TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
76/2018

Models for Modern Power Distribution System Planning

SAMBEET MISHRA

TTÜ PRESS

Tallinn University of Technology
School of Engineering
Department of Electrical Power Engineering and Mechatronics
The dissertation was accepted for the defence of the degree of Doctor of Philosophy (PhD) on 20 November 2018

Supervisor: Professor Ivo Palu
School of Engineering
Department of Electrical Power Engineering and Mechatronics
Tallinn University of Technology, Tallinn, Estonia

Co-supervisor: Dr. Chiara Bordin
Department of Energy Systems
SINTEF Energy Research, Trondheim, Norway

Opponents: Professor Maria Teresa Vespucci
Department of Information Technology and Mathematical Methods
University of Bergamo, Bergamo, Italy

Associate Professor Audrius Jonaitis
Department of Electrical Power Systems
Kaunas University of Technology, Kaunas, Lithuania

Defence of the thesis: 20 December 2018, Tallinn

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

Sambeet Mishra ______________________________
signature

European Union
European Social Fund

Investing
in your future

Copyright: Sambeet Mishra, 2018
ISSN 2585-6898 (publication)
ISBN 978-9949-83-362-7 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-363-4 (PDF)

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
76/2018

Kaasaegsete jaotusvõrkude planeerimise mudelid

SAMBEET MISHRA

TTÜ KIRJASTUS

"There is no such thing as the conclusion, so we stop writing . . . "

Contents

List of Publications

The list of author's publications is a selection of papers published by the author during the doctoral studies, on the basis of which the thesis has been prepared:

1. **S. Mishra**, M. Leinakse, and I. Palu "Wind Power Variation Identification using Ramping Behavior Analysis" Elsevier Energy Procedia, pages - 141:565–571, 2017
2. **S. Mishra**, M. Leinakse, I. Palu, and J. Kilter "Ramping Behaviour Analysis of Wind Farms" IEEE conference, EEEIC2018, pages - 1(1):1–7, 2018
3. **S. Mishra**, C. Würsig and I. Palu "Multivariate Scenario Generation - an ARIMA and Copula Approach, International Journal of Modeling and Optimization, 2018
4. **S. Mishra**, C. Bordin, J. Fornes, and I. Palu "Reliability Framework for Power Network Assessment" E3S Web of Conferences , 2018
5. **S. Mishra**, C. Bordin, and I. Palu "RNR: Reliability oriented Network Restructuring" IEEE conference-RTUCON, 2018

Forthcoming publication under peer review

1. **S. Mishra**, C. Bordin, A. Tomasgard and I. Palu. "Coordinated Microgrid Expansion Planning", 2018

The list of author’s publications during the doctoral studies, not included in the thesis

1. Mishra, S.; Bordin, C.; Pisciella, P.; Palu, I., “Micro-grid expansion a cooperative game theory approach, conference presentation”, Computational Management Science 2018
2. Mishra, S.; Palu, I.; Madichetty, S.; Suresh Kumar, L. V., “Modelling of wind energy-based microgrid system implementing MMC”, International journal of energy research, John Wiley & Sons, 2016
3. Mishra, S.; Palu, I.; “A user interface tool for ramping behavior analysis of renewable energy”, IEEE conference-PQ2016, 2016
4. Mishra, S.; Koduvere, H.; Palu, I.; Kuhi-Thalfeldt, R.; “Modelling of Solar-Wind hybrid renewable energy system architectures”, IEEE conference-ENERGYCON, 2016
5. Mishra, S.; Koduvere, H.; Palu, I.; Kuhi-Thalfeldt, R.; Rosin, A.; “Assessing demand side flexibility with renewable energy resources”, IEEE conference-EEEIC, 2016

Author's Contributions in

Manuscripts Published

I In 1, I was the main author, developed the concept of wind ramp events classification and applied to measured data from wind farm.

II In 2, I was the main author, improved the concepts and model. Applied to wind farm data and results are explained.

III In 3, I was the main author, identified the research questions, formulated the model and conducted the tests. The data is procured from industrial partners.

IV In 4, I was the main author, identified the real world power network reliability and adequacy indicators, developed the framework and applied to a IEEE 14 bus test case.

V In 5, I was the main author, developed the concept. Implemented the model adapting the framework in previous work to real world AC distribution network.

Forthcoming publication under peer review

I In 1, I was the main author, prepare the concept, formulated the model and tested with real data in collaboration with industrial partners.

Preface

This thesis is a collection of research works conducted by the author for the Doctoral studies at Tallinn University of Technology (TalTech), supported by the Norwegian University of Science and Technology (NTNU) and SINTEF Energy Research. This work is a collaboration between department of Electrical Power Engineering and Mechatronics at TalTech and the department of Industrial Economics and Technology Management at NTNU. This work is also consequential to *PUT533* - Impact of new types of electricity generating patterns to high voltage equipment and cable insulation, *LEP15070* - Cable networks and their effects on the function of the transmission network, *Flex4RES* - Flexibility of the energy system in the context of the integration of renewable energy sources. In addition, the au-thor spent two years of visiting research period at the Department of Industrial Economics and Technology Management, NTNU. Such exchange was partially funded by the Center for Sustainable Energy Studies *(CenSES)*.

