in highly populated settings, the safety and security of the people are among the main goals of most smart cities. The technology that has been widely used to address this issue is video surveillance, in which streaming cameras, i.e., closed-circuit television (CCTV), are installed to monitor and record possible crimes and accidents. Video surveillance is known to help solve crimes and, in some cases, also prevent them; therefore, it is common to see multiple CCTVs installed in malls, homes, and public places. The many CCTVs could provide a comprehensive view of each city; however, the fragmented ownership of the cameras impedes the integration of the information coming from different video flows. The nature of CCTV, which is owned by multiple and untrusted parties, has originated a few security and privacy challenges. First, it is prone to be maliciously attacked or manipulated in various ways (Costin 2016). Knowledgeable hackers can manipulate the cameras and the video repositories, making them useless in the events of crimes. Second, CCTVs pointed to inappropriate directions could violate the privacy of others, especially neighbours. We argue that the integrity of the video flows and the privacy of citizens have not been addressed at the city-wide level, considering multiple CCTVs as a distributed IoT system without centralized control. The last few years have witnessed a renewed interest towards distributed systems, thanks to the diffusion of the blockchain, a distributed ledger technology (DLT) that lays behind the popular cryptocurrency named Bitcoin. Blockchains enable peer-to-peer interactions where non-trusting participants can interact with each other without a trusted intermediary in a verifiable manner. The blockchain records immutable information due to the use of cryptographic hash and the distribution of the content among the network participants. While blockchain started in financial technology domain, it is believed to be a disruptive technology for many industries, including solving challenges in IoT (Christidis and Devetsikiotis 2016). For the IoT surveillance issue, the authors proposed BlockSee (Gallo et al. 2018), an immutable ledger of validated views of the city without a trusted intermediate. BlockSee uses metadata of surveillance systems

to help to build an authoritative and comprehensive view of the city in case of events and a verified occurrence of potential privacy violations due to changes in the settings of cameras.

Another critical element of the smart cities is automation in home and industrial environments. The Fourth Industrial Revolution, known as Industry 4.0, is pushing towards decentralization and self-regulation by leveraging several paradigms and technologies including cyber-physical systems (CPSs) and the Internet of Things (IoT) (Hofmann and Rüsch 2017). IoT is the network of things: physical devices, sensors, actuators, appliances connect each other and exchange data. The IoT paradigm allows improving the quality of life for households and workers, making it easier to automatically carry out actions that are usually carried out by the man. However, having so many interconnected devices introduces privacy and security issues related to unauthorized access, data tampering, denial of service. These issues require defining new mechanisms for access control that jointly maintain data integrity and supersede the limits of current centralized solutions, thus granting and revoking the rights to access data without single points of failure. In facts, distributed systems are more resilient to attacks and failures.

Furthermore, a fully-centralized approach works only in local environments; it is doomed to fail in a global IoT scenario because of the plethora of various solutions and verticals, whose convergence is difficult to happen. To shift the paradigm of Industrial and Domestic IoT from centralized to distributed, the authors proposed DeCyMo (Gallo et al. 2018), a Decentralized Cyber-Physical System for Monitoring and Controlling Industries and Homes. For managing information, DeCyMo relies on the blockchain, in particular on the Multichain platforms. This choice was motivated by the fact that Multichain supports streams, which provide a natural abstraction for data retrieval, timestamping and archiving. The Multichain streams are therefore used for managing message queues, substituting IoT brokering functionalities, and disseminating IoT topics. Furthermore, we integrate into our stream-based architecture also a method for adjusting privacy settings through fine-grained access control.

2 Related Work

The joint use of IoT and blockchain was proposed in (Pinno et al. 2017) with the introduction of ControlChain, a customized blockchain for using relations between devices. The fundamental idea is to use a decoder that translates data from the typical IoT models towards ControlChain while maintaining the access rules. The Blockchain Connected Gateway has been proposed in (Cha et al. 2017) as a mediator between users and IoT devices with the primary purpose of handling privacy policies and protect users from providing private data without their consent. Users obtain information about IoT devices and their privacy policies through the blockchain connected gateway. As preferences are immutably stored on the blockchain, they are used to resolve future disputes between users and IoT service providers. Limited

resources of the things characterize IoT scenarios. Therefore, scalability issues are dramatic and require the use of lightweight blockchains. In the literature, solutions for these lightweight blockchains employed a hierarchical structure with three levels, including smart home, the overlay network, and cloud storage (Dorri et al. 2017; Dorri et al. 2017; Bogdanov et al. 2011). The communication between these layers is enabled by blockchain transactions, which are used for storage, access control and monitoring. Furthermore, local storage is provided in a centralized ledger that is similar to a blockchain. Finally, in Rifi et al. (2017), it was proposed a Publisher-Subscriber mechanism where the IoT gateway dispatches messages between the smart home and the blockchain also using intermediate cloud servers. The authors propose to use smart contracts to save rules, authentication, and communication between nodes. The gateway uses IPFS (IPFS—InterPlanetary File System) for the off-chain storage of data arrived from the sensors, then generates the hash of the location in the database and add it to the blockchain. This kind of storage makes the transaction visible to all nodes. Our architecture differs in the role assumed by the IoT gateway and in the methodology, we use to make it more flexible and programmable to address the needs of a blockchain node. Following previous work, DeCyMo provides a publish-subscribe mechanism, but unlike in the past, it introduces an intermediate re-publishing on blockchain streams (a novel blockchain tool) and does not require smart contracts.

Recording the integrity of videos on the blockchain has been studied in the context of individual videos. Gipp et al. (Gipp et al. 2016) proposed to record the integrity of videos from a car's dashboard in the event of a collision, for using the video with provable integrity in court. In their android-phone-based system, when a collision is detected automatically using built-in accelerometers, the current video is cryptographically hashed, and the hash is recorded on bitcoin blockchain via the Origin-Stamp protocol. Similarly, the authors in Hemlin Billström and Huss (2017) also explored an android-phone-based system where hashes of videos are recorded on a blockchain to preserve their integrity. We extend the previous work on the integrity of individual videos by providing a whole architecture for recording the integrity of multiple videos, city-wide. BlockSee records integrity of multiple videos in the same blockchain network. Apart from being able to prove the integrity of each video flow, also their time ordering is preserved even if the clock time of individual videos is not set properly.

Privacy of data has been a concern for both blockchain (Kosba et al. 2016) and video surveillance (Upmanyu et al. 2009). In the blockchain domain, bitcoin protects the identity of users by using pseudonyms (the hash of their public keys) as account numbers for recording the amount of money transferred on the blockchain. As financial transactions are considered highly secret, Kosba et al. (2016) addressed the problem of transactional privacy in their proposed system Hawk. The concept of Hawk is to send encrypted information to the blockchain and relies on zero-knowledge proofs to enforce the correctness of the data and execution of the smart contract. Similarly, in the video surveillance domain, a video contains meaningful information and is beneficial to share among organizations. However, sharing raw videos can reach to sharing information that is not intended. Upmanyu et al. (2009)

proposed an approach to preserve surveillance privacy. Each image frame is split into random small images which by themselves contained no information. Then a secret sharing algorithm is applied to distribute and store the small images among the participants. In a similar spirit to earlier privacy-preserving research, this chapter also does not record any private data from videos but allows potentially interested users to know whether camera settings have been changed.

3 Background

3.1 Internet of Things

The IoT uses a centralized architecture composed of a few main elements: the IoT gateway, the cloud service as well as devices and nodes. Devices are generally low-resources sensors and actuators, without communication and storage capabilities. Therefore, they are helped by sensor nodes, which send data to the IoT gateway, which is the hub for data collected by sensors and processes. The gateway applies to filter, makes analysis, removes inappropriate values by analytical algorithms, then send data to the cloud server, which provides storage and algorithms for big data analysis. The cloud permits the interaction with the users, whereas the communication among the components described above is demanded to dedicated transport protocols such as HTTP, MQTT, and CoAP. The Hypertext Transfer Protocol (HTTP) and the Constrained Application Protocol (CoAP) are application-layer protocols for transferring data on the web. HTTP is an application-layer protocol initially designed for communication between browsers and web servers; however, its use has spread to many purposes; CoAP was explicitly designed for the needs of constrained devices and ran over UDP. Message Queuing Telemetry Transport (MQTT) is a publish-subscribe messaging protocol designed for lightweight M2M communications; it was initially developed by IBM and is now an open standard. MQTT has a client/server model over TCP, where every sensor is a client that connects and send data to the server (Stanford-Clark and Truongk 2013). The server provides brokering functionalities, and widespread implementation of an MQTT broker is Mosquitto (Light 2017).

3.2 Blockchain and Smart Contracts

The blockchain technology guarantees the immutability of validated records. Networked computers collaboratively maintain a growing list of grouped records (i.e., blocks) which are linked (i.e., chained) to the past blocks using cryptography. The blockchain is first known as the backbone technology of Bitcoin (Nakamoto 2008), the first popular cryptocurrency, which allows Internet users to transfer money

(in BTC) to each other in a peer-to-peer manner without the needs of a trusted intermediary. Blockchain, as implemented in Bitcoin, integrates cryptography and distribution of records to make it difficult to tamper with the transactions, which has proven its effectiveness as Bitcoin has been in operation since 2009 and no significant attack has been successful. While Bitcoin is a specific application of blockchain as a public financial application that supports the transfer of value from one participant to another, blockchain has been extended to use in other contexts. In the crossway between smart grids and smart cities, the blockchain enables transactive energy solutions that also include energy losses (Sanseverino et al. 2017) and permit the technical operation of the distribution network (Silvestre et al. 2018).

Applicability of blockchain to other use cases beyond Bitcoin has been extended with the concept of smart contracts on the blockchain. A critical mechanism in Bitcoin is the automatic verification that an account has sufficient funds before user A can transfer money to user B. To extend blockchain to other applications, the validation process has been included in smart contracts, which are custom user programs that can be recorded on the blockchain to automatically verify data and execute actions when predetermined conditions are met. An example of a smart contract in real life is the logic in vending machines. When a vending machine receives an order of a snack, it waits for the user to put money into the device. If the user does not put sufficient amount of money to cover the cost of the specified snack, it releases the money that was placed in the machine, if any. If a sufficient amount of money arrives, it releases the specified snack and the amount of money that exceeds the cost. Similar to data stored on the blockchain, smart contracts need to be agreed upon by the participants of the network, cannot be tampered with, and can be executed on all the participating nodes.

Computers that participate in a blockchain need to agree upon the content of the ledger; this is done through a consensus mechanism. Instead of trusting an authority to decide which content is the truth in a centralized manner, blockchain shares the content amongst everyone. This approach requires that the network maintains consensus around the information recorded on the blockchain. Many consensus mechanisms have been proposed and deployed. Each consensus mechanism has different trade-offs between the security of the data and the economics or performances of the system. Bitcoin utilizes Proof-of-Work (PoW), which is highly secure but performs slowly and wastes electricity. Examples of other popular consensus mechanisms include Proof-of-Stake (PoS), Practical Byzantine Fault Tolerance (PBFT) (Castro and Liskov 2002).

Bitcoin is a public blockchain, which means that anyone who wishes to use their computers to help maintain the records or to use Bitcoin as a cryptocurrency can participate. While the public approach allows for mass participation and increases trustworthiness, there are many use cases where the data cannot be a public record, even if the data is encrypted. Examples include data which impact government security or business decisions. Therefore, permissioned blockchains have been invented. In permissioned blockchains, the computers that participate in record keeping need to be granted permissions. This type of blockchain is more suitable for business and government uses, and this chapter uses permissioned blockchain. Since 2015,

there have been many blockchain platforms developed. Ethereum and Hyperledger Fabric are two examples of popular blockchain on the market with smart contract capability. Any system based on the blockchain relies on the following components: the *Blockchain actors* and their participation in the blockchain network. In public blockchain (e.g., bitcoin, Ethereum), anyone can use their computers to be a blockchain node. In permissioned blockchain (e.g., Hyperledger Fabric, permissioned deployment of Ethereum), only computers which are granted permissions can join the network. Multiple types of nodes coexist in the same blockchain network; for example, in Bitcoin, the miners are particular nodes that help to verify transactions and to add blocks to the growing chain, other nodes receive and maintain the records. *Transactions* indicate the transfer of values (money, energy, etc.) between actors, generally the nodes of the blockchain. Blockchain transactions indicate state transitions for the system under study, switching from an old state to a new one. In this chapter, we generalize the concept of transaction, including also the exchange of 'information' between nodes. *Smart contracts* are the custom logic that is recorded and executed on the blockchain. In Ethereum, smart contracts are implemented in Solidity language. In Hyperledger Fabric, smart contracts, which are called chaincodes, can be implemented in Go language or Java. *Validation* for transactions to be included in the blockchain may be run by smart contracts or by specific nodes (e.g., the miners in BitCoin). This is a key concept complementary to the immutability of the blockchain: immutable data in the digital world without a preliminary validation from the physical world would be of any use. In facts, perennial transactions have to represent correct statements in the physical domain.

3.3 DeCyMo Architecture for Blockchain and IoT

DeCyMo uses IoT and blockchain technologies for implementing cyber-physical systems for home and industries. Although IoT works with a centralized approach towards the gateway and the blockchain is completely distributed, DeCyMo decentralized architecture combines the two paradigms and provides flexibility and programmability by revising the role of the IoT gateway, as shown in Fig. 1.

The main difference between the archetypal IoT topology and DeCyMo is the role of the IoT gateway, which is the both an IoT broker and a blockchain node; it is the contact point between these two technologies, straddling the boundary between the centralized and distributed paradigms. DeCyMo topology is decentralized, the joint point between the centralized topology of IoT and the distributed approach of the blockchain. The IoT gateway requires, therefore, to be jointly compliant with several possible IoT protocols and with several blockchain mechanisms. Supporting all these aspects is unfeasible because of the lack of standardization and the full range of adopted solutions. Therefore, flexibility is required for the gateway to adapt to the scenario(s) in use, which depends on the IoT devices, the communication protocol, the kind of blockchain. The use of Docker containers permits to compartmentalize applications while guaranteeing interactions among them and enable the
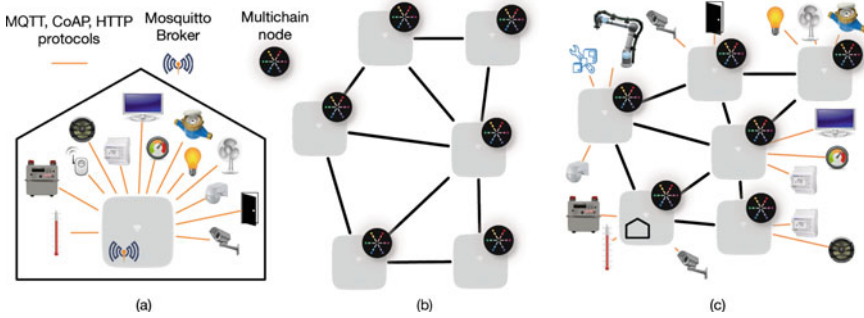
Fig. 1 Centralized topology on the IoT gateway at home or industry (**a**) typical distributed blockchain network topology (**b**), DeCyMo decentralized topology with IoT gateways as blockchain nodes (**c**)

DeCyMo decentralized infrastructure. Without DeCyMo, IoT data are made available to selected users through the Internet using publish-subscribe mechanisms or shared repositories in the cloud. Unlike the past, DeCyMo data flow consists of a subscription of sensors to IoT brokers, which in turn publish to Multichain streams. Users subscribe to the Multichain streams of their interest which they are authorized to.

DeCyMo uses lightweight clients; they do not contain the full history of transactions and use the publish-subscribe paradigm to the blockchain. DeCyMo works using a publish/re-publish/subscribe procedure where data are published on the broker by sensors, then the broker re-publish such data on dedicated Multichain streams (Coin Sciences Ltd. 2016) after encrypting them with symmetric key encryption. DeCyMo does not require any change for sensors and IoT nodes: they continue to publish their measurements on their broker. Finally, users subscribe to the data streams of interest.

The access to the stream is regulated by function *Auth* (*S*, *D*) that provides true if a subscriber *S* is authorized to read data stream *D*, this is depicted in the figure by green circles. The subscription function *Sub* (*P*, *S*) provides true or false depending on if the subscriber *S* subscribed to the publisher *P*, represented by black circles. We indicate with *S* the boolean subscription matrix, whose elements if the *j-th* user subscribed to the *i-th* publisher, 0 otherwise. The same notation applies to authorization matrix *A*, indicating if the *j-th* user is authorized to access the content published by the *i-th* publisher. Then, the actual access to streams is regulated by authorized subscriptions $W = A \wedge S$, where $\wedge$ is the logic AND. Confidentiality is required in the publish-subscribe paradigm to restrict the data access, only to the subset of blockchain nodes with authorized subscriptions. DeCyMo provides confidentiality using a well-known procedure that combines symmetric and asymmetric cryptography with the addition of two extra streams: one for disclosing public keys and one for giving the secret key only to authorized users. Data is encrypted with a symmetric key mechanism (AES), then is stored and timestamped on the blockchain on the streams corresponding to IoT topics (the black horizontal lines in Fig. 2. The AES secret key is encrypted with
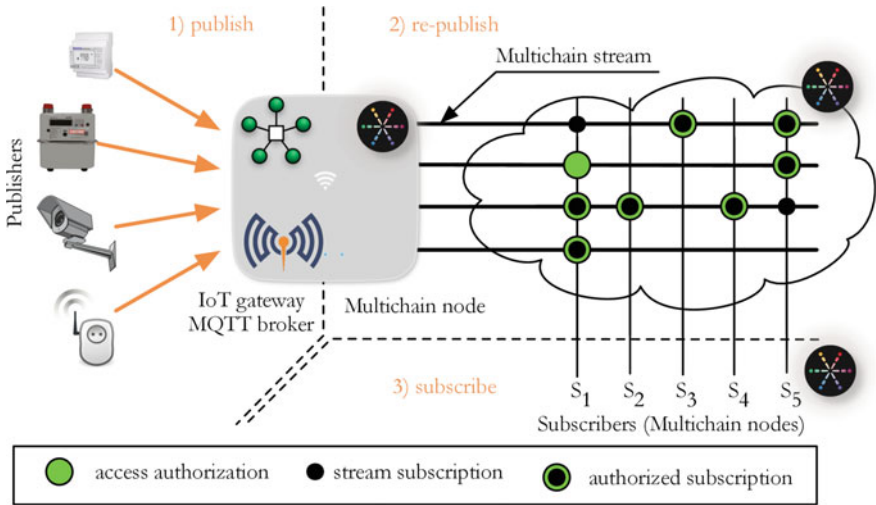
Fig. 2 The publish/re-publish/subscribe mechanism of DeCyMo; publish on the broker, re-publish on Multichain streams, subscribe to Multichain streams

the public key of the authorized nodes, then it is published on a dedicated stream. The public keys of nodes are publicly available on another dedicated stream. The use of blockchain streams for replacing legacy publish/subscribe tools is in agreement with the idea that the location of data is not essential; IoT data has the central role, and the IoT topic identifies it, wherever it is stored off-chain. The use of streams for replacing standard publish/subscribe tools has some drawbacks, as blockchains are much slower than databases. Therefore, performance is a crucial element to consider, as we will discuss in the next section.

3.4 Video Surveillance in Smart Cities

Video surveillance has recently obtained attention for its applications in private and public environments, for monitoring and protecting areas at different scales: houses, malls, or even metropolitan areas. A unified vision of different video sources is difficult as videos are produced and managed by heterogeneous actors, but a global view would be beneficial for applications such as smart transport, traffic monitoring, video surveillance, and forensic analysis. The benefits of jointly using multiple video flows are dramatic when heterogeneous sources are correlated in time and space. However, this introduces the problem of the availability of data and the preservation of citizens' privacy.

3.4.1 Camera Settings

BlockSee extends video-surveillance systems that focus on the integrity of video flows (Gipp et al. 2016; Hemlin Billström and Huss 2017) by validating and guaranteeing the integrity of camera settings including the position of the camera, the direction of view and the zoom level.

It is quite challenging to keep track of the orientation and the configuration of thousands of cameras spread over the city, as they can change over time. Besides well instructed and authorized technicians, camera owners as well as external factors and actors—including hackers—can manipulate digital and physical settings. The complexity of such settings explodes considering that the available tunable knobs depend on the camera model and vendor; the actual configuration depends on the untrusted owners, hackers, events. BlockSee solves the difficulty to track camera configuration, through a lightweight photogrammetry analysis of the camera frames as it arrives at the end-user, i.e., downstream of possible manipulations.

To explain our reference scenario, we refer to one frame from the stream that is publicly available at (City of Helmstedt—Webcam of the City Hall), that is shown in Fig. 3. This exemplary image explains BlockSee working principles through a fictional scenario that explains potential privacy violations and is representative of several disputes between neighbours. Marks A, B, C indicate the entrance doors to
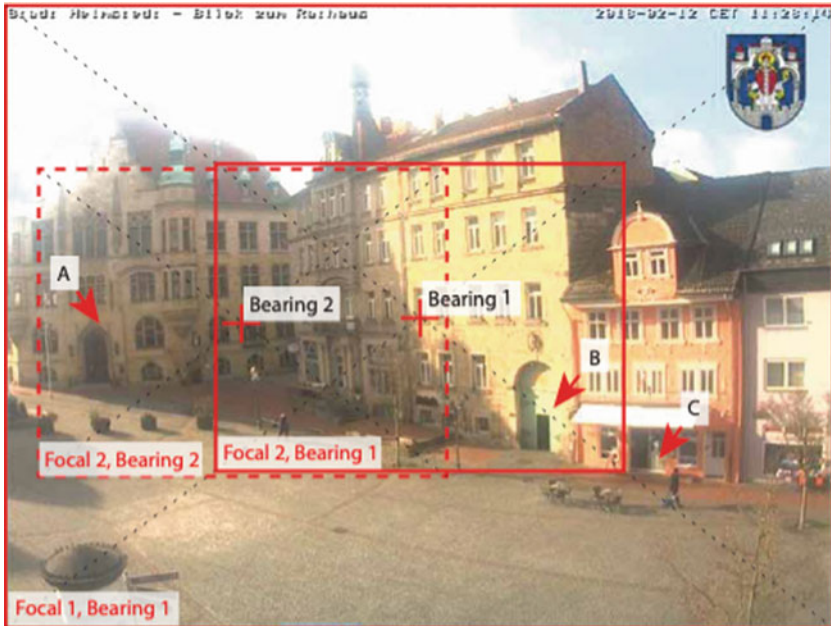
Fig. 3 Effects of settings modification on the field of view and the provided image: changes of the focal length are highlighted by solid red rectangles, changes of bearing angle are shown by the red crosses

three different areas of the building, pointed by as many red arrows. Let's assume that the camera is own by the household B (the one that uses door B) for monitoring the access to his venue. If B sets the camera for using Focal2 and Bearing1, only door B is watched, in the other cases, the camera monitors also doors A and C, offending the privacy of people that go through such doors. The *camera bearing* influences the centres of the FoVs (indicated by red crosses), the *camera focal length* influences the zoom (indicated by red rectangles).

In our example, door B is the intended target. Therefore, camera settings with Focal2 and Bearing1 are the only ones that do not introduce privacy violations. Even cameras that are—provably and nicely—deployed and configured (e.g., by a certified technician or under the supervision of a Judge's delegate), may be changed in their settings afterwards, without any evidence to the interested actors. A recent verdict of the Court of Cassation, Italy's supreme court, states that to modify the field of view of the camera or change its optical properties are simple operations that can be done out of control from the appellants and can lead to potential privacy issues (Cassazione). People that are potentially offended cannot demonstrate that cameras have been involuntarily or maliciously changed from their certified settings. Conversely, camera owners cannot demonstrate that their cameras are positioned well—all the times—without revealing the whole video stream.

3.4.2 Position

Changing the position of the camera permits to change the watched scene and therefore offend neighbours' privacy. The *direction of view* is given by bearing and tilt angles, which can be changed through remote commands in some models of cameras. Even fixed-mounted cameras can change their direction of view, due to intentional or accidental mechanical events (camera re-deployment, wind, vandalism, network attacks, etc.). As shown in Fig. 3, the *zoom level* has a role in defining what is monitored by a camera. Multi-focal cameras can change their level of zoom through a mechanic or digital adjustments. The *focal length* can be changed through physical knobs on the camera or piloting motorized focal lengths through remote commands. Furthermore, the *physical position* and the direction of view of the camera defines the watched scene and therefore offend neighbours' privacy. In particular, the bearing and tilt angles can be changed through remote commands in some models of cameras or can be changed even in fixed-mounted cameras due to intentional or accidental mechanical events (camera re-deployment, wind, vandalism, network attacks, etc.). Finally, as shown in Fig. 3, the *zoom level* influences the scene within the field of view. Multi-focal cameras can change their level of zoom through a mechanic or digital adjustments. The focal length can be changed through physical knobs on the camera or piloting motorized focal lengths through remote commands.

3.4.3 Image Analysis

The BlockSee key idea is to infer camera settings through a lightweight photogrammetry analysis on the video frames, then store them with the immutability and the fine-grained access of the blockchain.

First, BlockSee provides video frame segmentation, then it distinguishes the background and foreground (Brutzer et al. 2011) and creating a background model of the sequence. The immutable background is generally subtracted from the image in order to focus on new elements that enter or move in the scene (Bradski and Kaehler 2008), and segmentation is referred to in the literature as background subtraction. Unlike typical applications, BlockSee specifically uses the background for deducing camera settings. The choice of the specific segmentation technique is not relevant for our purposes. We consider *frame differencing* for its simplicity: the background is obtained as the part of the scene that does not change over time, computing the difference between frames. In the rest of the chapter, we will neglect the image foreground, which can be used to identify events occurring in the scene and trigger alarms.

3.4.4 Feature Extraction and Settings Inference

Relevant features are extracted from the background (e.g., corners, edges, etc.), then the distances of such features are computed from the borders of the frames. Features are extracted from the background using algorithms that are invariant to scale or rotations, such as the scale-invariant feature transform algorithm (SIFT) (Lowe 2004) and the Binary Robust Invariant Scalable Keypoints (BRISK) (Leutenegger et al. 2011). These features are the fingerprint of the image background. The quality of the matching between the last certified deployment and the current one says if the camera settings have changed since a certified deployment, revealing eventual uncertified reconfigurations. The algorithms mentioned above work well for small affine transformations between the certified background and the current background. This is not a problem as massive alterations of camera settings are easy to detect. BlockSee does not require any distortion model of the camera as the algorithm applies a differential analysis. This image analysis can be executed both by the camera or by smart contracts, however implementing computer vision on a smart contract is not an easy and safe task, especially under the number of actors that are explained in the following section.

3.5 *BlockSee Architecture*

BlockSee architecture is reported in Fig. 4, it includes three different levels: cameras, nodes, and the blockchain core. BlockSee is designed on the top of a permissioned blockchain: only authorized nodes can write on the blockchain by appending the

Fig. 4 BlockSee
architecture and its
heterogeneous nodes in
terms of computation,
storage capabilities, access
permissions
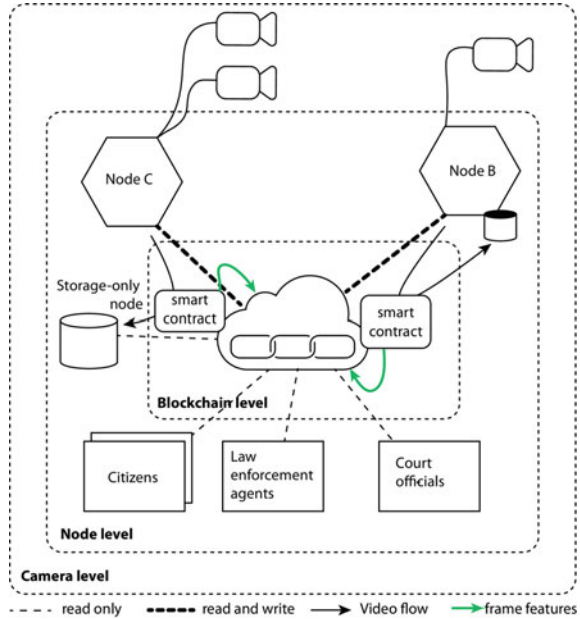
hashes of video segments as well as their metadata (camera settings, for fast data retrieval). The use of a permissioned blockchain prevents Sybil attacks (Douceur 2002); therefore, cheap consensus mechanisms can be used.

3.6   BlockSee Actors

BlockSee works with different actors under heterogeneous trust models, roles, and read/write permissions on blocks.

- *Camera manufacturers* are trusted if they pre-load BlockSee-compatible cameras with digital certificates for periodically signing the output frames. Legacy cameras that do not have digital certificates and do not sign (periodic) frames cannot provide content to BlockSee.
- *Cameras* provide the video flows. BlockSee-compatible cameras are assumed to periodically sign specific frames (e.g., a subset of the keyframes). No extra prior knowledge of camera specifications, settings, and capabilities is assumed (e.g., manual/motorized, fixed-focal/multi-focal, etc.). BlockSee does not require physical access to cameras (besides the deployment), nor a dedicated control channel (e.g., for receiving camera configuration parameters).
- The *Certified installers* of cameras hold a certificate to sign frames taken at the end of installation or maintenance. Even certified installers are not trusted; they are accountable for the installation and settings.

- *Nodes* have a threefold role in BlockSee: (i) they collect video flows from different cameras, (ii) validate video transactions, (iii) write hashes of videos and camera settings on the blockchain. Computation power, complexity, and storage capabilities of nodes are heterogeneous. Nodes are not trusted; they can be both legacy computers or NVRs (Network Video Recorders), which usually manage several cameras of an administrative domain (e.g., a shop). Nodes encrypt video flows and store them off-chain, in local or remote repositories, that can be centralized or distributed as shown in Fig. 4. Several examples of distributed repositories have recently appeared, including IPFS (IPFS—InterPlanetary File System), Swarm (Trón et al.), and StorJ (StorJ).
- *Citizens* have the interest to look up BlockSee, to check if their privacy has been violated. Citizens are untrusted and have access to camera settings, as they can read only frame features that permit them to check if their privacy has been violated. Citizens cannot watch the actual video flows, which are encrypted.
- *Policemen and court officials* are considered as 'moderately trusted'; they can watch the videos only in case of events after a Judge provides the authorization. Singularly, they cannot watch the content of the videos; they are enabled only if two over the three authorized actors agree (e.g., police and court officials, police and camera owners, court and camera owners). This is possible by using multi-signed keys for the encrypted video segments; the information can be decrypted when two out of three holders of private keys require viewing the flow.

3.7 BlockSee Transactions and Validation

In the blockchain terminology, transactions are movements of value between two parties (money, energy, etc.). BlockSee transactions are movements of video segments or just video frames from cameras to nodes; this is in line with the recent trend to consider data as a valuable good. Unlike financial transactions, which do not occur at regular intervals as they depend on human actions, BlockSee video-transactions occur periodically. BlockSee transactions contain the timestamp, the camera id, the hash of video segments H(S) for video integrity, and the features F of the background model, for the integrity of camera settings. Two extra bits permit to distinguish normal frames from accountable and certified ones, as discussed later on. When technicians install or maintain the cameras, they may apply changes in their settings. Installers and technicians are not trusted and may deliberately set cameras with offending configurations and exculpate themselves afterwards by attributing the malicious settings as occurred on a later time. Video-transactions that are signed by installers are accountable for camera settings as the installer is responsible for them. Transactions that are jointly signed by installers and court officials are certified. Even simple accountability is a valuable goal if associated with proof of immutability, as provided by BlockSee. Any unaccountable or uncertified modification of camera settings is permanently recorded on the blockchain and can be timely faced. Video transactions are packaged in blocks, then they are chained in BlockSee, as shown
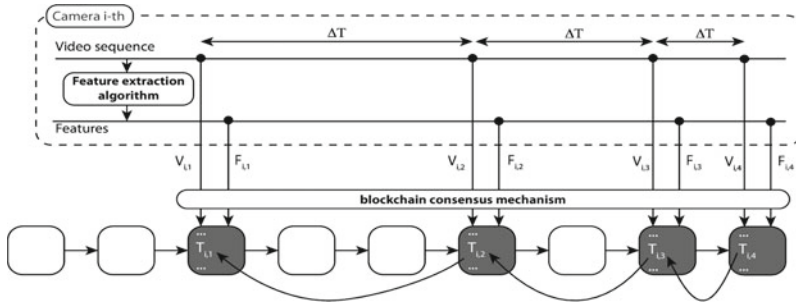
Fig. 5 Sequence of blocks in BlockSee and data coming from the i-th camera

in Fig. 5. The *i*-th camera is the source of views *V* and features *F*, which are the output of the feature extraction algorithm. Besides the blockchain chaining mechanism based on hash functions, indicated by left-to-right arrows, video transactions are chronologically linked each other and permit to track the status of a specific camera (as indicated by right-to-left arrows). Transactions come from multiple cameras, through a multitude of nodes, they are identified by a double index, the id of the camera *i* and the sequence number *j* for that camera. In the figure, the blocks that contain video transactions from the *i*-th camera are indicated in grey and give the full history of that camera.

3.7.1 Expression Missing

A simplified version of the video transaction is indicated in the expression above. Starting from a view, for example, the *j*th frame from camera *i*, the gneric transaction includes the features *F* extracted from the view and the hash of the view. All views are signed by the producing cameras; views are optionally signed by the technician that operated a change in configuration, due to first set up or maintenance. The information on accountability and certification, as well as all updates of camera settings, is available to all citizens that can make analysis over time. The validity of blocks is checked by a dedicated BlockSee smart contract; video segments are valid when signed by a camera vendor. Unsigned video flows are not processed by BlockSee as they are considered invalid; frames coming from legacy cameras cannot enter BlockSee, this provides backward compatibility. We note that a common attack on CCTV cameras is to send the old static image instead of the current view. The feature extraction part would not be able to detect this malicious attack, but the hash of the video frame would. Two images of the same scene over time should see minor changes and would not result in the same hash.

4 Experimental Results

4.1 DeCyMo Results for Blockchain and IoT

Our experimental setup permits to validate the DeCyMo architecture using a simple but effective deployment that mimics the typical home and industrial IoT infrastructure. The setup includes temperature and humidity sensors, smart plugs as actuators, one IoT node and one IoT gateway, as described in Fig. 5. The gateway is implemented through Docker containers to leverage virtualization, to make smooth data processing, and quickly adapt to the blockchain in use. DeCyMo IoT gateway is deployed using a Raspberry Pi 3. It uses a Docker engine and three containers that run as many applications: the MQTT broker, the NodeJS, and the Multichain node. The use of Docker permits DeCyMo to rely on *programmable IoT gateways.* Containers connect to each other using a virtual LAN inside Docker. Our IoT node is composed by an ESP8266, which is equipped with a microcontroller and a Wi-Fi chipset with full TCP/IP stack. The IoT node and the IoT gateway are associated with a Wi-Fi access point so sensors can publish data on the IoT gateway, which works as a broker. This MQTT Broker runs as a Docker container on the IoT gateway, it collects data from the IoT node using the MQTT Protocol. Another container runs NodeJS, which executes the monitoring and actuation logic. NodeJS is used to implement a lightweight web application for displaying graphs and statistics from the collected data and control actuators by sending commands to the IoT Node.

The IoT Sensor Node reads data from the sensor and publishes them to the MQTT Broker. The blockchain node inside the gateway subscribes to all topics on the broker and re-publish data on the appropriate Multichain stream and make them available to other nodes on the Multichain network. The NodeJS Container eventually stores data in a local database, handled by in another Docker container, then displays them on the website for real-time monitoring. The web application reports the status of the sensors and the relays, as well as the graph of temperature and humidity, as shown in Fig. 6. Actuators are automatically controlled by the control logic implemented in NodeJS. The user can bypass the control logic and force the manual control on actuators operating on the switches of the GUI, this sends commands to relays through the MQTT Broker, whose services are exposed on the LAN and the Internet.