Some of the case studies and data-set are procured through agreements with Industrial partners. The industrial partners associated with this research are: *TrønderEnergi AS*, Trondheim, Norway and BrooklynMG, *LO3 Energy*, NewYork, USA.

Introduction

Around the world, the power generation portfolio has increased the share of renewable energy resources in the total energy mix [1]. The factors behind this growth are increasing demand, increasing fossil fuel prices, and the necessity of reducing greenhouse gas emissions [2]. It is predicted that these reasons will not disappear in the forthcoming years even with the best of intentions to increase energy efficiency and decrease fossil-based energy production. Solar, wind, geothermal, and nuclear energy are among the highly cultivated resources.

In the real world it is both cost-effective and time-saving to construct a model for a complex concept. A model can be deemed as being a near-real-world symbolic representation of the eventual real-world concept. A model can be further classified into an abstract or concrete. The difference is that the former is more generalized and is therefore easier to adapt, while the latter is specific to the peculiarities of the concept. A mathematical optimization model, therefore, is a streamlined algebraic formulation which contains parameters and variables expressing relations and concepts. Optimization models in turn are a type of mathematical model. Note that the complexity of the models grows with the volume of input information. Therefore, depending on the intended purpose behind the models, the priorities are set *a priori* by means of assumptions.

The following sections introduce the power system planning process and the modelling approach. Following that, the main contributions of the thesis are outlined.

Modern Power Distribution System Planning

Electrical power is typically transmitted from the point of generation to the point of consumption. Note that this can involve one of the three processes: a) centralized b) decentralised c) the transactive model of the flow. A power system in which the power flows from a set of large-scale generator units in one location to the source of the demand, is a termed as centralized. Opposed to this is the decentralized mode in which the power flows from multiple large and/or small generation units, usually non-dispatchable generators which are situated in close proximity, to the source of the demand. Transactive energy refers to the flow of energy among and across producers, with this being classed as being a decentralized and distributed mode of supply. Power distribution can be understood as the final stage in the delivery of electric power from the transmission system to individual consumers. The modern power distribution system (PDS) can be referred to the distribution system wherein the power is produced locally and consumer participation. The objective of modern PDS planner is to optimally maintain the energy balance in the most economical, reliable and secure manner. The contours of modern power distribution system planning include intelligent distribution management system, coordinated planning, consideration of uncertainties, decomposition, and advanced optimization methods [3]. Therefore, planning for modern power distribution network is an integrated and coordinated decision-making process.

When considering a power network as shown in fig. 1, it can be observed that it has, as an example, fifty nodes and seventeen load demands. The network has twelve non-dispatchable power generation units (such as wind power) and four dispatchable generation units. It illustrates a section of a power distribution network with various transmission line capacities. PDS planning for the presented power network includes decisions in relation to the following:

- the sizing and siting of generation units and energy storage units
- the capacity expansion of PDS

- scheduling the non-dispatchable production units and maintenance
- flexibility in the PDS through demand side management
- the restructuring of the PDS

Optimal planning for such a network involves the challenge faced by the need for the near-accurate prediction of variables such as load demand and wind power production. Wind, being a form of local energy production, is crucial in order to avoid energy curtailment and to better utilize the power production. For example, an accurate wind power prediction would lead to optimal planning for the expansion of transmission capacity. Predicting natural phenomena is often more challenging than that of the dispatchable generation units. However, identifying significant events, such as events that involve high or low power production levels, leads to the better control of the wind power plant. Subsequently, modelling PDS planning becomes highly complex owing to the variables, uncertainties, scale (granularity) and context all having to be considered. For instance, the variability of wind energy production and the inherent uncertainty are examples of elements in that complexity. In chapter 1 the variability of wind power production is characterized and the developed model is presented. Solving a PDS planning model with hourly granularity becomes more computationally intensive with the volume of data that needs to be processed. When consider the capacity expansion model, the context of a distribution system planner can vary from that of a production system.