Our setup considers a simple Multichain network composed by two virtual machines running on Virtualbox with 1 CPU and 2 GB of RAM on a physical machine equipped with an Intel i7-2600 CPU @3.4 GHz and 8 GB of DDR3 system memory @1333 MHz. We ran different experiments in several conditions, implemented a dedicated script on Python 2.7 using libraries that are specific for IoT and blockchain. The virtual machines run our customized Multichain client, which uses paho-MQTT library 1.3.1, to subscribe to the topic of our MQTT broker, and the Savoir library version 1.0.6, to republish data on Multichain streams. Savoir is also used to query Multichain internal data. We re-publish homogeneous data elements on the Multichain stream that are long 23 bytes: 4 bytes for data and remaining 19 bytes for meta-data. Figure 7 shows the number of queued transactions, i.e., the number
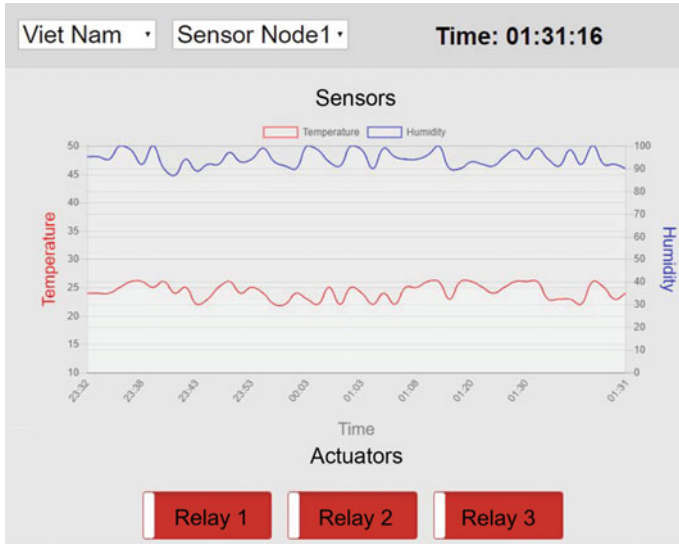
Fig. 6 The graphical user interface of the exemplary control panel with temperature and humidity graphs and relay status and control
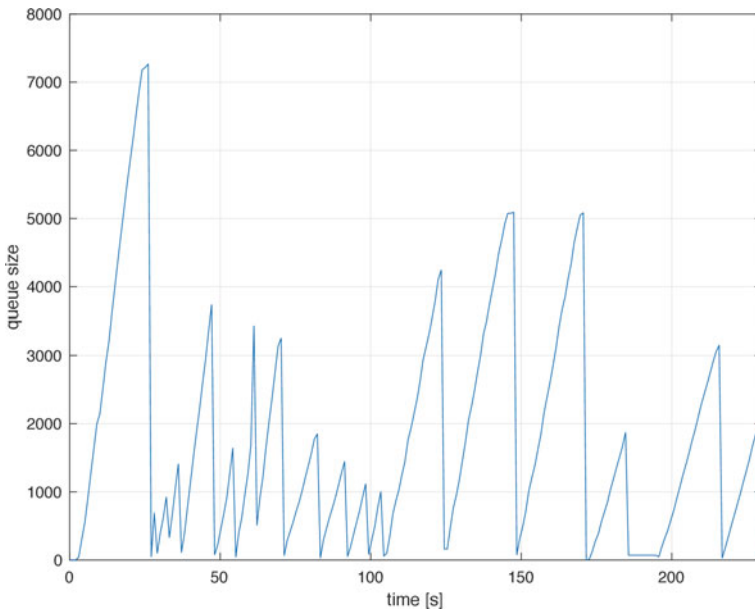

Fig. 7 DeCyMo experimental results, Queue size of uncommitted transactions overtime with a constant rate of new transactions
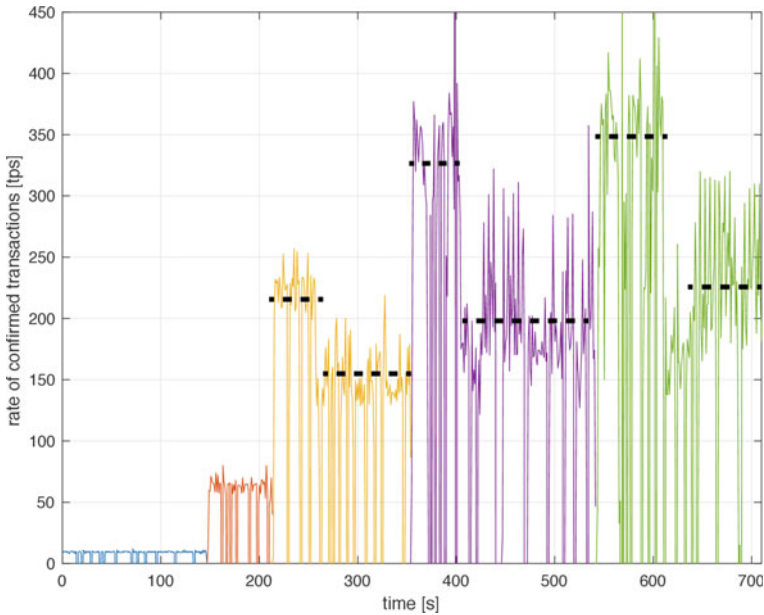
Fig. 8 Rate of confirmed transactions published on a stream under 5 DeCyMo re-publishing rates

of validated transactions that have not yet been confirmed on the chain, which corresponds to the memory pool size. We use a rate of 1000 publications per second on the stream. For retrieving this number of queued transactions, we use the function *getmempoolinfo* () of the Multichain API.

Figure 8 reports the rate of confirmed transactions where the publishing rate changes during consecutive time intervals. It worths noting that the publishing rate of IoT devices and nodes may be higher than the re-publishing rate (the latter corresponds to the transaction rate), in such conditions, the MQTT broker might also work as data buffer by adding a dedicated Docker container. In our experiment, we offered 10 transactions per second (tps) during the interval $[0; 145]s$, then in $[146; 214]s$ 100 tps, after that during $[215; 353]s$ 1000 tps, then in $[354; 542]s$ 10,000 tps, finally during $[543; 780]s$ 100,000 tps. The average rate of confirmed transactions increases with the rate of offered transactions. However, while for low rates all the offered transactions are confirmed, for higher values the rate of confirmed transactions falls behind the offered transactions demonstrating that the blockchain is not fast enough to address all requests of re-publishing IoT data. With an unconfirmed rate of 10 tps, there is a particular rate of 10 tps; at 100 tps in input correspond 60 tps in output. For higher rates, there are two phenomena: first, there is a transitory period when the queue is filled; then the confirmed rate slows down to a steady-state. Second, a saturation phenomenon is visible, as indicated by dashed horizontal segments in black: for rates that are respectively 146, 190, and 242 confirmed transactions per second in the stream.

Fig. 9 BRISK applied to a BlockSee frame to extract features and reveal rototranslations of the camera FoV. Frame shot on 12 Feb 2018

4.2 BlockSee Results for Video Surveillance and Blockchain

A visual representation of the application of the BRISK algorithm to BlockSee frames is reported in Figs. 9 and 10. These two figures show that, between the two shots, the camera was slightly turned to the right. In facts, the two frames taken at different days and BlockSee is able to recognize this translation. First, it extracts the selected features from a Region of Interest (RoI) defined once for the specific camera to remove eventual timestamps and watermarks, which are not significant for the understanding of the background and the analysis of the field of view.

5 Conclusions

DeCyMo and BlockSee are two architectural solutions that use the Blockchain technology together with IoT and video surveillance for applications in the smart cities of the future. DecyMo is a flexible architecture that handles IoT transactions and data streams for domestic and industrial applications maintaining devices, protocols, and actors as in legacy IoT deployments but rethinking the role of the IoT gateway. DeCyMo introduces new methodologies for disseminating data among nodes while maintaining privacy and access control. DeCyMo permits to add decentralized intelligence, modularity, flexibility, backward compatibility, as well as resilience to attacks and faults to the IoT arena. On the other hand, BlockSee uses the blockchain not only

Fig. 10 BRISK applied to a BlockSee frame to extract features and reveal rototranslations of the camera FoV. Frame shot on 15 Apr 2018

for guaranteeing video integrity but, more important, the integrity of the configuration of the whole surveillance system. Furthermore, videos are made available to police and court officials in case of events, without an uncontrolled disclosure and without the risk to be tampered with. The integrity of camera settings has an impact on privacy; it permits citizens to check possible violations of their privacy. BlockSee laid at the intersection between blockchain technology and computer vision and pointed our focus on the generalization of the transaction, including the transfer of information, here expressed as video segments.

References

Bogdanov A, Knežević M, Leander G, Toz D, Varıcı K, Verbauwhede I (2011) SPONGENT: a lightweight hash function. In: International workshop on cryptographic hardware and embedded systems, pp 312–325

Bradski G, Kaehler A (2008) Learning OpenCV: computer vision with the OpenCV library. O'Reilly Media, Inc.

Brutzer S, Höferlin B, Heidemann G (2011) Evaluation of background subtraction techniques for video surveillance. In: CVPR 2011, pp 1937–1944

Cassazione, Sezione Civile VI-2, Ordinanza Cron. 12139/15 R.G.N. 17939/13

Castro M, Liskov B (2002) Practical Byzantine fault tolerance and proactive recovery. ACM Trans Comput Syst 20(4):398–461

Cha S-C, Tsai T-Y, Peng W-C, Huang T-C, Hsu T-Y (2017) Privacy-aware and blockchain connected gateways for users to access legacy IoT devices. In: 2017 IEEE 6th global conference on consumer electronics (GCCE), pp 1–3

Christidis K, Devetsikiotis M (2016) Blockchains and smart contracts for the Internet of Things. IEEE Access 4:2292–2303

City of Helmstedt—Webcam of the City Hall

Coin Sciences Ltd. (2016) Multichain Streams. (Online). Available: https://www.multichain.com/blog/2016/09/introducing-multichain-streams/

Costin A (2016) Security of CCTV and video surveillance systems: threats, vulnerabilities, attacks, and mitigations. In: Proceedings of the 6th international workshop on trustworthy embedded devices, pp 45–54

Di Silvestre ML, Gallo P, Ippolito MG, Sanseverino ER, Zizzo G (2018) A technical approach to the energy blockchain in microgrids. IEEE Trans Ind Inform

Dorri A, Kanhere SS, Jurdak R (2017) Towards an optimized blockchain for IoT. In: Proceedings of the second international conference on Internet-of-Things design and implementation, pp 173–178

Dorri A, Kanhere SS, Jurdak R, Gauravaram P (2017) Blockchain for IoT security and privacy: the case study of a smart home. In: 2017 IEEE international conference on pervasive computing and communications workshops (PerCom Workshops), pp 618–623

Douceur JR (2002) The Sybil attack. In: International workshop on peer-to-peer systems, pp 251–260

Gallo P, Pongnumkul S, Nguyen UQ (2018) BlockSee: blockchain for IoT video surveillance in smart cities. In: Proceedings—2018 IEEE international conference on environment and electrical engineering and 2018 IEEE industrial and commercial power systems Europe, EEEIC/I and CPS Europe 2018

Gallo P, Nguyen UQ, Barone G, Van Hien P (2018) DeCyMo: decentralized cyber-physical system for monitoring and controlling industries and homes. In: IEEE 4th international forum on research and technology for society and industry (RTSI), RTSI 2018—proceedings

Gipp B, Kosti J, Breitinger C (2016) Securing video integrity using decentralized trusted timestamping on the bitcoin blockchain. In: MCIS, p 51

Hemlin Billström A, Huss F (2017) Video integrity through blockchain technology. Kth Royal Institute of Technology

Hofmann E, Rüsch M (2017) Industry 4.0 and the current status as well as future prospects on logistics. Comput Ind 89:23–34

IPFS—InterPlanetary File System. IPFS is the distributed web

Kosba A, Miller A, Shi E, Wen Z, Papamanthou C (2016) Hawk: the blockchain model of cryptography and privacy-preserving smart contracts. In: Proceedings of 2016 IEEE symposium on security and privacy, SP 2016, pp 839–858

Leutenegger S, Chli M, Siegwart RY (2011) BRISK: binary robust invariant scalable keypoints

Light RA (2017) Mosquitto: server and client implementation of the MQTT protocol. J Open Source Softw 2(13)

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110

Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system

Pinno OJA, Gregio ARA, De Bona LCE (2017) ControlChain: blockchain as a central enabler for access control authorizations in the IoT. In: GLOBECOM 2017–2017 IEEE global communications conference, pp 1–6

Rifi N, Rachkidi E, Agoulmine N, Taher NC (2017) Towards using blockchain technology for IoT data access protection. In: 2017 IEEE 17th international conference on ubiquitous wireless broadband (ICUWB), pp 1–5

Sanseverino ER, Silvestre MLD, Gallo P, Zizzo G, Ippolito M (2017) The blockchain in microgrids for transacting energy and attributing losses. In: 2017 IEEE international conference on Internet of Things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE Smart Data (SmartData), pp 925–930

Stanford-Clark A, Truongk HL (2013) MQTT for sensor networks (MQTT-S) protocol specification version 1.2, Nov 2013

StorJ—Distributed, Open-source, Encrypted, Sustainable Cloud Storage

Trón V, Scher A, Nagy AD, Felföldi Z, Johnson N, Swarm

Upmanyu M, Namboodiri AM, Srinathan K, Jawahar CV (2009) Efficient privacy preserving video surveillance. In: 2009 IEEE 12th international conference on computer vision, pp 1639–1646

Innovative Antennas for Next Generation of Communication Systems in Vietnam

Andrea Massaccesi, Michele Beccaria, Ho Manh Linh, Nguyen Huu Trung, Nguyen Khac Kiem, and Paola Pirinoli

Abstract Vietnam is one of the first countries in the world to successfully test the 5th Generation Mobile Network or simply 5G service. At the test, mobile internet speed has reached 1.7 Gbps, much faster than 4G and equivalent to the speed of commercial cables. 5G is recognized as the foundation for implementing the Industrial Revolution 4.0 to deploy digital economy, smart cities, smart environment, electronic citizens and government in countries like Vietnam. To satisfy all these requirements, a notable effort in the research and exploration of novel and revolutionary technologies was done and it is still on going. In particular, the need of high data rates can be satisfied increasing the available frequency bands and the spectral efficiency. At its turn, it can be obtained with the use of massive Multiple-Input Multiple-Output (MIMO) architectures. MIMO antennas are characterized by a number of radiating elements larger than the served user terminals, so that it is possible to enhance the throughput capacity. The design of an efficient MIMO system involves two main aspects, both considered here: the design of the actual antenna and of the beamforming algorithm. Different possible solutions are considered, and their performance are compared.

1 Introduction

The new standard of communication systems will be the 5G, that is expected to achieve higher performance than the previous protocols, i.e. 3 or 4G. The high demand of customers to exchange a huge quantity of data, the need of high speed of streaming and buffering (for HD movies, heavy online videogames etc.) should be served by a sophisticated technology which would be able to perform an increased rate of 1000 times with a data rate of 1 Gb/s, a latency lower than 1 ms, reduced

A. Massaccesi · M. Beccaria · P. Pirinoli (✉)
Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
e-mail: paola.pirinoli@polito.it

H. M. Linh · N. H. Trung · N. K. Kiem
Department of Aerospace Electronics, Hanoi University of Science and Technology, Hanoi, Vietnam

energy consumption and costs. Moreover, 5G will be a system also suitable to provide a lifeline communication in case of a natural disaster connecting at the same time a lot of users without the interruption of the services (Andrews et al. 2014; Dahlman et al. 2014). Another important aspect of the user experience is the Cloud service that is expected to exponentially grow (above 90%) (Vannithamby and Talwar, 2017) in the next years, because of the interest in users to manage big data rates everywhere with their "own device" without an expensive hardware.

In view of all these requirements, the use of a technology able to increase wireless channel capacity without additional power and spectrum is almost mandatory. One of the most promising solution is represented by the use of a "Multiple-Input Multiple-Output" (MIMO) antenna system, whose higher efficiency is due to the use of multipath to increase the channel throughput.

To understand how a MIMO system works, let consider the case of an array, made up of N elements, that sends data to M separate users. If its beamforming is designed in such a way that the antenna radiates several patterns (i.e. it shows a multibeam behavior), each one transmitting data just to a single user, and if each radiation pattern encodes a different data stream, the entire system can transmit simultaneously to all the users with a spectral occupancy equal to that of a single transmission. Something similar happens in MIMO antennas: if between the transmitter (Tx) and the receiver (Rx) there are several possible paths as sketched in Fig. 1, each of them can transmit a different data stream, increasing the channel throughput. This is called Spatial Multiplexing (SM). Within the transmitter, each data stream may be separately applied to the N transmitting antennas and propagates over multipath channel to a receiver. Error correction coding may be applied to each of the data streams separately or in a combined space-time coding method. MIMO technology also offers diversity gain and array gain. MIMO provides more robustness against signal interference by beamforming and/or nulling techniques. For a fixed overall
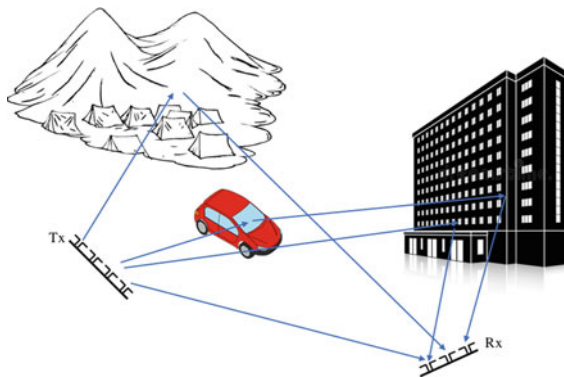

Fig. 1 Sketch of an outdoor environment, where several transmitting and receiving devices are present. The wireless connection is carried on taking advantage from the multiple reflections as in a MIMO system

transmitted power, the capacity offered by a MIMO configuration may scale with the increased Signal-to-Noise Ratio (SNR). Array gain increases the received Carrier-to-Noise-Ratio (CNR) with coherent signal combination at the receiving side. Channel State Information (CSI) is required at both transmitter and receiver with the help of pilot signals (Marzetta 2010).

Massive MIMO systems employing a large number of transmitting and receiving antennas (usually greater than 100 elements), have been of great interest in recent years because of their potential to dramatically improve spectral efficiency of future wireless systems and increase the transmission data rate through spatial multiplexing to deliver multiple streams of data within the same resource block (time and frequency). Massive MIMO systems exploit multipath propagation to improve the system reliability in terms of Bit Error Rate (BER) performance, without the expense of additional bandwidth. Moreover, massive MIMO can increase the power efficiency by scaling down the transmit power of each terminal inversely proportional to the number of elements of antenna array at base stations. It can steer multiple beams to a number of user ends to enhance SNR (Dahlman et al. 2014).

The design of a MIMO system involves both that of the actual antenna as well as that of the beamforming. An appropriate beamforming algorithm is needed to transmit the signal to desired users or receive the signal in a desired direction in order to overcome frequent LOS blockages and rich scattering multipath propagation in the mm-wave channel. The very small wavelengths of mm-wave frequencies allow to incorporate a large number of antenna elements in a compact structure to generate sharp beams corresponding to large array gains. For example, when using squared flat antenna array with half-wavelength uniform spacing between elements of the array in 20 GHz frequency band, the dimension of the 256-element array is only 12 cm^2.

The radiation features of the MIMO system can be obtained adopting a multibeam antenna, able to cover different angular regions and to radiate inside each of them a sufficiently high gain. They can be realized exploiting different technologies (Hong et al. 2017; Jokanovic et al. 2017) that could be grouped in two different classes, that of the antenna using active components and the other ones that are passive structures. In Active Multibeam Antennas (AMBAs) the direction of maximum radiation is controlled electronically: the resulting performance is generally excellent, but the use of beamforming networks and active components make the antenna difficult to design and expensive to manufacture. In Sect. 2, considerations on multibeam phased arrays and in particular on the beam forming are discussed.

The second alternative is represented by Passive MultiBeam Antennas (PMBAs) whose main advantages are that of a lower cost and easier manufacturing. Among the possible types of PMBA there is that of adopting a passive aperture antenna illuminated by the field radiated by a feed-array: even if the most traditional configurations use a shaped reflector or a dual reflector system to improve the antenna performance, other possible solutions consist in substituting the metallic, bulky reflector with a planar Reflectarray (RA) (Chou and Liu 2018), or in using a device working in transmission, i.e. a lens or a planar Transmitarray (TA) (Jiang et al. 2017; Plaza et al.

2015), presenting the advantage that it could be designed with a centered feeding system, since it does not suffer for problems related to the blockage.

In Sect. 3 some results on the possibility to use a TA for the realization of an efficient PMBA will be presented; the focus is on the requirements that a transmitarray has to satisfy to overcome its intrinsic limitations in multibeam applications. Some considerations on the effect of the unit cell will be firstly presented, then possible techniques for the efficient design of the TA will be discussed.

2 Multi-beam Phased Arrays

Multiple beamforming is a technique that uses array antennas to produce a number of simultaneously available adjustable radiation patterns, which can point to the desired coverage areas and minimize the impact of unwanted noise and interference, thereby improving the quality of the desired signal. Basically, beamforming is an optimal spatial filter (Marzetta 2010). Antenna arrays using a beamforming technique can eliminate interferences having a direction of arrival different from that of the desired signal. Multi-polarized arrays can also eliminate undesired signals having a polarization different from that of the desired signal, even if the signals have the same direction of arrival.