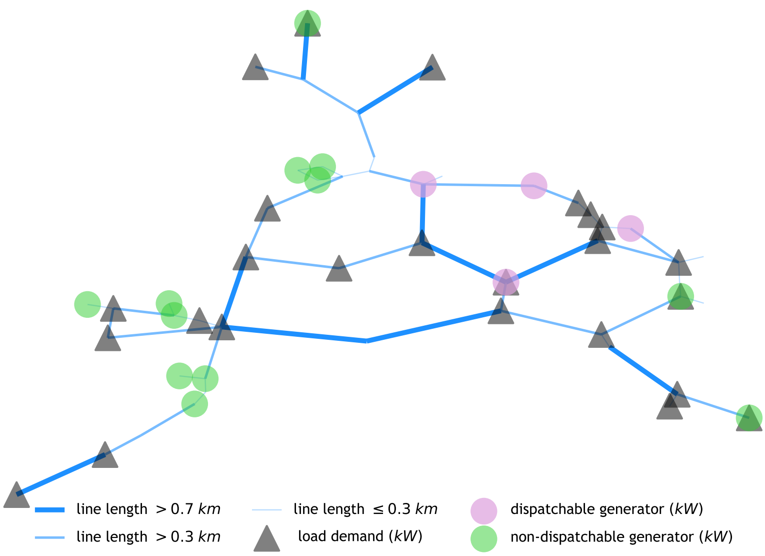


Figure 1 – A cross-section of power network

Note that the line capacities pose the problem of network congestion. The congestion refers the hindering of power supply from the point of generation to the point of load demand along the path of least resistance. However, the congestion as shown in the diagram can happen both in terms of new non-dispatchable energy production units and dispatchable ones. In order to obtain optimal results, such a problem needs to be addressed with high degrees of resolution. In addition, the local energy production and line capacities should also be considered when it comes to devising a both feasible and optimal power system plan.

The Generation Expansion Planning (GEP) model deals with the investments into the expansion and operation of power generation systems. The deteriorating power generation units and growing levels of demand are the two primary motivations for modelling the GEP models with assumptions of the transmission capacity of the existing power network. In particular, with an existing portfolio for the generation units from dispatchable and non-dispatchable resources, the production planner can determine optimal and feasible investments in terms of sizing, operational strategy (such as economic dispatch and unit commitment), siting (involving the optimal selection of locations for the installation of the generation units) and a resource mix of generation units which are able to meet the future demand forecast, along with topological changes in the electrical energy system. The Generation and Transmission Expansion Planning (GTEP) models are combination of generation and transmission expansion planning models. Expansion planning can be a decision for a long or short-term time horizons, depending upon the type of generation units being used. For a dispatchable energy resource, a short-term planning horizon is more relevant, whereas for a non-dispatchable a long term one is more suitable.

Transmission Expansion Planning (TEP) models are a subclass of the power system capacity expansion model with the objective of confirming the power transmission capacity of the network in terms of supplying power to meet the demand at any point in time. The objective is to ensure that demand will be met even in critical circumstances such as: peak demand situations and a failure of any of the generation units. The TEP has two dimensions in the model, namely the economical transaction of energy and the reliability of the network. When considering the mathematical formulation from a central planner's perspective, the GEP is formulated either as a profit maximization model within a market framework in which the trading of energy takes place among agents, or as a cost minimization model. The TEP is formulated either as a centralized or regulated competitive planning model. The objective function for a TEP model considers optimal network expansion in order to achieve economic objectives (such as a reduction in load shedding costs and generation costs) and an increment in system reliability for demand mitigation. A long term planning horizon is adopted for a TEP by considering the growth rate of demand, and the existing generation portfolio.

Any inherent randomness in the natural phenomena and/or incomplete information in terms of the systems give rise to uncertainty. For example, demand is dependent upon weather patterns, as is wind and solar energy production. In the decision theory, GTEP involves variables such as demand, prices and wind which all serve to introduces uncertainty. Computational techniques that can be used to manage uncertainties can be broadly characterized either as sensitive or less sensitive. The latter consists of *a priory* assumption to restrain the variables and the former includes the quantification of possible scenarios. Essentially the model can be classified by basing it on a single scenario and multiple scenarios. With uncertainty comes the dimension of risk in terms of decisions. To elaborate, the optimal expansion plan which derives the highest levels of profit for a given set of scenarios may, be prone to generating higher profit levels but may as well change significantly in another set of scenarios. Conditional value-at-risk (CVar) and chance-constrained techniques are one of the prominent risk management metrics when it comes to hedging the risk that can be associated with uncertainties in the energy domain [4].

The energy market is an environment in which energy transactions can be carried out while ensuring the system's balance and reliability. Energy markets can be regulated or deregulated. The proposed GTEP model in chapter 2 has a nodal pricing scheme for market cleaning, similar to [5]. The Nodal pricing scheme ensures a Nah-equilibrium for energy prices. The regulation ensures market equilibrium where no producer would change their position given the information. Note that incomplete information leads to the possibility of the existence of multiple equilibrium. The emergence of intelligent and responsive tools which are driven by techno-economical advancements with environmentally conscious practices, has highlighted the distributed energy systems as the better solution over the traditional hierarchical power network structures. This work considers a network environment with distributed energy nodes, one which permits interactive energy and information transactions at all levels of generation and consumption. An example of this is demand-side participation which refers to interactions between the price signals and energy consumption. Value added energy transactions among agents within and across power systems via economic and control signals is defined as transactive energy. Notice that the complexity of such a transactive energy-based system grows exponentially due to increase in control points in contrast to traditional ones. This in turn invokes decision-making with the importance of considering detailed system information, granularity, and actors becoming key to the process. In addition, partial information based energy transactions in an energy market in which the energy prices are revealed following the submission of bids with quantity, gives rise to uncertainties. Similarly, the lapse between the time frame involved in physical transactions and that for virtual transactions adds up to uncertainty and complexity.

The context of a decision making model is central to the formulation and interpretation of the model results. For example a market regulatory body with an intention of policy making is more inclined towards avoiding market inflation by ensuring healthy competition and environmental concerns. At the national level in a developed economy (such as in OECD countries) a central regulatory body, typically governments, determines the adequate policies for optimal system welfare. The power infrastructure in a country constitutes transmission system operators (TSO), distribution system operators (DSO), consumers and prosumers. Although there is quite often only one major TSO, there are significantly more numbers of regional DSOs. The emergence of transactive energy platforms and associated new actors (for instance prosumers, plug-in electric vehicles) is bringing with it increased levels of attention to the distribution network. It is clear that the shifts will have to take place in the power distribution network space. For instance one aspect is peer-to-peer energy transactions among and across prosumers. One practical application appeared as the first start-up on block-chain based peer-to-peer energy transaction by the Brooklyn-MG initiative from the LO3 Energy in New York, USA [6]. Depending upon the size and structure of the power network, the central power grid is followed by the micro, nano and pico grids. One of the clear advantages at a central grid level is power system stability. However the past few decades entails many stability improvements in terms of decentralized control architectures that help to tackle the engineering issue. Note that the history registers the economy as being the driver of technological advancements, for example the communication industry.

The fig. 2 depicts a power distribution network. This distribution network is divided into zones of various sizes and capacities of load demands and power productions. Splitting the network into zones has the following effects: a) reduces the total complexity of the overall system, b) facilitates the tracking of minute changes and impacts, c) is computationally less expensive to solve smaller instances of the problem, d) the optimal solution may be a local one rather than a global one, e) it is arguably more practical in terms of the

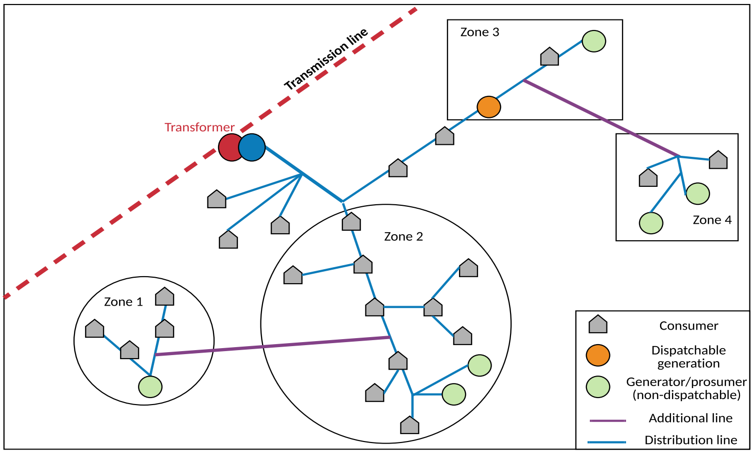


Figure 2 – Power distribution network divided into multiple zones

local network and power system elements and f) makes it possible to consider in more detail the properties of the network (e.g, adequacy and reliability).

In this scenario the problem is a decentralized and distributed one. The granularity of the problem increases however more practical and applicable results. A generation and transmission expansion problem in this case considers multiple zones in the region. The DSO owns the network and facilitates both the energy flow and ancillary services such as demand side participation and prosumers. Consequently an optimal solution that can be obtained from such an approach is feasible on a local level. Considering the context explained, the model proposed in chapter 2.1 focuses on the distributed and decentralized energy infrastructure (DDEI) using transactive energy flow. Chapter 2.2 focuses on employing a decision support system for interactive visualization of the results. In a DDEI power distribution system, networks becomes more relevant with any significant increase in the number of power generation entities, prosumers, community grids. A zone, in this context, consists of modular energy production, dynamic load demand and smart control mechanisms which all serve to form a local power distribution system. Note that a zone can both be grid integrated and isolated. In this context there are many zones and therefore, even individual decisions are more interdependent or in other words there is a tractable and cause-effect mechanism that is present. Be it a zone or microgrid (MG), a community grid or a prosumer each actor has an individual profit maximization objective that may overlap with those of its neighbours.

In this context a *best-of-both-worlds* decision would be to coordinate the decisions so as to maintain the system stability. Note that each MG has controls and information that are limited to its region. Getting a unanimous consensus among the actors in this context is a somewhat perilous undertaking that could lead to ambiguity. This work is an initial step towards understanding and therefore solving this multi-dimensional decision problems. A multi-dimensional model often requires an inter-disciplinary approach to the solution. The coordinated information is limited to information that is of mutual interest for instance, for non-dispatchable generation and any increase in load demand. The internal system information can cover areas such as the remaining utilization lifetime of power apparatus, power outages, and substation-wise power consumption which are not shared or owned by the local distribution network operator. The acquisition of such data has became more accurate owing to significant developments in information and communication techniques, such as smart energy meters. However, it is not within the scope of this research to investigate the associated privacy issues that may be related to the use of such data - that is something for future investigation. Furthering the coordinated decision making in chapter 2, the system reliability and investment within the zone is explored in chapter 3. Power system reliability refers to the operational state of a power system. Power system adequacy is associated with static conditions and the facilities that are needed in order to meet the load demand. Therefore an optimal system plan includes (but not be limited to) the components that include: optimal power flow, remaining utilization life span of power system apparatus and consumer satisfaction. A framework to evaluate the reliability and adequacy and a mathematical optimization model formulation which serves to address optimal investment in power network restructuring for the power distribution network with alternating current optimal power flow (AC-OPF) are both presented in chapters 3.1 and 3.2 respectively.