Beamformers use an array of antenna elements that are individually phased in such a way as to form beams (or nulls) in a desired direction. Let consider a multiple beamforming system. The inter-element distance is d. The system has N antenna elements for multiple beamforming. The system model is illustrated in Fig. 2. If
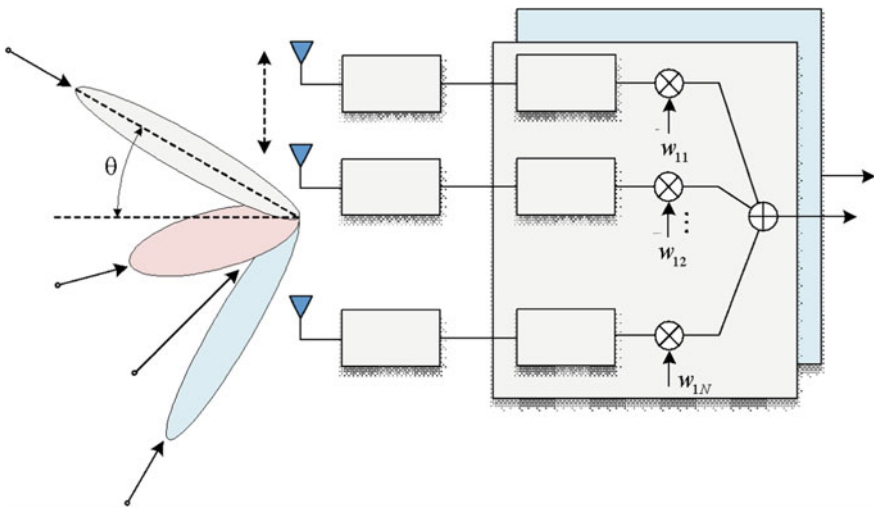
Fig. 2 Multi-beam phased array architecture

$s_i(t)$ is the signal associated to the i-th beam and the pointing angle associated with $s_i(t)$ is θ_i, the vector of array transmitting from N elements at time instant t is expressed as:

$$x_i(t) = a_i(\theta_i, \omega_i) s_i(t) \tag{1}$$

where $x_i(t) = [x_{i1}, x_{i2}, \ldots, x_{iN}]^H$ collects the signal samples received at the antennas and $a_i(\theta_i, \omega_i)$ is the steering vector according to $s_i(t)$:

$$a_i(\theta_i, \omega_i) = \left[1 e^{j\omega_i d \sin(\theta_i)/c}, e^{j\omega_i 2d \sin(\theta_i)/c}, \ldots, e^{j\omega_i (N-1)d \sin(\theta_i)/c}\right]^H \tag{2}$$

In the above equation, ω_i indicates the carrier frequency and c the speed of light. The steering vector depends on the direction of departure and the frequency. For simplicity, we denote $a_i(\theta_i, \omega_i)$ with a_i. The single beamforming model is therefore expressed as $x_i(t) = a_i s_i(t)$ and the multiple beamforming model becomes:

$$x = As \tag{3}$$

where $= [a_1(\theta_1, \omega_1), a_2(\theta_2, \omega_2), \ldots, a_P(\theta_P, \omega_P)]$, $x = [x_1, x_2, \ldots, x_P]^T$ and $s = [s_1, s_2, \ldots, s_P]^T$ for the P beams.

There are two general beamforming systems: narrow band beamforming and broadband beamforming. In narrow band beamforming model, the output signal $y_i(t)$ of the beamformer i ($i = 1, \ldots, P$) at the time instant t is obtained by linear combination of the signals of N elements as:

$$y_i(t) = \sum_{j=1}^{N} w_{ij}^* x_{ij}(t) \tag{4}$$

For broadband model, the output signal is expressed as:

$$y_i(t) = \sum_{j=1}^{N} \sum_{p=0}^{K-1} w_{ij,p}^* x_{ij}(t-p) \tag{5}$$

where $K - 1$ is the number of delay stages of the equalizer at each channel of the i-th element of the array. The transmitted signal is expressed as:

$$y_i(t) = w_i^H x_i(t) \tag{6}$$

where x_i is the signal vector. The vector w_i^H of length N represents the weights as:

$$w_i^H = \left[w_{i1}^*, w_{i2}^*, \ldots, w_{iN}^*\right] = \left[w_i^H\right]^* \tag{7}$$
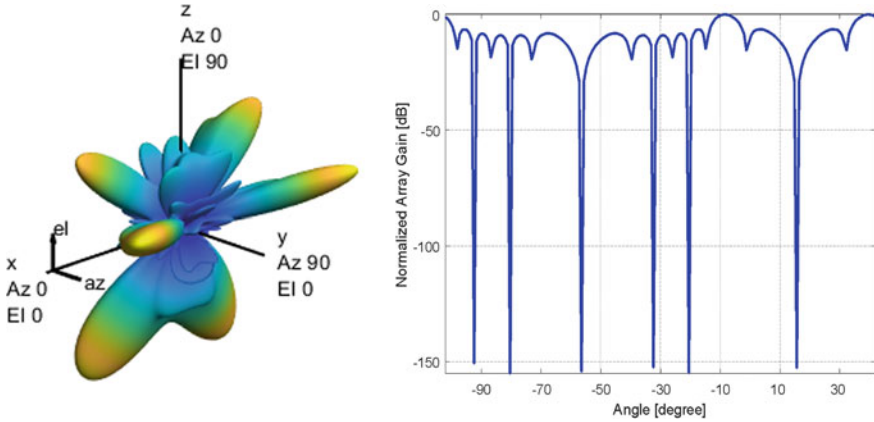
Fig. 3 3D beampattern (left) and array gain with nulling (right) of a multi-beam phased array

and so the response of the i-th beamformer is expressed as:

$$r_i(\theta_i, \omega_i) = \boldsymbol{w}_i^H \boldsymbol{a}_i \tag{8}$$

The beam pattern is defined as the squared magnitude of $r_i(\theta_i, \omega_i)$. Note that each of the weight coefficient in the weight vector impacts on the response of the beamformer in terms of time and space (Fig. 3).

2.1 Millimeter-Wave Beamforming

The obvious advantage of using millimeter-wave (mm-wave) bandwidth is the availability of under-utilized contiguous spectrum. These bands allow wide bandwidth up to 400 MHz and support up to 1.2 GHz of instantaneous bandwidth with carrier aggregation (Vannithamby and Talwar 2017).

Depending on the beamforming architecture, the beamforming weights required to form the beam pattern could be applied in the digital or analog domain. For MIMO systems, all-digital beamforming is usually implemented in the form of digital precoding that multiplies a particular coefficient to the modulated baseband signal per RF chain. On the other hand, for analog beamforming, complex coefficients are applied to control the directive beam through phase shifters with variable gain amplifiers if needed. In general, all-digital beamforming offers a higher degree of freedom and better performance at the expense of increased complexity and cost. Analog beamforming, on the contrary, is a simple and effective method of generating high beamforming gains from a large number of antennas but less flexible than the other method.

The all-digital beamforming architecture for sub-6 GHz MIMO platform and the advanced signal processing algorithms for spatial multiplexing was first extended to mm-wave beamforming. However, the baseband and MIMO precoding structure in the sub-6 GHz system cannot be applied directly into the mm-wave systems (Larsson et al. 2014) because the baseband precoding requires a complete dedicated RF chain with very high-speed ADC/DAC for each antenna element at the transmitter or the receiver, which is expensive and highly complex.

To obtain the advantages of massive MIMO and also provide high array gain to compensate for the large path loss at mm-wave bands, the hybrid beamforming structure has been proposed as an enabling technology for 5th Generation Mobile Network (Hong et al. 2017). This is a trade-off solution between flexibility and performance and bring the mm-wave beamforming to practical implementation. The hybrid beamforming structure for the mm-wave transmitter is illustrated in Fig. 4.

As it can see in the block diagram, N_s data streams are fetched to the MIMO precoding, normally spatial multiplexing and then the output is through the N_{RF} RF chains. The beamformer function splits the RF signal into P beams to feed each active element of the phased array. It provides high-resolution phase and amplitude weighting, which is needed to synthesize beamforming patterns and adaptively null potential interferences. The RF chain blocks for OFDM modulation often include ITTF, P/S (parallel to serial), CP (cyclic prefix insert) and DAC to form output signal. The mixer converts the baseband signal up to the carrier frequency. The final block of RF chain which is close to the antenna is the front-end, which contains a high-power
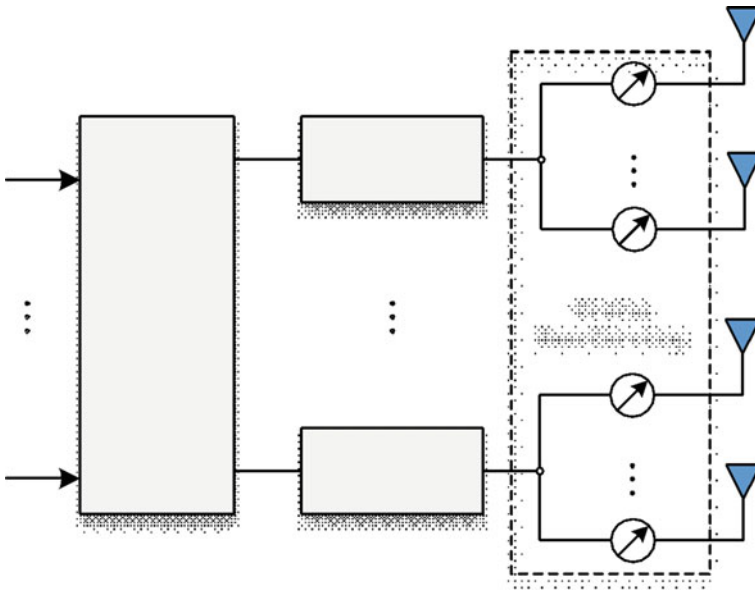

Fig. 4 Hybrid beamforming structure for the mm-wave transmitter

and high-efficiency transmit PA, a transmit/receive switch, and low-noise amplifier (LNA) (Jokanovic et al. 2017).

3 Multibeam Transmitarray

A transmitarray antenna is a structure composed by a planar transmitting surface and a feed source, that in fixed beam applications consists in most of the cases in a horn (see e.g. Abdelrahman et al. (2017) and references therein). The TA surface is designed in such a way that, for a given position and type of the feed, it transforms the incoming spherical wave front in a planar phase front. To overcome the lack of curvature, the TA surface is discretized with a proper number of Unit Cells (UCs), having size lower or equal to half of the wavelength ($\lambda_0/2$) at the design frequency f_0: each of them can control the phase and the amplitude of the transmission coefficient S_{21}, that are locally changed varying one or more geometrical parameters of the unit cell itself, in such a way to adjust the phase of the incident field, that at its turn depends on the relative position of the feed and the TA. The resulting antenna is generally characterized by a high gain in the selected direction of maximum radiation.

In the case of a multibeam system, things are a little more complex. If the TA is a passive configuration, the generation of the different beams can be obtained substituting the single feed with a focal feed array: as it can be seen from the drawing in Fig. 5, it generates simultaneously more beams, that impinge on the TA surface with a different incident angle and therefore the transmitarray radiates a beam in a different direction for each of them.

Ideally, the TA might be able to radiate in almost the same way in all the directions, i.e. the radiation patterns generated by beams arriving from the different incident direction must present almost the same gain, Half Power Beam Width (HPBW) and Side Lobe levels (SLL). Unfortunately, this is difficult to obtain, mainly for the two following reasons:

- the behavior of the unit cell depends on the direction of arrival of the incident wave;
- the procedure for the TA design described above guarantees a good performance for a fixed position of the feed, but not when it is moved.

In view of these considerations, the following analysis has been carried on. First, two different types of unit cell have been considered, and their effect of the entire antenna performance has been studied and compared; then the effect of the position of the feed with respect to the TA has been analyzed and finally the possibility to use a different design procedure has been exploited. Some of the obtained results are summarized in the following. They have been got with the full-wave numerical analysis of a reduced size transmitarray ($10\lambda_0 \times 10\lambda_0$), carried on with CST Microwave Studio® (MWS). Since the multibeam transmitarray is thought as the possible radiating system in a MIMO configuration for 5G, the selected operating frequency is
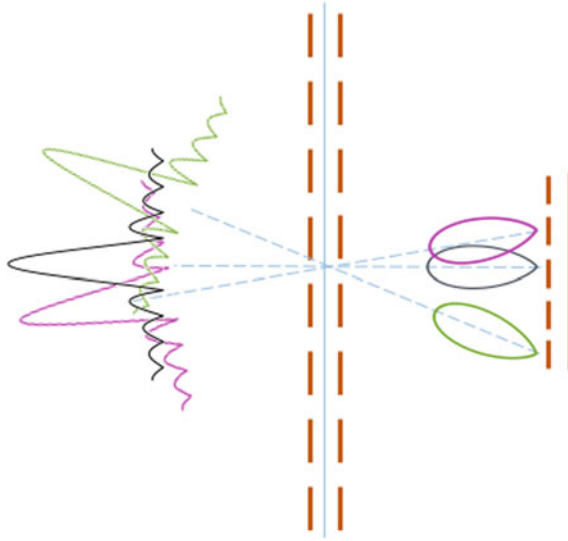
Fig. 5 Working principle of a multibeam TA: each element in the feed array radiates a beam in a different direction, that impinges on the transmitarray with a different angle of incidence that generates a beam in the consequent direction

in one of the bands that will be used for this service, and it is equal to $f_0 = 5.8$ GHz (Beccaria et al. 2018a, b).

3.1　Transmitarray Unit-Cell

In order to investigate the relation between the behavior of the unit cell as a function of the direction of arrival of the incident field and the capabilities of beam steering of the TA, two different UCs have been analyzed.

The first one is the classical multilayer printed structure shown in Fig. 6. Each layer is characterized by a relative dielectric constant $\varepsilon_r = 2.574$, a thickness $h = 0.5$ mm, negligible losses; the spacing between two following layers are chosen in such a way that the total distance between two printed elements is almost $\lambda_0/4$ at the design frequency f_0 to maximize the transmission coefficient. The unit-cell size is $\lambda_0/2$, with the aim to avoid grating lobes. The variation of the transmission coefficient is controlled through the side W of the printed elements, that have the shape of a Malta cross.

In Fig. 6, the variation of the phase and the amplitude of S_{21} with W, and for different angles of incidence, is plotted. They have been computed using the Floquet excitations in CST MWS®. Because of the presence of higher order mode resonances, it happens that for some values of W the phase of the transmission coefficient presents discontinuities and its magnitude strongly decreases, this phenomenon is already
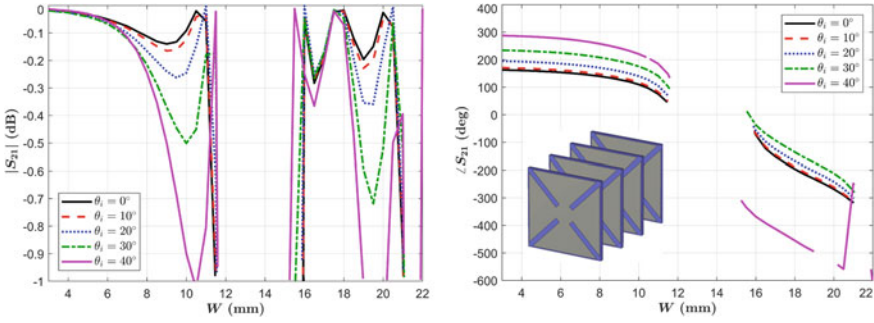
Fig. 6 Variation of the amplitude (left) and phase (right) of S_{21} versus W for several angles of incidence. Inset: unit cell

noticeable in the case of normal incidence ($\theta = 0°$), but it is more evident for higher values of the incident angles. To avoid the problems that such a behavior could have it has been decided to "cut" the curves in correspondence of these resonances, and the corresponding values of W are discarded. While in case of normal incidence the unit cell guarantees a phase variation of 370° and losses not greater than -1 dB, for other values of θ, both the phase and the amplitude of S_{21} changed drastically, and therefore a different behavior of a TA using this kind of element for different positions of the feed is predicted.

This hypothesis is confirmed by the radiation patterns shown in Fig. 7. They refer to a $10\lambda_0 \times 10\lambda_0$ transmitarray, designed using the multi-layer Malta cross as unit
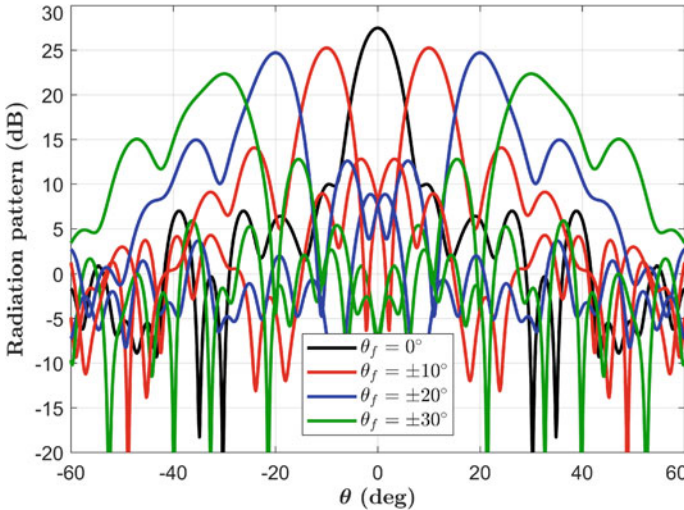
Fig. 7 Radiation patterns in the E-plane relative to the Malta cross TA, illuminated by a horn that rotates along an arc centered in the center of the transmitarray aperture
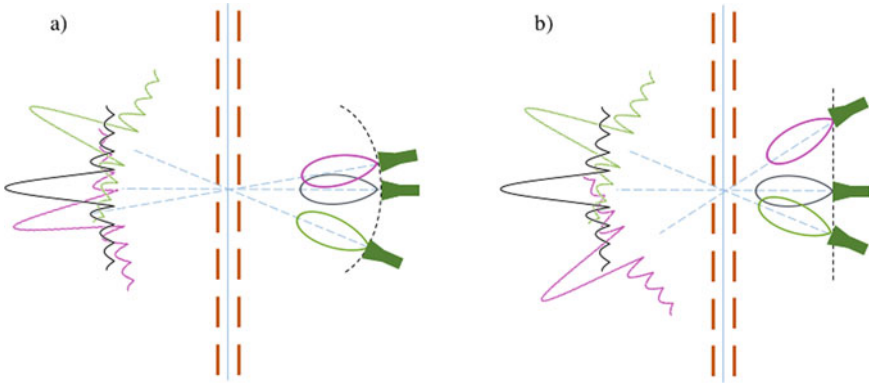
Fig. 8 The two analyzed TA configurations: (**a**) the feed rotates along a circular arc; (**b**) the feed moves parallel to the TA surface, as in a feed array

cell. The value of W for each of them has been fixed assuming the feed centered with respect to the TA and located at $9\lambda_0$ from its surface. To obtain the beam in different directions, the feed has been rotated along a circular arc, as it can be seen in Fig. 8, on the left.

The curves in Fig. 7 show that there is a noticeable degradation of the radiation patterns, starting from the broadside direction to the maximum scanning angle. The gain loss is of almost 5 dB, and the SLL increases drastically.

The second unit cell, considered here, has been recently introduced in Massaccesi and Pirinoli (2018) and experimentally validated in Massaccesi et al. (2018, 2019). It is composed by three dielectric layers of the same material (see inset in Fig. 9) and has a periodicity W in both x and y directions. The central layer is characterized by a square hole of height H_{sqh} and size d, that is used to control the phase of the transmission coefficient. The two external, identical layers present a hole that has a truncated pyramid shape with height H_{tap} and they act as a broadband matching
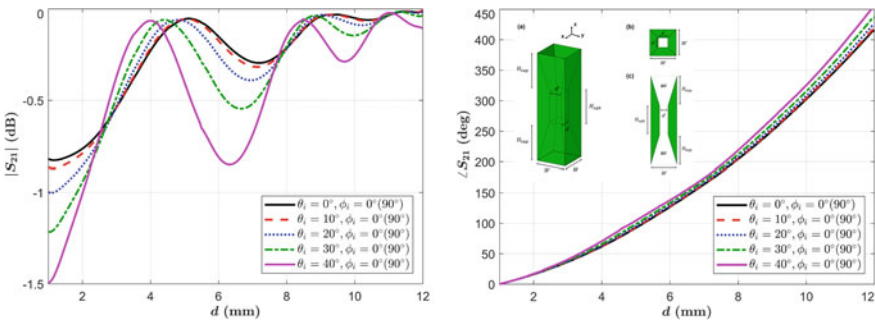

Fig. 9 Variation of the amplitude (left) and phase (right) of S_{21} versus W for several angles of incidence. Inset: unit cell

circuit, allowing to improve the performance of the dielectric unit-cell in terms of bandwidth and to reduce the dependence from the angle of incidence. Moreover, the dielectric UC has the advantage that it can be realized using Additive Manufacturing (AM) techniques, as demonstrated in Beccaria et al. (2018a, b).

In view of its advantages, the selected material for the realization of the TA has been the commercial 3D-printable Preperm® L700HF, which is characterized by $\varepsilon_r = 7$ and $\tan \delta = 0.001$. This material can be provided as filaments and printed through Fused Deposition Modeling (FDM) processes. The high value of the dielectric constant is favorable for reducing the thickness of the cell and therefore the dielectric losses. The periodicity of the cell is $W = 13$ mm $\simeq \lambda_0/4$, while the heights are $H_{sqh} = H_{tap} = 24$ mm $\simeq 0.46\lambda_0$, corresponding to a total cell thickness $T = H_{sqh} + 2H_{tap} = 72$ mm $\simeq 1.38\lambda_0$. The hole size d is varied from 1 to 12 mm to obtain a full phase range of 360°.

Also in this case, the transmission coefficient of the unit-cell has been computed using Floquet excitations in CST MWS®. Figure 9 shows the variation of the transmission coefficient magnitude (left) and phase (right) with d, evaluated for different incidence angles (θ_i). The magnitude slightly degrades when the angle of incidence increases, with a reduction not greater than 0.7 dB for $\theta_i = 40°$, while from the curves of the phase variation it emerges that all of them provide a full phase range of 360° and remains linear and quite similar for the different θ_i. They maintain almost the same slope until half of the range of d, then, the slope slight changes, arriving to have the maximum difference at the upper limit of the range. These results suggest that the designed unit cell is not particularly sensitive to the angle of incidence, making it a potential good candidate to be used in the design of transmitarrays with multibeam capabilities.

As for the other unit cell, a reduced size TA has been designed, with geometrical parameters (size, focal distance, feed movement) comparable with those of the TA using Malta cross UC. In Fig. 10, the radiation patterns in the E-plane (where the rotation of the feed occurs), for the scanning in the overall considered angular region are shown. From their comparison with the patterns in Fig. 7, it clearly emerges that the low sensitivity of the unit cell from the direction of arrival of the incident field has a strong impact on the antenna performance. In fact, here the radiation patterns stay almost the same till to $\theta_f = \pm 30°$, showing a small reduction of the gain and a reduced increase of the SLL.

3.2 Transmitarray Feed Movement

The results presented in the previous sub-section have been obtained assuming that the feed rotates along a circular arc, as drawn in Fig. 8a. If this hypothesis is sufficiently accurate for what was of interest there, it is necessary to point out that the elements of a feed array illuminate the TA surface in a different way. To mimic more accurately this situation, it is possible to consider a configuration as in Fig. 8b, where the feed moves linearly and always stays parallel to the TA. This means that
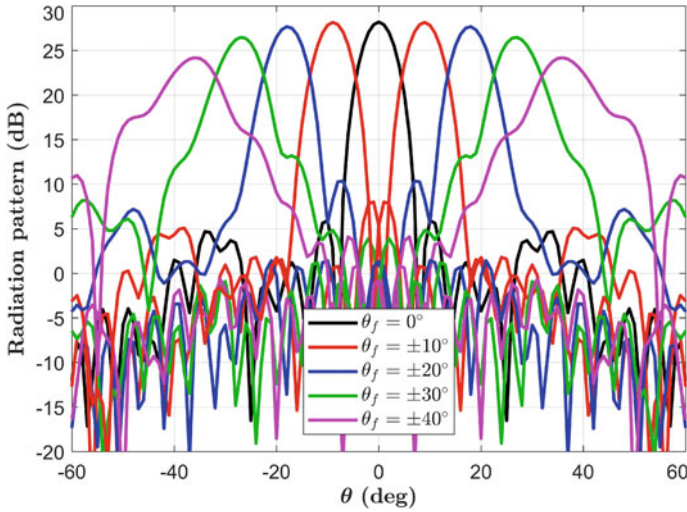
Fig. 10 Radiation patterns in the E-plane relative to the dielectric TA, illuminated by a horn that rotates along an arc centered in the center of the transmitarray aperture (configuration in Fig. 8a)

the distance between the feed and the transmitarray is no longer constant, and that it is lower for smaller angles of incidence.

Using this illumination mechanism, a new TA with the same size of the configurations discussed in the previous section has been designed and numerically analyzed. The maximum distance of the feed from the TA is $9\lambda_0$ and it is achieved for the scanning angles $\theta_f = \pm 30°$, while the minimum distance of $7.8\lambda_0$ occurs for $\theta_f = 0°$. Figure 11 shows a comparison between the radiation patterns in E-plane relative to the conventional TA described in the previous section and this new configuration. The radiation patterns corresponding to the same value of θ_f are very similar, with comparable HPBW and SLL.

3.3 Enhanced Techniques for the Design of a Multibeam TA

The above results prove the feasibility of a multibeam TA, provided that a suitable unit cell is adopted, but also highlight that if the TA is designed assuming a fixed position of the feed, its performance is good over a limited angular range. To overcome this limitation, a different design procedure must be adopted, able to take into account the different directions of arrival of the field radiated by the feed.

A possible solution is that of using a global optimizer; in Chap. 18 results on the application of two different algorithms to the design of a beam-scanning reflectarray are presented: similarly to what is described there, the same approach can be also used for the transmitarray, and it is expected that in this way it would be possible to
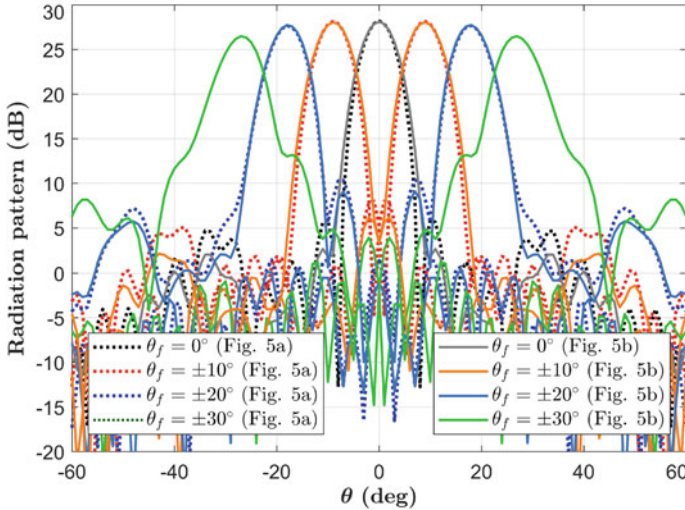
Fig. 11 Radiation patterns in the E-plane relative to the dielectric TA, illuminated by a horn that translates parallel to the TA as in Fig. 8b (dotted line) and compared with the case in which the horn rotates along a circular arc as in Fig. 8a (solid line)

enlarge the beam steering capabilities of the antenna at least to an angular range of 80°.

Another solution, investigated in the following, consists in designing the TA as it was a bifocal lens. The main steps of this procedure are the following:

- the feed is assumed to be in an offset position, with the main beam that forms a certain angle θ_1 with the z-axis, and the required phase distribution φ_1 is computed;
- then, the feed is rotated forming an angle θ_2 while keeping the same distance of the feed from the center, and another phase distribution φ_2 is computed;
- since the unit cells can compensate only one phase requirement, they cannot compensate both φ_1 and φ_2, and therefore each element can provide a phase φ_{BFM} that is the mean value between φ_1 and φ_2.

In Fig. 12, the phase distributions φ_1, φ_2 and φ_{BFM} relative to a TA adopting as unit cell the dielectric one, with size and focal ratio equal to those of the configurations already analyzed, are shown. The radiation patterns for several values of the angle of incidence, varying between $-40°$ and $+40°$ have been computed through a full-wave simulation of the entire antenna, and in Fig. 13 those in the E plane are plotted.

The solid curves represent the results obtained with the traditional design, the dotted are referring to the bifocal method. These results prove that the bifocal TA has greater scanning capabilities, for angles of incidence not larger than 30° the gain patterns have almost the constant values and low side lobes. When the direction of maximum radiation approaches 40°, the gain decreases somewhat, but the main beam keeps its shape, and the side lobes stay low.
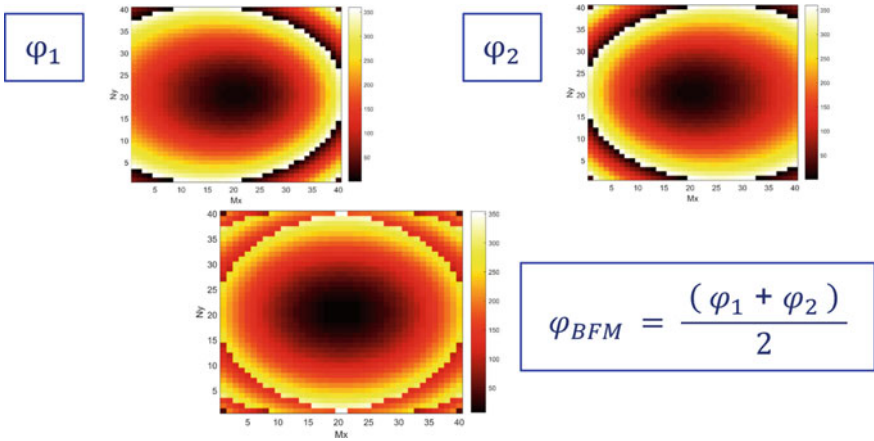
$$\varphi_{BFM} = \frac{(\varphi_1 + \varphi_2)}{2}$$

Fig. 12 Example of phase distributions used by the bifocal approach to design the TA
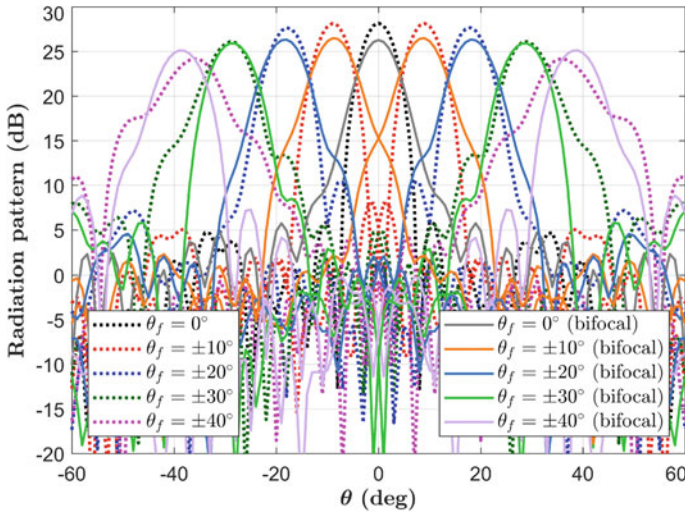

Fig. 13 Radiation patterns in the E-plane relative to the dielectric TA, designed with the bifocal method, compared with those obtained with the configuration in Fig. 8a

The variation of the gain as a function of the scanning angle is finally reported in Fig. 14, with that of the conventional TA illuminated with the feed moving along an arc (solid line) or linearly (dashed line). These latter are almost equal and have a higher gain in the broadside direction, but they decrease rapidly for the other directions of maximum radiation. On the contrary, the gain for the bifocal TA varies of almost 1 dB in the entire angular range.
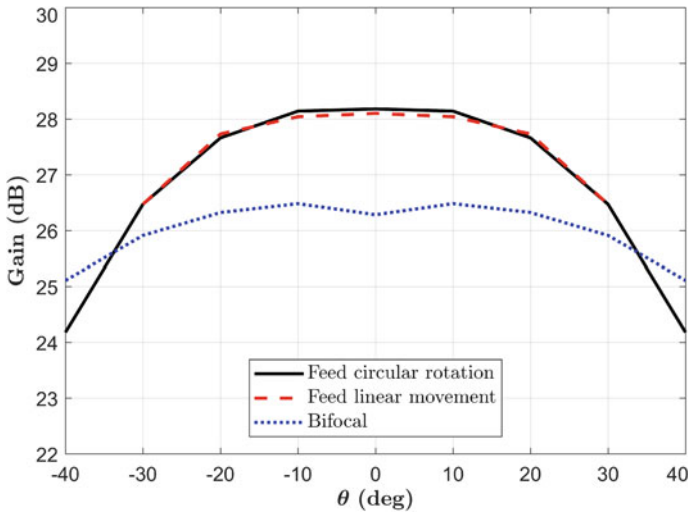
Fig. 14 Variation of the gain versus scanning angles. Comparison among the conventional TA, illuminated with a feed that moves along a circular arc (solid line) or linearly (dashed line), and the bifocal TA (dotted line)

References

Abdelrahman AH, Yang F, Esherbeni AZ, Nayeri P (2017) Analysis and design of transmitarray antennas. M&C Publishers

Andrews JG et al. (2014) What will 5G be? IEEE J Sel Areas Commun 32(6):1065–1082

Beccaria M, Massaccesi A, Pirinoli P, Manh LH (2018a) Multibeam transmitarrays for 5G antenna systems. In: 2018 IEEE seventh international conference on communications and electronics (ICCE), Hue, Vietnam, 18–20 July 2018, pp 217–221

Beccaria M, Massaccesi A, Pirinoli P, Kiem NK, Trung NH, Manh LH (2018b) Innovative MIMO antennas for 5G communication systems. In: 2018 IEEE international conference on environment and electrical engineering and 2018 IEEE industrial and commercial power systems Europe (EEEIC/I&CPS Europe, Palermo, Italy, 12–15 June 2018, pp 1–4

Chou H, Liu JW (2018) Synthesis and characteristic evaluation of convex metallic reflectarray antennas to radiate relatively orthogonal multibeams. IEEE Trans Antennas Propag 66(8):4008–4016

Dahlman E et al (2014) 5G wireless access: requirements and realization. IEEE Commun Mag 52(12):42–47

Hong W et al (2017) Multibeam antenna technologies for 5G wireless communications. IEEE Trans Antennas Propag 65(12):6231–6249

Jiang M, Chen ZN, Zhang Y, Hong W, Xuan X (2017) Metamaterial-based thin planar lens antenna for spatial beamforming and multibeam massive MIMO. IEEE Trans Antennas Propag 65(2):464–472

Jokanovic B, Milosevic V, Radovanovic M, Boskovic N (2017) Advanced antennas for next generation wireless access. In: 2017 13th international conference on advanced technologies, systems and services in telecommunication (TELSIKS), Nis, 2017, pp 87–94

Larsson EG, Edfors O, Tufvesson F, Marzetta TL (2014) Massive MIMO for next generation wireless systems. IEEE Commun Mag 52(2):186–195

Marzetta TL (2010) Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wireless Commun 9(11):3590–3600

Massaccesi A, Pirinoli P (2018) Enhancing the bandwidth in transmitarray antennas using tapered transmission line matching approach. In: 12th European conference on antennas and propagation (EuCAP 2018), London, UK, 9–13 Apr 2018. IET, pp 1–4

Massaccesi A et al (2018) 3D-Printable dielectric transmitarray with enhanced bandwidth at millimeter-waves. IEEE Access 6:46407–46418

Massaccesi A, Dassano G, Pirinoli P (2019) Beam scanning capabilities of a 3D-printed perforated dielectric transmitarray. Electronics, MDPI 8(4):379

Plaza EG, Costa JR, Fernandes CA, León G, Loredo S, Las-Heras F (2015) A multibeam antenna for imaging based on planar lenses. In: 9th European conference on antennas and propagation (EuCAP), Lisbon, pp 1–5

Vannithamby R, Talwar S (2017) Towards 5G: applications, requirements and candidate technologies. John Wiley & Sons

Optimization Strategies for Efficient Antenna Design

Michele Beccaria, Ho Manh Linh, Andrea Massaccesi, Alessandro Niccolai, Nguyen Huu Trung, Nguyen Khac Kiem, Riccardo Zich, and Paola Pirinoli

Abstract The conventional procedures for the design of antenna systems often yield a solution that is sub-optimal: this occurs especially when the configuration to be designed is complex, as could be a multi-beam antenna, since many parameters have to be managed and several goals, sometimes competing each other, have to be achieved or when the antenna constraints are not known "a priori" but could be just estimated during the optimization process itself. A possible solution to overcome these limitations consists in using a global optimizer. Here, several different approaches are considered and compared in terms of their performances when applied to different classes of problems.

1 Introduction

Optimization is a key aspect in the engineering system design. It can be faced with different methods: one of most basic and widely used method is the so-called *trial and error*, a typical approach to problem solving based on testing several system configurations until a reasonable solution is reached (Young 2009). This method can be hardly applied to very complex problems because the solution space is too large.

Evolutionary Optimization Algorithms (EAs) are a very valid alternative because they do not require any special knowledge about the shape of the cost function, i.e. the function that mathematically models the problem to be solved (on the other hand, this is required for linear and quadratic programming (Dantzig 2016)), they do not

M. Beccaria · A. Massaccesi · P. Pirinoli (✉)
Department of Electronics and Telecommunications, Politecnico di Torino, Torin, Italy
e-mail: paola.pirinoli@polito.it

H. M. Linh · N. H. Trung · N. K. Kiem
Department of Aerospace Electronics, Hanoi University of Science and Technology, Hanoi, Vietnam

A. Niccolai · R. Zich
Department of Energy, Politecnico di Milano, Milan, Italy

need any initial guess close to the desired minimum (as is required by traditional non-linear optimizer like the simplex method (Kolundzija and Olcan 2006)). Moreover, Evolutionary Algorithms do not require neither continuity nor derivability of the cost function (Simon 2013).

Several evolutionary optimization algorithms are available in literature. The first one that was introduced is the Genetic Algorithm (GA) (Goldberg and Holland 1988). It was firstly implemented for binary problems, and then it has been also adapted successfully to real-valued problems. Another important EA is the Particle Swarm Optimization (PSO), an algorithm native for real-valued problems (Kennedy 2010).

These two algorithms are the most established ones, but many others have been implemented and have shown very good optimization capabilities, finding a proper tradeoff between exploration and exploitation (Simon 2013). Some of them are the Differential Evolution (DE) (Storn and Price 1997), the Biogeography Based Optimization (BBO) (Simon 2008), the Fireworks Algorithm (FA) (Tan and Zhu 2010).

Antenna optimization problems often involve many degrees of freedom, whose management becomes difficult when a deterministic procedure is applied for its design and therefore the resulting configuration could be a sub-optimal solution. For this reason, EAs have been widely adopted to problems involving the design of an antenna, either of a single radiating element either of the entire system. The most widely EAs used for the optimization of antenna systems are undoubtedly the GA and the PSO. However, the increasing complexity of the problems to be optimized pushes the researchers to investigate and to develop new approaches with improved features, in terms of convergence, computational cost and reliability. In Sect. 2, the capabilities of two innovative algorithms, the Social Network Optimization (SNO) (Grimaccia et al. 2017; Niccolai et al. 2015) and the M_QC_{10}-BBO, that is an enhanced version of the BBO (Mussetta et al. 2013; Pirinoli et al. 2017), are studied when applied to a complex antenna problem as the design of a scanning beam Reflectarray.