Methodology Used for Modelling in the Thesis

This section provides a brief introduction to the methods, and mathematical programming that are, used for modelling in the thesis. The objective is to provide a comprehensive background on the following: Mixed Integer Linear Programming (MILP), Non-Linear Programming (NLP) and stochastic programming.

Linear and Mixed Integer Linear Programming Models

LP is arguably the most widely used constrained optimization model. The reason being efficient solution algorithms, applicability and presence of extensive theory [7–9].

A LP model can be expressed as

$$\underset{X}{\textit{minimize}} \quad \sum_{t=1}^{T} \sum_{i}^{M} c_i x_t \tag{1a}$$

subject to:

$$\sum_{t=1}^{T} \sum_{i=1}^{M} a_{it} x_t \geq b_i \qquad \forall\, t \in T, i \in M \tag{1b}$$

$$x_t \geq 0 \qquad \forall\, t \in T \tag{1c}$$

In (1) T, M, N are interpret as scalars within R and c_i, x_t, b_i are vectors within R^m. Moreover a_{it} is a matrix within R^{m,t}. Here x_t is a decision variable. In LP decision variables are *continuous* in nature. The above formulation has a linear objective function as in (1a) and a set of inequality constraints as in (1b) and (1c). Consider that objective function can be changed from *minimize* to *maximize*, in this case the inequalities become $\leq$. Standard algebraic solution techniques are inapplicable in the presence of inequalities and thus methods such as Simplex algorithm is utilized to solve.

Let us consider a change as $x_t \in$ Z $\forall\, t \in$ T; in this case the LP becomes an ILP (Integer Linear Programming) formulation. Thereby a MILP is a combination of LP and ILP wherein a subset of variables are integers. For instance $x_t \in \{0,1\}\, \forall\, t \in$ T restricts x_t to be binary. Note that the complexity of the model is increased with an integer variable as opposed to continuous since the variable can only take an integer type value and not decimal. For example it can not take a value of 1.1. This has a wide range of use in practice in case of energy sector, such as scheduling of dispatchable generation units used in chapter 2. Semi-continuous variables are those that can take the value zero or any value between its lower bound and its upper bound. Note that the semi-continuous upper bound need not be finite but the lower bound need to be finite.

This is used in chapter-2 to restrict the power flows in the power transmission lines within the line capacities. LP models are Polynomial models meaning $\exists$ an ensured polynomial time solution. While MILP are NP-hard models meaning $\nexists$ ensured polynomial time solution. The objective function contours and boundaries of the feasible region are necessarily straight lines that is $y = mx + b$: where y is stating how far up, x how far along, m = slope or gradient implying how steep is line and b is intercept stating where the line intersects the y-axis.

The (1) is also a deterministic model because it satisfies the properties: proportionality, additivity and certainty. Proportionality refers to the contribution of each decision variable both in objective function and constraints. Additivity requires the sum of of the individual contributions of each variable to be same as the contribution of all the variables in the objective function. Certainty means that the formulation has known constants or average value approximations of the probabilistic distributions. Note that it is usual practice to make such assumption under the criterion that the standard deviations of these distributions are sufficiently small. If the standard deviation is large then a sensitivity analysis is performed to record the variations.

Non-Linear Programming

In chapter 1 and 2 the models developed are Non-Linear Programming (NLP) models. This section introduces the fundamental concepts for a NLP.

A NLP model consists of an algebraic objective function and constraints. The algebraic refers to operations of addition, subtraction, division, multiplication and exponentiation, etc are applicable to the variables excluding differentiation and integration. The objective function contours and constraint boundaries are need not be straight lines that make it very difficult to solve. A NLP can be classified as convex or non-convex. A region of space is deemed to be convex if the portion of the straight line between any two points in the region also lies in the region. Convexity can be expressed as

$$f : \psi \rightarrow \mathrm{R} \qquad \forall\, x_1, x_2 \ldots \in \psi \tag{2}$$

$$\forall\, i \in [0,1] : \qquad f(ix_1 + (1-i)x_2 \ldots) \leq if(x_1) + (1-i)f(x_2) \tag{3}$$

Thus any point $x^* \in \psi$, satisfies $f(x^*) \leq f(x)\ \forall\, x \in \psi$. A convex model maintains the properties of a LP. The primary difference between convex and non-convex optimization model is that

- A convex model has a unique solution that is the global optimal or $\nexists$ a feasible solution to the model. In addition, the local solution is both locally and globally optimal.
- A non-convex optimization might have multiple locally optimal solutions and it is time consuming to distinguish if $\exists$ no solution or global one.
- Convex models are easier to solve and very efficient solution algorithms exist to solve it in linear times.
- If a convex problem is solved multiple times to the optimality using different solution algorithms or solvers the solution would always be the same. However for a non-convex problem it is not guaranteed. In fact, it highly dependent on the algorithm and initial guess.