If global optimizers as the evolutionary algorithms are efficient tools for the design of a system that has to satisfy requirements that are known "a priori" and that are used to mathematically model the optimization problem itself, in some other cases it occurs that not all the problem constraints are defined. This is for instance what happens in modern wireless communication systems such as fifth-generation (5G) mobile communication systems, that utilize massive MIMO (multiple input multiple output) configurations (Elkashlan et al. 2014). MIMOs consist of two sets of antennas, one placed at the base station and the other assembled in a small device, such as a mobile phone, where design space is very limited. Between these two sets of antennas there is a rich scattering multipath fading environment that must be modeled in the most accurate way, since it affects the system performance and in particular the definition of the beamforming algorithm used by the antennas to generate the multibeams: as better as it can predict the behavior of the channel as higher is the performance of the entire system. However, the structure of the propagation environment is not known "a priori", but could only be communicated to the transmitter by the receiver, after its estimation, and this operation drastically increases the complexity of the algorithm; therefore, it becomes necessary to adopt an optimization process able to

estimate in a sufficiently accurate way the propagation environment, trying at the same time to reduce its computational cost (Elkashlan et al. 2015). In Sect. 3 some examples of the most commonly optimization techniques adopted in this contest are presented.

2 Evolutionary Optimization-Based Design of a Beam-Scanning Antenna

Optimizing the design of a system means to select a proper set of physical variables (called design variables and represented by the vector d) that can modify the performance of the system itself.

The design variables are represented in the EA by means of a set of optimization variables (vector x) that can be mapped in a biunivocal manner to the design variables themselves.

The performance of the system should be mapped in a numerical value (the cost C) that should be minimized; this mapping procedure is performed by the *cost function.* The entire relation that occurs between the cost C and the optimization variables is called *objective function.*

The optimal set of parameters obtained after the optimization process is indicated with the symbol b.

In the specific case of the antenna optimization, the objective function consists in most of the cases by the calculation of the antenna radiation pattern and the cost represents how far it is from the constraints it is asked to satisfy. At each iteration, the algorithm computes the cost relative to the different set of optimization variable that represents the EA population. This means that the objective function is computed, throughout the optimization process, thousands of times, and therefore it is necessary to describe the problem with a model sufficiently accurate but also not too much computationally expensive. In case of an antenna, this means that the use of a full-wave approach for its characterization is not feasible, since it would result in an unaffordable increase of the computational time. Therefore, other techniques, as the representation of the radiating elements with their equivalent circuital model or the use of approximated methods for the computation of the antenna radiating features are generally adopted.

Once the optimal solution is determined, at the end of the EA process, it is finally analyzed with a full-wave approach, to check the correctness of the optimization. In Fig. 1 the entire optimization process is sketched. It is a general scheme, independent from the type of the used algorithm and the type of antenna problem to be solved.

Fig. 1 Summary of the
optimization procedure
adopted in the antenna
optimization
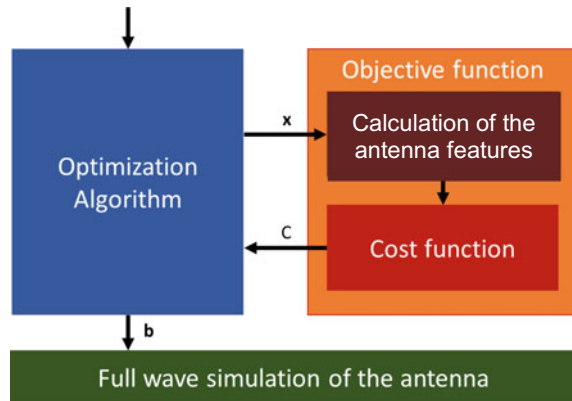
2.1   SNO

The Social Network Optimization (SNO) is an evolutionary algorithm based on
the interaction and influence process that takes place in online social networks
(Grimaccia et al. 2017).

The algorithm is based on a population of individuals (users of the social network)
and on the posts available online. These two data structures are the basis of the
algorithm and they drive the information exchange process: users write posts with
their opinions, these posts are read by other individuals and they are influenced and,
thus, they change their opinions (Niccolai et al. 2015).

Each user is characterized by a set of *opinions*. This is an array with the same size
of the optimization variables vector. The interaction among users is driven by two
kinds of networks: a friend network and a trust network. For defining these networks,
a friend list and a reputation list are associated to each user. The friend list is the set
of user ID of all the friends of the user, while the reputation list contains a reputation
value for all the other users. All these user's information evolves during time thanks
to the algorithm operators.

The interaction in network is a key aspect of this algorithm because it drives
the tradeoff between exploration and exploitation. The two interaction networks are
deeply different: the friend network is symmetric, the connections are particularly
strong, and its evolution depends on events in the real world. On the other hand,
the trust network is not symmetric, i.e. trust is not reciprocal, the connections are
weaker, and its evolution depends only on online relations (Cameron et al. 2011).

Also, the posts are complex data structures: their main content is the status, i.e.
the transposition of the opinion of the user. Other metadata are added to this content:
the ID of the user that posted it, the posting time, and a visibility value. This is very
important because posts with high visibility can be read more often, and, thus, their
impact on the other users is greater.

The post is the structure that interacts with the objective function: in fact, the
status is the candidate solution (the vector of the optimization variables **x**), while the
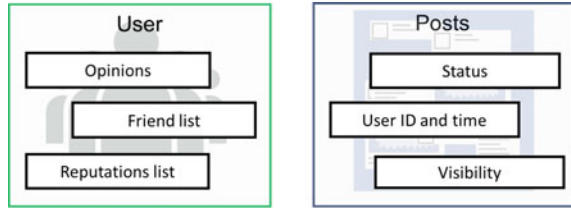
visibility value is the cost value C assigned by the objective function to the candidate solution. Figure 2 shows a summary of the data structures of SNO.

The evolution of these data structures is obtained by means of several algorithm operators. The main ones are the linguistic transposition, the reputation update, the trust network creation, the friend network evolution, the influencer selection, the crossover, and the idea contagion function.

Figure 3 shows all the operators in the loop of the algorithm. The red squares are the operators, while the blue and green rectangles represent posts data structures and users' ones, respectively. The orange box is the optimization problem, and its interaction with the algorithm is underlined by the dashed lines.

2.2 M_QC_{10}-**BBO**

The M_QC_{10}-BBO is a modified version of the BBO, aimed to improve its performance (Mussetta et al. 2013; Pirinoli et al. 2017). As the original BBO, it takes inspiration from the migration process of the species among islands, also named habitats. Each of them represents a possible solution, while the species are the optimization variables and the Habitat Suitability Index (HIS) represents the goodness of a solution, i.e. its fitness score.

The moving of species among islands is regulated by how good the habitat itself is, i.e. by its HIS, through the emigration and the immigration rates, that in the standard BBO depend linearly on the number of species present in an island. A high HIS characterized a crowded habitat and therefore the less performing species are pushed to migrate to more favorable islands, i.e. where the number of species already present is small. On the other hand, a low performing habitat has a low emigration rate and high immigration rate and therefore it receives species from better islands. The consequence of this mechanism is that the best habitats have low probability to share information with other high performing solutions and to converge to the optimum.

Despite of these considerations, the algorithm has been shown to perform well for some specific problems, like power distribution (Bhattacharya 2010), while for antenna optimization it stagnates in local minima (Mussetta et al. 2012). For this reason, some improvements have been introduced in BBO (Mussetta et al. 2013).
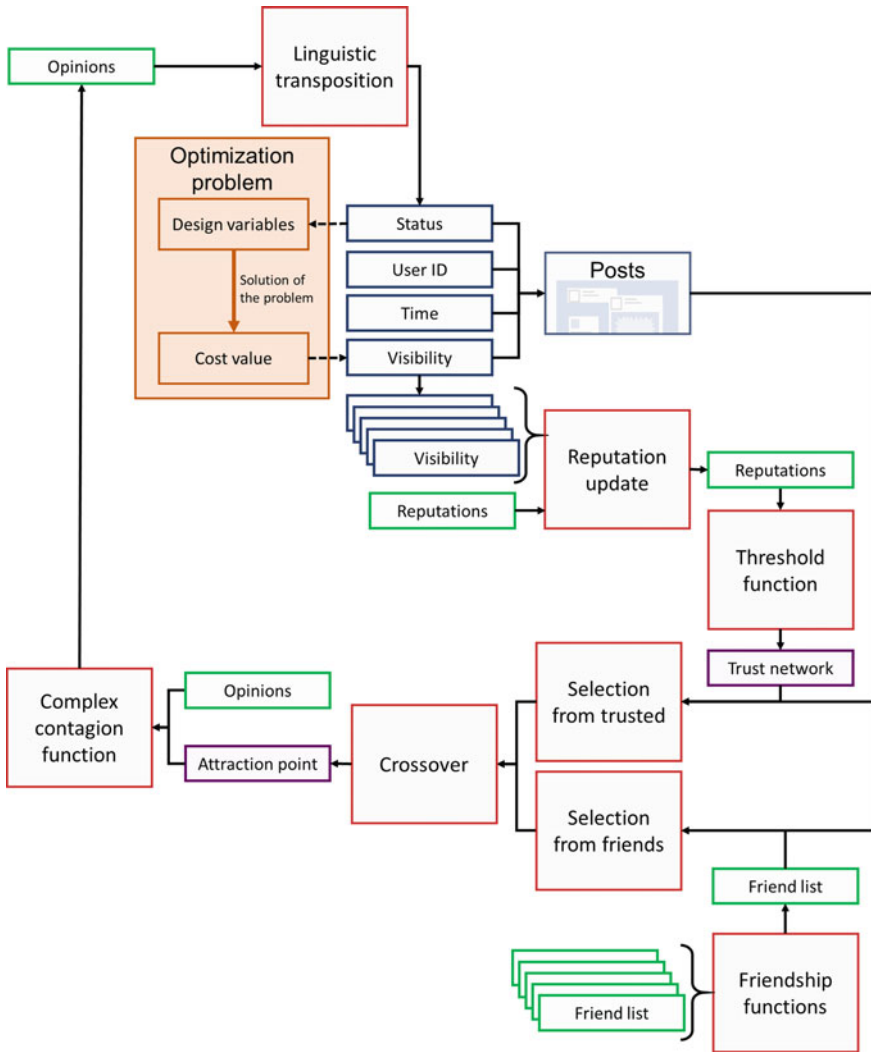
Fig. 3 SNO evolution of data structures by means of the algorithm operators

The first one is relative to the model that describes the relation between the number of species in a habitat and the emigration and immigration rates, not yet assumed to be linear but quadratic.

Moreover, it has been observed that the BBO is too much "deterministic" and therefore a novel implicit restart procedure, named "cataclysm" has been introduced: when the best among all the solutions did not improve in the last N iterations, all the solutions are destroyed (cataclysm) and new ones are randomly generated. In order

to preserve the best solution, elitism applies and no cataclysm occurs again before at least N generations have passed.

The information about the use of the quadratic model as well as of the cataclysm and the number of iterations between two following events is codified in the name of the algorithm: "M_Q" indicates that the migration model is the quadratic one, "C" informs about the presence of the cataclysm while its numeric subscript n is related to the minimum number N of iterations between two cataclysms, being $N = 5n$. Several tests have been done to fix the value of n: if it is too small the cataclysms are too close each other and the algorithm becomes a random search, while if too many iterations occur it falls back into the BBO. At the end, it was concluded that a reasonable value is $n = 10$.

2.3   Application of SNO and $M_Q C_{10}$-BBO to the Design of a Beam-Scanning Reflectarray

Here, the SNO and the $M_Q C_{10}$-BBO are applied to the design of a planar beam-scanning Reflectarray (RA).

RAs represent a good compromise between reflector antennas and arrays, and therefore they are suitable for high gain applications (Huang 2017; Nayeri et al. 2018). The antenna system is composed by a feed (usually a horn antenna) and a planar reflector. In order to compensate the absence of curvature of a conventional reflector, the surface of the RA is divided in a proper number of square unit cells (with size lower than or equal to $\lambda_0/2$, being λ_0 the wavelength computed at the design frequency f_0). Each unit cell includes one or more re-radiating elements, whose selected geometrical parameters are varied to compensate the phase of the incident field and to obtain the desired radiation pattern.

The RA design procedure consists in a first step in which the unit cell behavior is analyzed and the curves representing the variation of the phase of the reflected field in correspondence of the unit cell, assumed embedded in a period structure, as a function of the selected geometrical parameters, are obtained. Then, the reflectarray is designed considering the feed characteristics, the feed position and the desired features of the entire antenna: the proper value of the geometrical parameters of each cell is selected in such a way they provide the phase necessary to obtain the desired radiation pattern for a fixed position of the feed.

Things are different when a beam-scanning antenna, based on the use of a passive RA, must be designed. In fact, in this case, the degrees of freedom of the unit cells are not enough to provide the proper value of the phase for each of the considered directions of maximum radiation that are obtained for instance rotating the feed along a circular arc, as sketched in Fig. 4. It is therefore necessary to find the values that give the best trade-off among the radiation patterns for all the considered pointing directions.
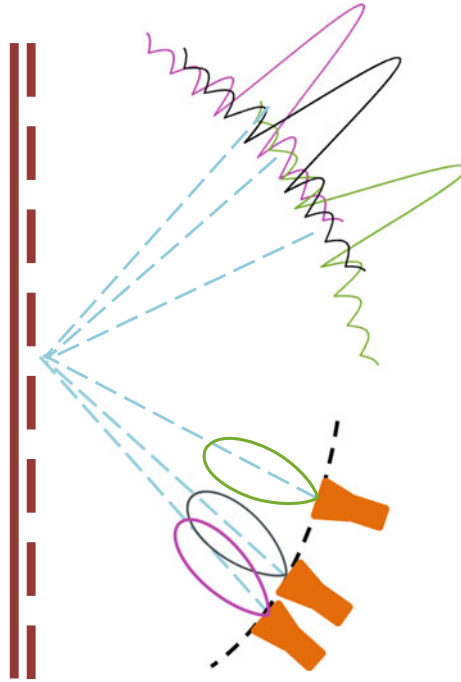
Fig. 4 Side view of a scanning beam reflectarray: the feed moves along a circular arc, and changing its position also the direction of maximum radiation varies

In view of the large number of unknowns whose values must be determined, this is a typical problem that could be conveniently solved with the aid of an evolutionary algorithm.

Here, a microstrip reflectarray is considered, made of a dielectric substrate with relative dielectric constant $\varepsilon_r = 2.55$ and height $h = 0.8$ mm. The RA surface is discretized with 24×24 unit cells with size $\lambda_0/2$. The feed is a standard horn, located at a distance $F = 10.8\lambda_0$ from the center of the reflectarray, and it can be moved along an arc, covering the angular range that corresponds to have a beam scanning between $-40°$ and $+40°$.

Each unit cell includes a square patch (see inset in Fig. 5), whose size d is used to control the phase and the amplitude of the reflection coefficient S_{11} provided by the cell itself. Their variation with d is plotted in Fig. 5. S_{11} is computed with CST MWS$^{®}$, carrying on a full-wave simulation of the unit-cell embedded in a periodic structure and for normal incidence.

The total number of variables, i.e. the size d of the 24×24 square patches, thanks to the double symmetry, reduces to 148.

The estimation of the radiation patterns has been performed with the Aperture Field Method (Huang 2007), discretizing the space with 91 samples along the θ
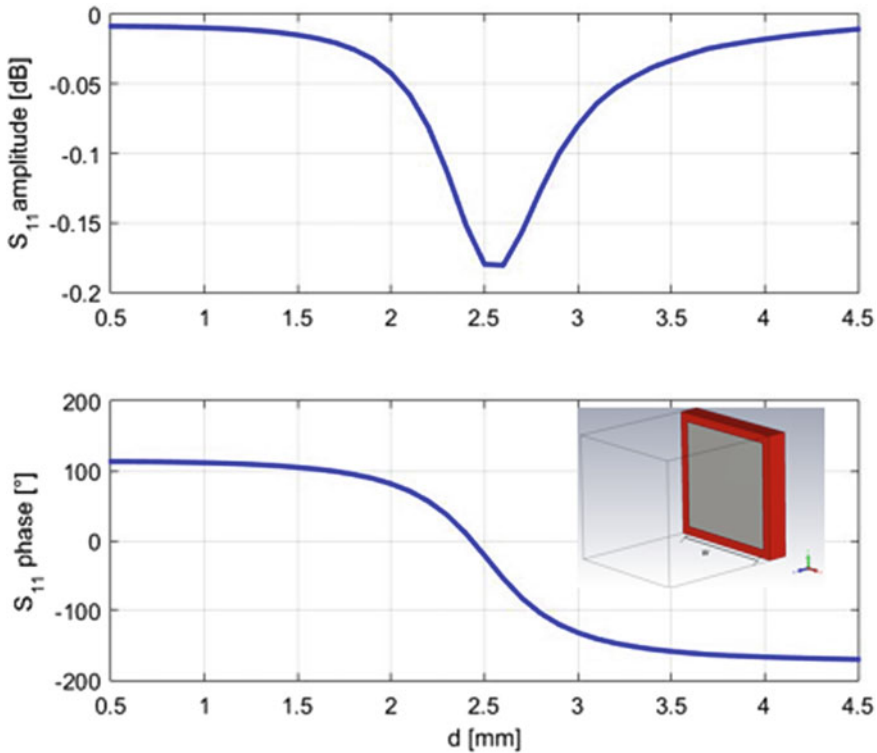
Fig. 5 Variation of the phase (top) and amplitude (bottom) of the reflection coefficient provided by each unit cell, as a function of the size d of the square patch. Inset: sketch of the unit cell

coordinate and 35 along ϕ. This choice is a very good trade-off between the accuracy of the model and the computational time required by the optimization.

The aim of the optimization is to achieve the most constant gain for all the scan angles, controlling at the mean time the Side Lobe Level (SLL). This has been codified imposing that in correspondence of four different directions of maximum radiation the antenna radiation pattern stays below a different mask, in which the SLL and the half beam width are limited. The values of the constraints have been selected for achieving a good behavior of the gain. It is important to notice that the numerical code used for the optimization is not able to calculate the gain that can be assessed only after the entire optimization process with the full-wave simulation of the antenna itself.

The optimization process has been stopped after 50,000 objective function calls for both the algorithms. Moreover, since the EAs are stochastic techniques, 12 independent trials have been done to check their reliability.

Figure 6 shows the curves of convergence of the two algorithms. The thin lines represent the convergence of each independent trial, while the darker thick line is the average value. It is possible to show that the M_QC_{10}-BBO has a faster convergence in
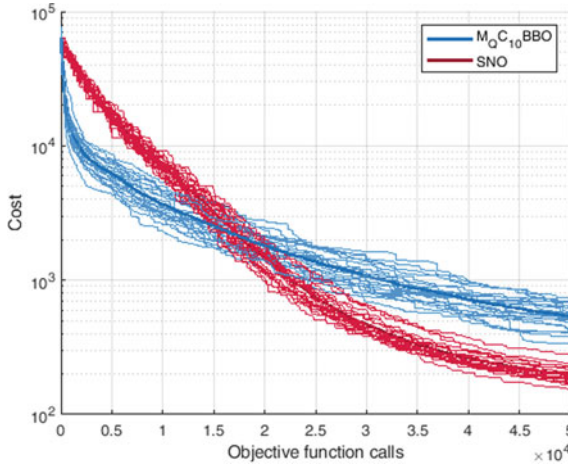
Fig. 6 Curves of convergence of the two algorithms

the first third of the optimization process, while SNO can keep a good convergence rate for all the time. Both the algorithms have a good reliability, as proved by the low dispersion of the curves corresponding to the single trials.

Figure 7 shows the optimal RA geometries obtained with the M_QC_{10}-BBO and the SNO.

To verify the effectiveness of the optimization process, these two configurations have been finally analyzed with the full-wave method implemented in CST MWS® and their radiation patterns are computed for the feed four different positions that correspond to the values of the scanning angle equal to 10°, 20°, 30° and 40°. The radiation patterns relative to the RA optimized with the M_QC_{10}-BBO are plotted in Figs. 8 and 9, while in Figs. 10 and 11 that for the SNO configuration are shown. In all the figures, the masks used for the optimization are also represented.
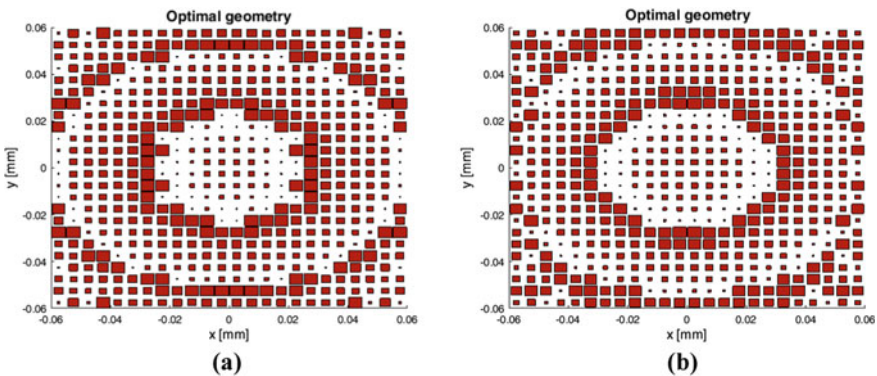

Fig. 7 Optimal geometries obtained by (**a**) M_QC_{10}-BBO and (**b**) SNO
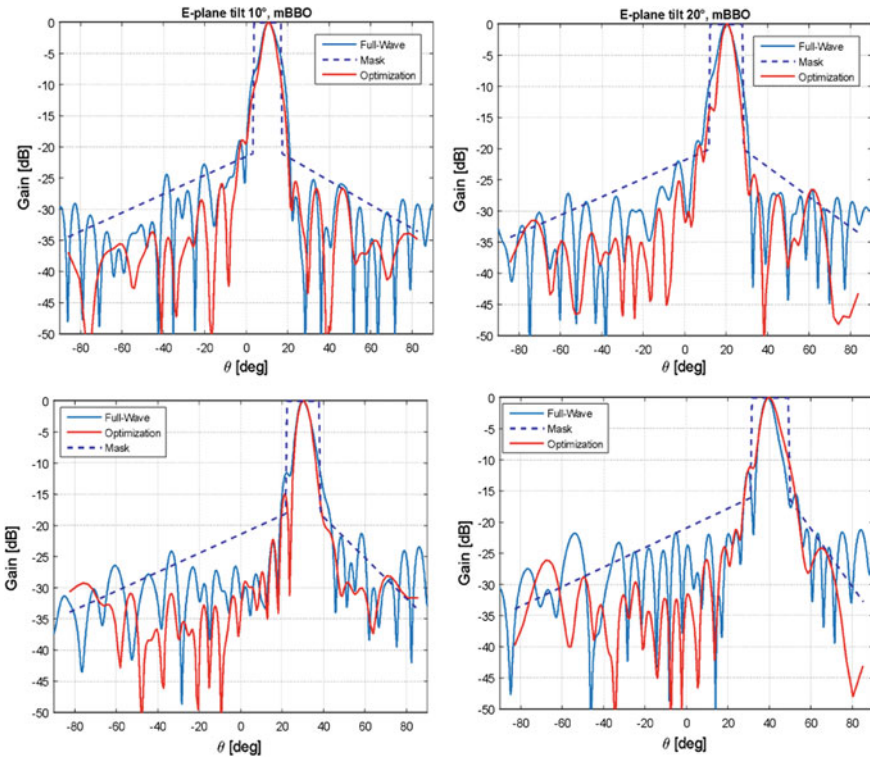
Fig. 8 Radiation patterns for the M_QC_{10}-BBO solution, E-plane

It is possible to see that in all the cases the radiation patterns well satisfy the constraints, since they are almost everywhere below the masks.

Finally, in Fig. 12 it is represented the gain as a function of the scan angle for the two antennas designed with the two algorithms. From this plot clearly emerges that both the RAs present a variation of the gain lower than 2 dB in the entire scanning range, and this definitively proves the effectiveness of the adopted optimization procedure, and of the two algorithms.

3 Optimization for MIMO Antenna

Antennas for modern wireless communication systems such as 5G systems make use of small element sizes at high frequencies to incorporate a larger number of radiating elements. The number of antenna elements in massive MIMO configurations is defined to be larger than 100 elements. These antenna arrays are essentials for beamforming operations that play a vital role in modern wireless communication systems.
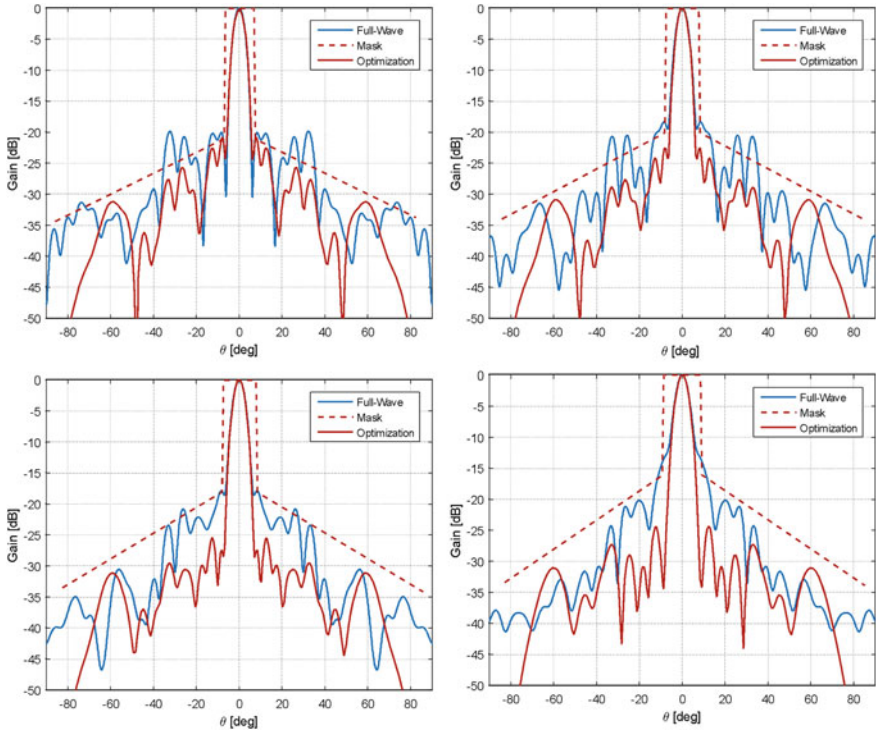
Fig. 9 Radiation patterns for the M_QC_{10}-BBO solution, H-plane

The beamforming system consists of an array processor and a linear or planar array of radiating elements. It is basically a spatial filter that is used to radiate or receive the maximum power to/from a predefined direction. Recently, two possible types of beamforming systems have been studied as candidates for next generation wireless networks: they are the digital beamformer and the hybrid analog-digital beamformer whose block diagrams are shown at the top and the bottom of Fig. 13 (Sohrabi et al. 2013).

Digital beamforming allows multiple stream transmission and serves multiple users simultaneously. However, it may not always be ideally suited for practical implementations regarding 5G applications. The very high hardware complexity may significantly increase size, cost, energy consumption and the complexity of the integration in mobile devices. On the other hand, it is well-suited for use in base stations. Hybrid beamforming has been proposed as a solution able to combine the advantages of both analog and digital beamforming architectures (Hefnawi 2019; Sohrabi and Yu 2013; Larsson 2015).

The criteria used to optimize beamforming systems include maximization of the Signal to Noise Ratio (SNR), Minimum Mean Squared Error (MMSE), Linearly
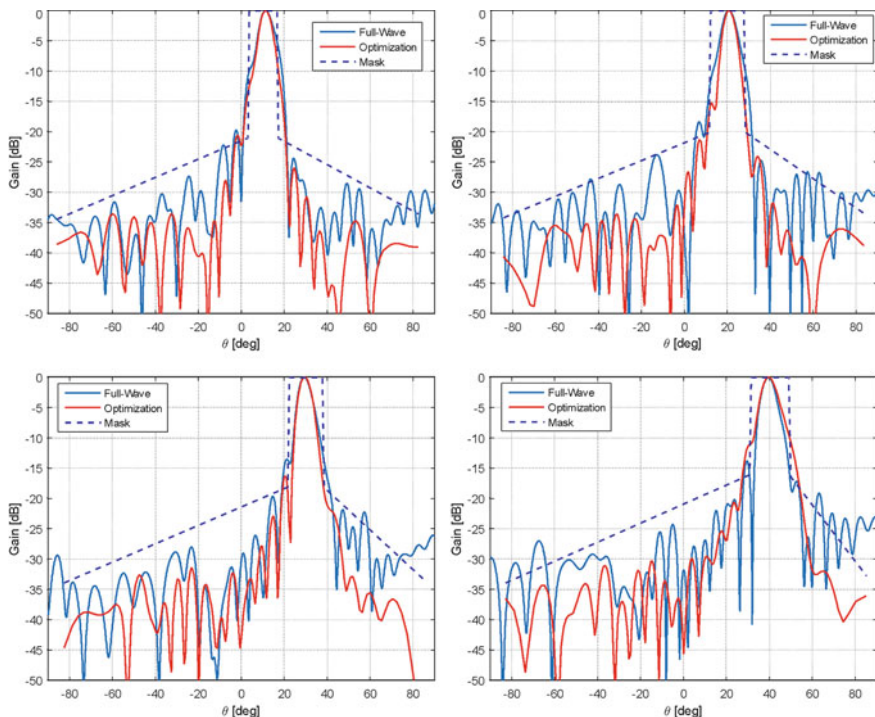
Fig. 10 Radiation patterns for the SNO solution, E-plane

Constrained Minimum Variance (LCMV) and robust optimization by Min-max criteria (Y.C. Eldar 2005).

Consider a beamforming system with M antenna elements. Denoting $s(t)$ the transmitted signal, the Direction Of Arrival (DOA) angle of the wavefront plane associated with $s(t)$ is θ; therefore, the vector of array observation from M elements at time instant t is expressed as:

$$\boldsymbol{x}(t) = \boldsymbol{a}(\theta, \omega)s(t) + \mathbf{i}(t) + \mathbf{n}(t) \tag{1}$$

where $\boldsymbol{a}(\theta, \omega)$ is the steering vector, $\mathbf{i}(t)$ is the interference and $\mathbf{n}(t)$ is a Gaussian noise vector.

$$\boldsymbol{a}(\theta, \omega) = \left[1 e^{j\omega d \sin(\theta)/c}, e^{j\omega 2d \sin(\theta)/c}, \ldots, e^{j\omega(M-1)d \sin(\theta)/c} \right]^{H} \tag{2}$$

representing d the distance between the two elements, ω is the carrier frequency and c the speed of light. The steering vector depends on the direction of arrival and the frequency. For simplicity, we denote $\boldsymbol{a}(\theta, \omega)$ with $\boldsymbol{a}$. The beamforming model is expressed as
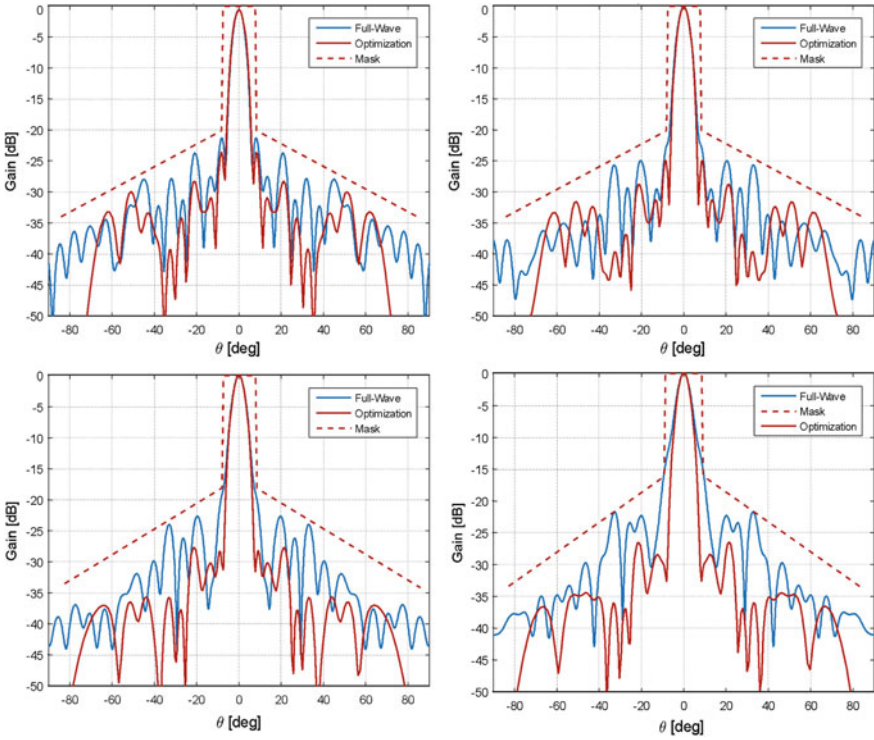
Fig. 11 Radiation patterns for the SNO solution, H-plane
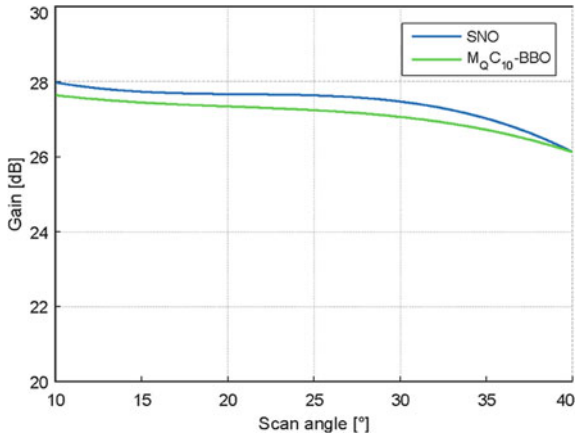

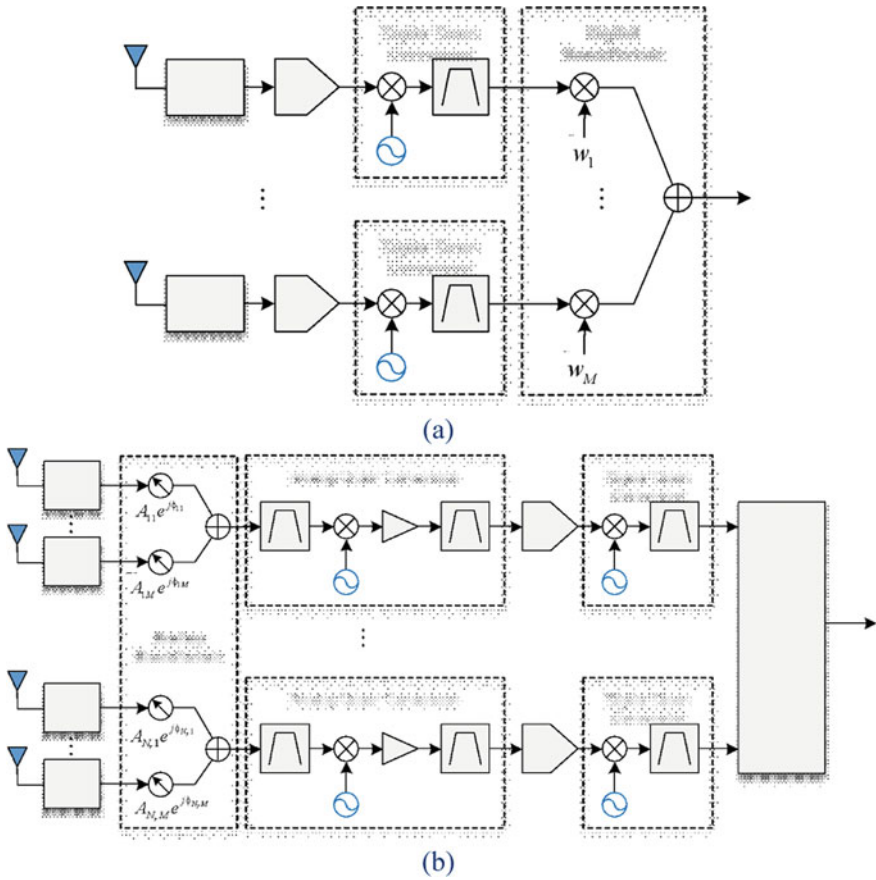**Fig. 12** Gain as a function of the scan angle

Fig. 13 Digital (**a**) and hybrid (**b**) beamforming architecture at receiver

$$\mathbf{x}(t) = \mathbf{a}s(t) + \mathbf{i}(t) + \mathbf{n}(t) \tag{3}$$

There are two general beamforming systems, including narrowband beamforming and broadband beamforming. In the narrowband beamforming model, the output signal of the beamformer at time instant t is $y(t)$ obtained by the linear combination of signals of M elements as:

$$y(t) = \sum_{i=1}^{M} w_i^* x_i(t) \tag{4}$$

For the broadband model, the output signal is expressed as:

$$y(t) = \sum_{i=1}^{M} \sum_{p=0}^{K-1} w_{i,p}^* x_i(t-p) \tag{5}$$

where $K-1$ is the number of delay stages at each channel of the ith element of the array. The output signal is expressed as:

$$y(t) = \boldsymbol{w}^H \mathbf{x}(t) \tag{6}$$

where $\boldsymbol{x}$ is the received signal vector. The vector $\boldsymbol{w}$ of length M represents the weights as:

$$\boldsymbol{w}^H = \left[w_0^*, w_1^*, \ldots, w_{K-1}^*\right] = \left[\boldsymbol{w}^T\right]^* \tag{7}$$

The response of a single beamformer is therefore:

$$r(\theta, \omega) = \boldsymbol{w}^H \boldsymbol{a} \tag{8}$$

The beampattern is defined as the squared magnitude of $r(\theta, \omega)$. Note that each of the weights in the vector $\mathbf{w}$ impacts to the response of the beamformer in terms of time and space. Output power or variance of estimated signal is determined as:

$$E\{|y|^2\} = \boldsymbol{w}^H E\{\boldsymbol{x}\,\boldsymbol{x}^H\}\mathbf{w} \tag{9}$$

where $E\{.\}$ denotes the mean value.