Consider the following formulation:

$$\underset{X}{\text{minimize}} \quad \sum_{t=1}^{T}\sum_{i}^{M} c_i x_t^2 \tag{4a}$$

subject to:

$$\sum_{t=1}^{T}\sum_{i=1}^{M} a_{it} x_t \geq b_i \qquad \forall\, t \in T, i \in M \tag{4b}$$

$$x_t \geq 0 \qquad \forall\, t \in T \tag{4c}$$

The change from x_t to x_t^2 in (4a) turns it a NLP. A NLP is inherently more difficult to obtain optimal solution. Because it is hard to distinguish between global and local optimal, optimal points are not restricted to extreme points, there might exist multiple disconnected feasible regions, different starting points may lead to different final solutions, difficulty to identify feasible starting point, different algorithms can arrive at different solutions for a given model. Vast literature exists on the NLP model and solution strategies as in [8, 10]. The heuristic technique is a process to reach a near optimal, feasible and practical solution. The heuristic begins with an educated guess to reduce the feasible region. The RBA model presented in chapter 1 is a heuristic technique to identify singnificant events in a time-seris data.

Stochastic model

In real world uncertainty is more practical, even in daily life decisions. Stochastic programming (SP) is a mathematical optimization structure that considers the underlying uncertainty in the real world. Uncertainty in this context indicates the absence or incomplete information. In general the uncertainty in future events. The cause of uncertainty can be partial observations, stochastic environment. Stochastic scenarios are forecasts based on past trend with probability of occurrence. In general, they are divided into three categories: most probable, least probable and remaining scenarios. The objective of SP is to find a feasible solution that is optimal to the case (i.e, optimal for the considered scenarios). In many cases the probability distribution of the data is either known or can be expected. This presents an advantage to generate scenarios with certain probabilities of occurrence.

Consider a two-stage SP which is convex in nature as follows:

$$\underset{X}{\text{minimize}} \quad \sum_{t=1}^{T}\sum_{i}^{M} c_i \sum_{s} \Omega_s(x_{ts}) \tag{5}$$

Where Ω_s is the probability distribution for scenario s. Note that SP maximizes the expectation of the objective function and the random variable. The advantage being the recourse action in second stage as a response to compensate any bad decisions taken in the first stage. Further literature available at [11].

In chapter 1 the heuristic technique based RBA model is presented. A novel stochastic multi-variate scenario generation method using ARIMA and copula is presented in chapter 2. The CoMG model discussed in chapter 2 of the thesis is a two-stage stochastic, risk neutral, MILP formulation to address generation and transmission expansion models for power distribution network. In this case the uncertainty in power demand, energy utility price and wind power production are considered as stochastic variables. In the chapter 3

RNR model is a NLP model formulation that takes in to account the Alternating Current-Optimal Power Flow (AC-OPF) model of the power system.

Tools, Software and Solvers for Modelling

Various mathematical optimization models are solved using mathematical algorithms. For example, a *branch-and-bound* technique is a mathematical algorithm that is used to solve the MILP model. The model is solved through iterations and continuous relaxation with the suppression of integer constraints. When it comes to solving the MINLP model in chapter 3, an outer approximation algorithm is utilized. The outer approximation technique utilizes the principles of decomposition, outer-approximation and relaxation. Then the original MINLP model becomes relaxed MILP sub-problems. In this technique a solution is reached through iterations. A detailed explanation of the outer approximation technique is presented in [12–14].

AIMMS (Advanced Interactive Multidimensional Modeling System) is a tool which is used to formulate and solve large scale mathematical optimization models. It provides an easy interpret-able graphical user interface with an integrated development environment for mathematical optimization models. AIMMS provides a combination of declarative and imperative programming styles. Not only it contains a range of pre-installed solvers but also it automatically chooses best solver with additional model specific tactics. It was primarily developed by Johannes J. Bisschop and currently available under licence from Paragon Decision Technology [15]. AIMMS was selected with the objective of being easy-to-use in development, primarily for prototyping both the CoMG and RNR models discussed in chapter 2,3 respectively. GAMS (General Algebraic Modelling System) is a mathematical programming tool. It is used to develop prototype of the RNR model presented in chapter 3.

Guido van Rossum developed the open-source python programming language. It is a high-level programming language for general-purpose programming. It is easy to read with a clear syntax and is object oriented. A range of packages along with a large support base are backing up the project. Pyomo is an open-source python package which was developed by William Hart, Jean-Paul Watson and David Woodruff as an algebraic language for formulating optimization models [15–17]. The Python based Pyomo is used with an objective of enabling an open-resource and open research strategy for modelling the RBA, CoMG, and RNR models.