If the signal is wide sense stationary, the covariance matrix $\mathbf{R}_x = E\{\boldsymbol{x}\,\mathbf{x}^H\}$ is statistically independent over time. Although the signal statistic is not often stationary, the performance of the optimized beamforming is computed under the hypothesis that it is wide sense stationary.

The covariance matrix of the narrow band signal $s(t)$ at frequency ω_0 is:

$$\mathbf{R}_x = \sigma_s^2 \boldsymbol{a}(\theta, \omega_0)\boldsymbol{a}^H(\theta, \omega_0) = \sigma_s^2 \boldsymbol{a}\boldsymbol{a}^H \tag{10}$$

where σ_s^2 is the average signal power.

Beamforming is an important technique in array processing in order to optimize the desired signal while minimizing interferences. The design of the beamformer under statistically optimal method requires statistical properties of the source and the statistical characterization of the channel.

3.1 Maximization of SNR

The weight vector is the solution of the maximization of the SNR problem:

$$\boldsymbol{w}_{MaxSNR} = \arg\max_{\boldsymbol{w}} \frac{\boldsymbol{w}^H \boldsymbol{R}_s \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{R}_n \boldsymbol{w}} \tag{11}$$

General solution $\mathbf{w}_{MaxSNR}$ requires both $\mathbf{R}_s = E\{\mathbf{ss}^H\}$ and $\mathbf{R}_n = E\{\mathbf{nn}^H\}$ i.e. the covariance matrices of signal and noise. Depending on applications, the calculation of $\mathbf{R}_s$ and $\mathbf{R}_n$ are different. For example, $\mathbf{R}_n$ can be estimated during the absence of signal, while $\mathbf{R}_s$ is estimated from signal and known DOA by Eq. (10). Note that the SNR does not change if the weight vector is multiplied by a scaling factor. Since the steering vector $\boldsymbol{a}(\theta, \omega)$ is fixed for a given signal, it is possible to choose a weight vector to satisfy $\boldsymbol{w}^H \boldsymbol{a}(\theta, \omega) = c$, where c is a constant. The problem of the SNR maximization can be rephrased in terms of minimizing the interference:

$$\boldsymbol{w}_{MaxSNR} = \arg\max_{\boldsymbol{w}}\{SNR\} = \arg\min_{\boldsymbol{w}}(\boldsymbol{w}^H \boldsymbol{R}_n \boldsymbol{w}), s.t. \ \boldsymbol{w}^H \boldsymbol{a} = c \tag{12}$$

Using the method of Lagrange multipliers, the solution of the equation (Simon 2008) is therefore:

$$\boldsymbol{w} = c\frac{\mathbf{R}_n^{-1}\boldsymbol{a}}{\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a}} \tag{13}$$

3.2 Minimum Mean Squared Error

The MMSE method minimizes the error signal between the transmitted signal and a reference signal $d(t)$. In this model, desired user assumes to transmit this reference signal, i.e. $s(t) = \alpha d(t)$ where α is the amplitude of the reference signal $d(t)$ and $d(t)$ is known at the receiver. The aim of the output signal of the beamformer is to track the reference signal. MMSE method seeks the weight to minimize the average error signal power:

$$\boldsymbol{w}_{MMSE} = \arg\min_{\boldsymbol{w}} E\{|e(t)|^2\} \tag{14}$$

The average error signal power is:

$$\begin{aligned}
E\{|e(t)|^2\} &= E\{|\boldsymbol{w}^H \boldsymbol{x}(t) - d(t)|^2\} \\
&= E\{|\boldsymbol{w}^H \boldsymbol{x}\boldsymbol{x}^H w - \boldsymbol{w}^H \boldsymbol{x}d^* - \boldsymbol{x}^H \boldsymbol{w}d + dd^*|^2\}
\end{aligned}$$

$$= \boldsymbol{w}^H \mathbf{R} \mathbf{w} - \boldsymbol{x}^H \mathbf{r}_{xd} - r_{xd}^H \boldsymbol{w} + dd^* \tag{15}$$

where $\mathbf{r}_{xd} = E\{\mathbf{x}d^*\}$.

Computing the derivative of 15 with respect to $\boldsymbol{w}^H$ and setting it to zero it is possible to obtain:

$$\frac{\partial E\{|e(t)|^2\}}{\partial \mathbf{w}^H} = \mathbf{R}\mathbf{w} - \mathbf{r}_{xd} = 0 \tag{16}$$

whose solution is:

$$\mathbf{w}_{MMSE} = \mathbf{R}^{-1}\mathbf{r}_{xd} \tag{17}$$

known as optimal Wiener filter. This method requires reference signal to train the beamformer.

3.3 Linearly Constrained Minimum Variance

The LCMV method consists in minimizing the output power of the beamformer methods. It keeps the response according to the direction of arrival of the desired signal fixed in order to preserve the desired signal while minimizing the impact of the undesired components, including noise and interference that come from other directions than the desired one.

The output response of the signal source with direction of arrival θ and frequency ω is determined by $\boldsymbol{w}^H \boldsymbol{a}(\theta, \omega)$. Linear constraint for the weights satisfies the constraint $\boldsymbol{w}^H \boldsymbol{a}(\theta, \omega) = c$, where c is a constant to ensure that all the signals with frequency ω come from the direction of arrival θ are passed with response c. The minimization of the output due to interferences is equivalent to minimizing the output power (minimum output power):

$$\boldsymbol{w}_{MoP} = \arg \min_{\mathbf{w}} E\{|y|^2\} = \arg \min_{\mathbf{w}}\{\boldsymbol{w}^H \boldsymbol{R}_x \boldsymbol{w}\}, s.t. \ \boldsymbol{w}^H \boldsymbol{a}(\theta, \omega) = c \tag{18}$$

Using the method of Lagrange multipliers, find $\min_{w}[\mathcal{L}(\mathbf{w}; \lambda)]$, where:

$$\mathcal{L}(\mathbf{w}; \lambda) = E\{|\boldsymbol{w}^H \mathbf{x}|^2\} + \lambda(\boldsymbol{w}^H \boldsymbol{a} - c) = w^H \boldsymbol{R}_x w + \lambda(\boldsymbol{w}^H \boldsymbol{a} - c) \tag{19}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}^H} = \mathbf{R}_x w + \lambda \boldsymbol{a} \tag{20}$$

The solution of the equation is:

$$\mathbf{w}_{LCMV} = -\lambda \mathbf{R}_x^{-1} \boldsymbol{a} = c \frac{\mathbf{R}_x^{-1} \boldsymbol{a}}{\boldsymbol{a}^H \mathbf{R}_x^{-1} \boldsymbol{a}} \tag{21}$$

In practice, the uncorrelated noise component ensures $\boldsymbol{R}_x$ is invertible. If $c = 1$, the beamformer is called Minimum Variance Distortionless Response (MVDR) beamformer. The solution of the MVDR beamformer is equivalent to maximizing the SNR solution by replacing $\sigma^2 \boldsymbol{a}(\theta, \omega) \boldsymbol{a}^H(\theta, \omega) + \boldsymbol{R}_n$ by $\boldsymbol{R}_x$ and applying the inverse matrix lemma $[\mathbf{A} + \mathbf{BCD}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}[\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1}]^{-1}\mathbf{D}\mathbf{A}^{-1}$, we have:

$$\mathbf{R}_x^{-1} = \left[\mathbf{R}_n + \sigma_s^2 \boldsymbol{a}\boldsymbol{a}^H\right]^{-1} = \mathbf{R}_n^{-1} - \frac{\mathbf{R}_n^{-1}\boldsymbol{a}\boldsymbol{a}^H\mathbf{R}_n^{-1}}{\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a} + \sigma_s^{-2}}$$

$$\Rightarrow \tag{22}$$

$$\mathbf{R}_x^{-1}\boldsymbol{a} = \mathbf{R}_n^{-1}\boldsymbol{a} - \frac{\mathbf{R}_n^{-1}\boldsymbol{a}\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a}}{\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a} + \sigma_s^{-2}} = \mathbf{R}_n^{-1}\boldsymbol{a} - \frac{(\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a})\mathbf{R}_n^{-1}\boldsymbol{a}}{\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a} + \sigma_s^{-2}} = \left(\frac{\sigma_s^{-2}}{\boldsymbol{a}^H\mathbf{R}_n^{-1}\boldsymbol{a} + \sigma_s^{-2}}\right)\mathbf{R}_n^{-1}\boldsymbol{a}$$

$$= c\mathbf{R}_n^{-1}\boldsymbol{a} \tag{23}$$

3.4 Robust Optimization by Min-max Criteria

A robust optimization algorithm finds the beamforming weight solution that minimizes the worst case (the best of the worst conditions) on a set of signals $\mathbf{r}_f$ (in time domain or frequency domain for frequency beamformers) and by $MSE(\mathbf{r}_f, \hat{\mathbf{r}}_f)$ criteria, with a constant $q > 0$ and a positive define matrix Q (Alluhaibi et al. 2016). The problem is stated in min-max optimization as:

$$\boldsymbol{w}_{MNM} = \arg\min_{\mathbf{w}_r} \max_{\mathbf{r}_f : \mathbf{r}_f^H \mathbf{Q}\mathbf{r}_f \leq q^2} MSE(\mathbf{r}_f, \hat{\mathbf{r}}_f)$$

$$= \arg\min_{\mathbf{w}_r} \max_{\mathbf{r}_f : \mathbf{r}_f^H \mathbf{Q}\mathbf{r}_f \leq q^2} E\left\{\left|\hat{\mathbf{r}}_f - \mathbf{r}_f\right|^2\right\}$$

$$= \arg\min_{\mathbf{w}_r} \max_{\mathbf{r}_f : \mathbf{r}_f^H \mathbf{Q}\mathbf{r}_f \leq q^2} \left\{\boldsymbol{w}_r^H \overline{\mathbf{R}}_x \boldsymbol{w}_r + |\mathbf{r}_f|^2 \left|1 - \boldsymbol{w}_r^H \mathbf{A}_r\right|^2\right\} \tag{24}$$

where the covariance matrix of observation vector is $\overline{\mathbf{R}}_x = E\left\{(\mathbf{r}_f \hat{\mathbf{r}}_f)\right\}$.

$$\overline{\mathbf{w}}_{MNM} = \arg\min_{\mathbf{w}_r} \max_{\substack{\mathbf{r}_f : \mathbf{r}_f^H \mathbf{Q}\mathbf{r}_f \leq q^2 \\ \overline{\mathbf{R}}_x : \sum \max\left\{tr\left(\overline{\mathbf{R}}_x \overline{\mathbf{R}}_x^H\right)\right\}}} MSE(\mathbf{r}_f, \hat{\mathbf{r}}_f) \tag{25}$$

The problem solution is determined by the method of Lagrange multipliers:

$$\overline{\mathbf{w}}_{MNM} = q^2 \frac{\overline{\mathbf{R}}_x^{-1} \mathbf{A}_r}{1 + q^2 \mathbf{A}_r^H \overline{\mathbf{R}}_x^{-1} \mathbf{A}_r} \tag{26}$$

The approximate solution and weight vector can be found by adaptive methods such as steepest descent, conjugate direction, gradient, conjugate LMS (Least Mean Squares) and interactive LMS (Deyan et al. 2016).

The performance of the beamforming systems under various optimization criteria is analyzed by means of the Monte-Carlo simulations. The simulations estimate the influence of some parameters on the performance of the system, including SNR and SIR (Signal to Interference Ratio). Array configuration is ULA, number of antennas $M = 64$, difference DOA angle between transmitted signal and interference $\Delta\theta = 10°$.

The system performance is evaluated with the Normalized Root Mean Square Error (NRMSE) and the final value is the average value of all Q values after each simulation:

$$NRMSE = \underset{1 \leq q \leq Q}{mean} \left\{ \frac{\sqrt{\frac{1}{N} \sum_{k=1}^{N} \left| \hat{x}_q(k) - x_q(k) \right|^2}}{\left| \max(x_q(k)) - \min(x_q(k)) \right|} \right\} \tag{27}$$

The simulation results are presented in Fig. 14 according to the SNR ranges. It is possible to see that when SNR assumes a low level, in weak signal range, the beamforming based on min-max optimization yields significant results, and the system is more robust. Moreover, when the signal level is stronger, MVDR provides a better result against interference. For signal estimation when SIR varies, we see that the performance of the system according to MVDR and LCMV methods give best results, meanwhile the min-max optimization yields bad results when SIR changes.

Figure 15 shows the beampattern of the hybrid beamformer with four RF chains and a number of the antennas of 64 (a) and 100 (b). We can see that the optimized beamformer has about four dominant beams. This beampattern means that the data streams can be successfully transmitted through those beams.
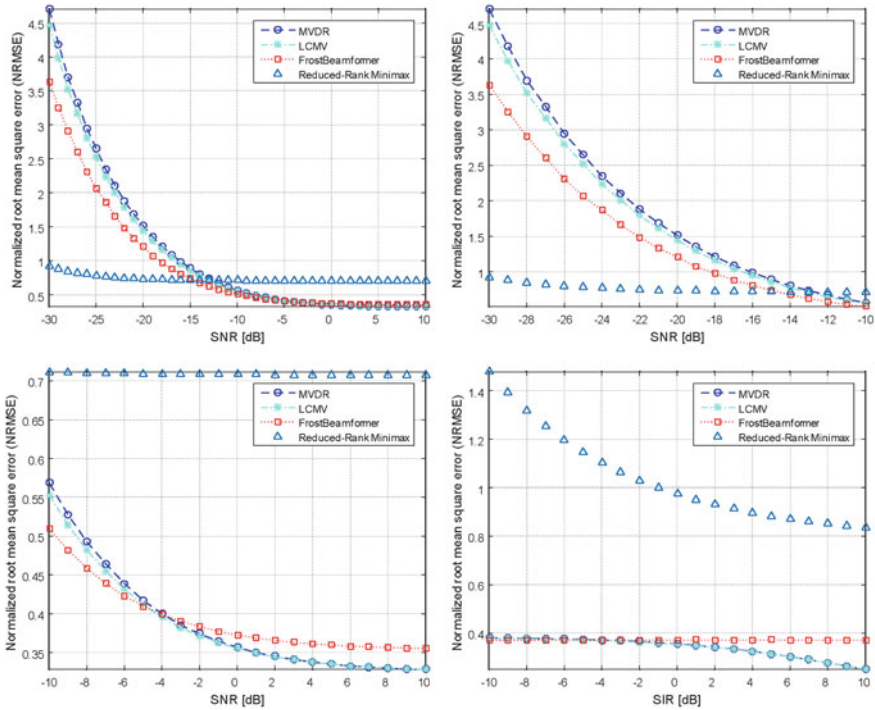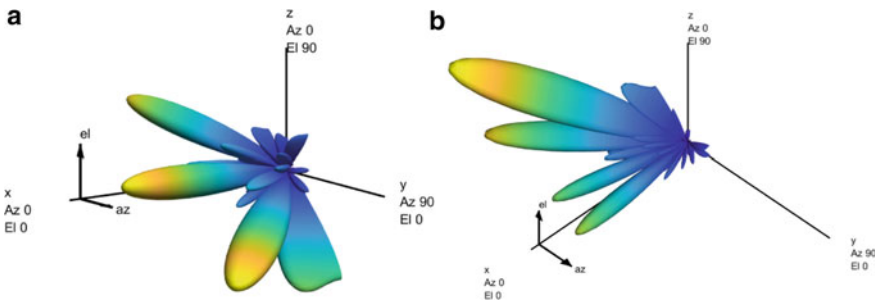
Fig. 14 NRMSE according to SNR and SIR

Fig. 15 Beampattern of the hybrid beamformer with the number of the antenna elements is 64 (**a**), 100 (**b**)

References

Alluhaibi O, Ahmed QZ, Pan C, Zhu H (2016) Capacity maximisation for hybrid digital-to-analog beamforming mm-wave systems. In: IEEE global communications conference (GLOBECOM)

Bhattacharya A, Chattopadhyay PK (2010) Hybrid differential evolution With biogeography-based optimization for solution of economic load dispatch. IEEE Trans Power Syst 25(4):1955–1964

Cameron JJ, Leung CK-S, Tanbeer SK (2011) Finding strong groups of friends among friends in social networks. In: 2011 IEEE ninth international conference on dependable, autonomic and secure computing. IEEE

Dantzig G (2016) Linear programming and extensions. Princeton University Press

Deyan P, Shouqiang D, Jingjie J (2016) The smoothing Fletcher-Reeves conjugate gradient method for solving finite minimax problems. Sci Asia 42:40–45

Eldar YC (2005) Minimax MSE estimation of deterministic parameters with noise covariance uncertainties. IEEE Trans Signal Process 54(1):138–145

Elkashlan M, Duong TQ, Chen H-H (2014) Millimeter-wave communications for 5G: fundamentals: Part I [Guest Editorial]. IEEE Commun Mag 52(9):52–54

Elkashlan M, Duong TQ, Chen H-H (2015) Millimeter-wave communications for 5G–Part 2: applications. IEEE Commun Mag 53(1):166–167

Goldberg DE, Holland JH (1988) Genetic algorithms and machine learning. Mach Learn 3(2):95–99

Grimaccia F et al (2017) Design of tubular permanent magnet generators for vehicle energy harvesting by means of social network optimization. IEEE Trans Ind Electron 65(2):1884–1892

Hefnawi M (2019) Hybrid beamforming for millimeter-wave heterogeneous networks. Electronics 8:133. https://doi.org/10.3390/electronics8020133

Huang J, Encinar JA (2007) Reflectarray antennas. Wiley, New York

Kennedy J (2010) Particle swarm optimization. Encycl Mach Learn 760–766

Kolundzija BM, Olcan DI (2006) Multiminima heuristic methods for antenna optimization. IEEE Trans Antennas Propag 54(5):1405–1415

Larsson EG (2015) Joint beamforming and broadcasting in massive MIMO. In: Proceedings of IEEE workshop on signal processing advances in wireless communications (SPAWC)

Mussetta M, Grimaccia F, Zich RE (2012) Comparison of different optimization techniques in the design of electromagnetic devices. In: 2012 IEEE congress on evolutionary computation. IEEE

Mussetta M, Pirinoli P, Zich RE (2013) Application of modified BBO to microstrip filter optimization. In: Proceedings of 2013 IEEE AP-S/URSI international symposium, pp 410–411

Nayeri P, Elsherbeni AZ, Yang F (2013) Radiation analysis approaches for reflectarray antennas [antenna designer's notebook]. IEEE Antennas Propag Mag 55(1):127–134

Nayeri P, Yang F, Elsherbeni AZ (2018) Reflectarray antennas—theory, designs and applications. Wiley, New York

Niccolai A et al (2015) Sparse array design by means of social network optimization. In: 2015 IEEE international symposium on antennas and propagation & USNC/URSI national radio science meeting. IEEE

Pirinoli P, Massaccesi A, Beccaria M (2017) Application of the M_m C_n-BBO algorithms to the optimization of antenna problems. In: 2017 international conference on electromagnetics in advanced applications (ICEAA). IEEE

Simon D (2008) Biogeography-based optimization. IEEE Trans Evol Comput 12(6):702–713

Simon D (2013) Evolutionary optimization algorithms. Wiley, New York

Sohrabi F, Yu W (2013) Hybrid digital and analog beamforming design for large-scale antenna arrays. IEEE J Sel Top Signal Process 3476–3480

Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359

Tan Y, Zhu Y (2010) Fireworks algorithm for optimization. In: International conference in swarm intelligence. Springer, Berlin

Young HP (2009) Learning by trial and error. Games Econ Behav 65(2):626–643