Contributions by the Thesis

The primary research area of this thesis is optimal planning of modern power distribution network. The planning issues addressed in this investigation concerns both technical and economic aspects for making an optimal decision. An inter-disciplinary approach (mathematical optimization, data science, energy economics, electrical power system, multi agent system) has been taken to formulate the models. This thesis contributes through development of novel models and framework for the optimal planning of modern power distribution systems. Specifically, the optimal network capacity expansion of decentralized and distributed power system. The developed models are tested with real-life data and scenarios for practical decision making. A selection of models are included in the thesis. The chapter-wise contributions are outlined as follows:

Classification and Modelling of Wind Power Variations

- A novel concept to classify the Wind power swings into a series of significant events.

- A ramping behaviour analysis (RBA) algorithm is developed and applied to wind power generation from a wind farm.
- An evolutionary genetic algorithm is used to combine the time-series data. The extracted events are studied to practical implications for wind farm operator are outlined.

Modelling Generation and Transmission Expansion Planning

- A novel stochastic multi-variate scenario generation technique is developed and applied to stochastic variables- wind, demand and price.
- A novel bottom-up two-stage chance constrained stochastic model is developed and presented for addressing GTEP (CoMG).
- A coordinated decision making framework with multi-agent-systems for strategic decision making is developed and applied to distribution network.
- A decision support system is developed and presented that interactively presents the results of the model for industrial applications.

Reliability Oriented Network Restructuring and Expansion Planning

- Power distribution system reliability and adequacy components for a power distribution network are discussed.
- A novel reliability oriented network restructuring (RNR) framework is proposed considering power system adequacy and reliability aspects.
- A non-linear AC-OPF model (RNR) is developed with the objective of enabling power distribution network restructuring taking into account any investment into maintenance costs.

1 Classification and Modelling of Wind Power Variations

"A scientist who learns one [field of] science alone can not be sure of his own science and, for this reason, the scientist has to be versed in many sciences."

Sushruta Samhita [18]

Wind energy is stochastic in nature due to wind flow being the product of multiple natural phenomena. When the intermittent wind power is introduced to the power system, the inherent uncertainty in the resource introduces challenges to maintain the demand-supply balance. Moreover, the dispatchable power generation units or energy storage units which cover this variability, require additional capacity to be able to do so. In addition, the optimal utilization of power produced from a non-disputable resource such as wind is a key challenge. Forecasting is one of the many possible solutions to this problem, as well as an interconnected grid, energy storage technologies, demand-side management such as electric vehicles. Forecasting aims to model the uncertainties inherited by the grid through wind power production and thus is a necessary and cost-effective element for the optimal integration of wind power into energy systems. However, forecasting is never accurate and literature suggests providing bounds for the forecasts or confidence intervals [19–21].

A wind power ramp event can be termed as a sudden change in the output power over a predetermined threshold. Mathematically, the absolute difference between power produced P_t in time t and $(t+\Delta t)$ that is above the set threshold $\overline{P}$ is a ramp event as in (6). However the threshold value is subjective.

$$|P_{(t+\Delta t)} - P_t| > \overline{P} \tag{6}$$

The system operator (SO) has to keep the system balanced meaning that power generation must meet the demand at each point in time. Wind ramp events can be either positive or negative based on the generation swings. If it is positive, then the wind turbine has to shut down to avoid accidents or damage to the system whereas if the swing is negative the SO has to find a replacement to mitigate the demand. From the economic point of view, both energy not used and energy from an alternative resource are crucial.

Wind farm planners predict the wind speed and power production levels by using historical data over time, with an objective to determine potential investment and operations. In long-term forecasting, events become insignificant due to the stretch of time being looked at, while the short-term forecasting of events is usually more accurate. Furthermore, the time interval Δt, is typically ten minutes for ramp events. The $\overline{P}$ is either set to an absolute value for a wind park or a certain percentage of the quantity of power produced depending on the installed capacity. The problem with this practice is that the peak generation capacity varies through seasons, additionally being mitigated by factors such as turbine maintenance or new installations. Although the threshold is subjective to the peculiarities of a wind park, the methods being used to classify ramp events are generic. This thesis places its focus on the procedure being used to detect ramp events. In addition, it demonstrates the application when it comes to real-world data from a wind park.

1.1 Classification of Wind Power Variations

The wind power variations are classified by means of wind ramp events. To capture the variations a mathematical model is developed. The RBA (Ramping Behaviour Analysis) is model containing distinct and inter-related functions. Each function has a distinct objective and the model is sequential and linear time. The sequential means that the functions must be executed in the same order as ramp events that are based on peak power. This algorithm has four components: rise time, fall time, ramp-up rate and ramp-down rate. The graphical representation in fig. 3 describes a ramp event and the associated components. The angle (θ) references the peak point for both ramp-up and ramp-down events. The persistence presents the amount of time over which the peak event persists. The $\Delta w_s^{up}, \Delta w_s^{down}$ refers to the change in amplitude during a significant event.

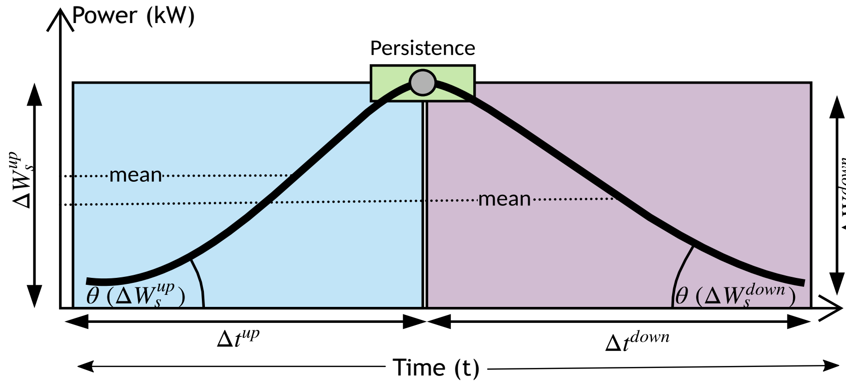


Figure 3 – Wind ramp event classification

The data is procured from the wind park with a time resolution of ten minutes. All of the data points coincide with each other date-wise and time-wise for each turbine. Fig. 4 presents the power production levels from the wind park for the winter season. It can be observed that the data has many power swings while preserving the a time-wise correlation between turbine power generation.

Wind production of a turbine is expressed through capacity factor: the ratio between the net power generation and the calibrated as in $\frac{P_{net}}{P_{rated}}$. The data are registered with a time stamp, however the relation between two subsequent observation is not. Again, the degree of noise content that is inherent in the data is substantial for event detection purpose. For example one outlier could either be a significant event or a noise. Therefore the raw data is smoothed out through various smoothing techniques so that any significant events can be extracted.

Exponential moving average filter technique is used for smoothing the data as in (7). Here f, c, p, w stands for exponential moving average, current value, previous value and $w = 2/(N+1)$ weight factor wherein N is number of periods respectively. The results are presented in the subsequent publication [22].

$$f(c) = \left[\{c - f(p)\}w\right] + f(p) \tag{7}$$

A given function or signal can be transformed from the time to the frequency domain

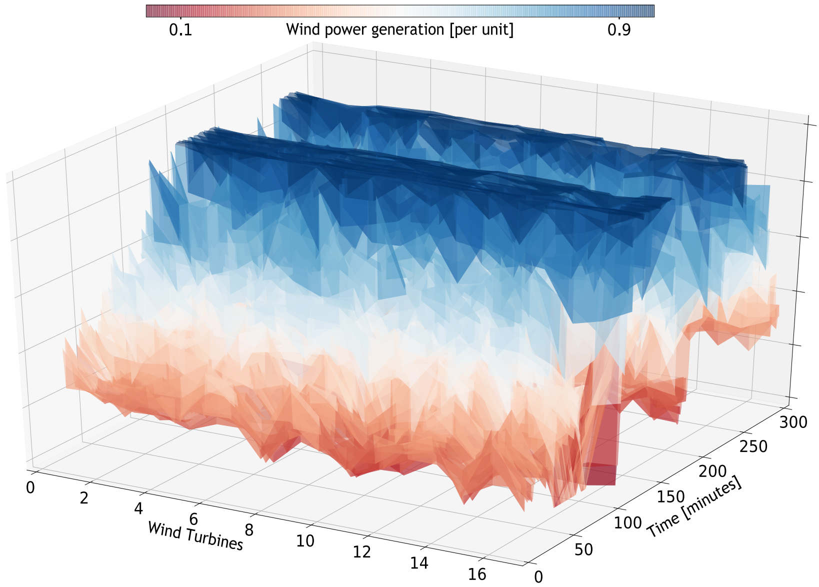


Figure 4 – Wind power generation from wind farm

and the other way round. The frequency domain transforms the linear differential to algebraic equations which are easy to solve. Furthermore, the latter provides the qualitative behavior of the system: such as in terms of bandwidth, frequency response, gain, phase shift, power spectral density, and eigenvalues to name but a few. The focus in this thesis is limited to spectral density and (FFT) Fast Fourier transformation. The Fourier transform of a function contains all of the information about the original signal, and with this information, it is possible to reconstruct the function entirely by an inverse Fourier transformation. This information includes the amplitude and phase of each frequency that is present in the function.

The Fourier transformation of a discrete-time signal $x[n], n = 0, \ldots, N$ is called the discrete-time Fourier transformation (DTFT), which provides a mathematical approximation of the full integral solution, and yields a periodic frequency spectrum. The DTFT of the sequence $x[n]$ denoted in (8a) is a function of a continuous frequency variable ω and $X(e^j\omega)$ and is always periodic with period 2π. And (8b) represents the inverse DTFT of $x[n]$. DFT (Discrete Fourier Transform) can be obtained from the DTFT by evaluating through a discrete set of equally spaced frequencies [23].

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \tag{8a}$$

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})e^{-j\omega n}d\omega \tag{8b}$$

A finite number of samples are selected in order to determine the spectrum. Then a window is generated by a multiplication of $x[n]$ by another sequence $w[n]$. A Blackman window is selected for the study. A time-domain representation of the same is presented in (9) where N is the *length* of the Blackman window.

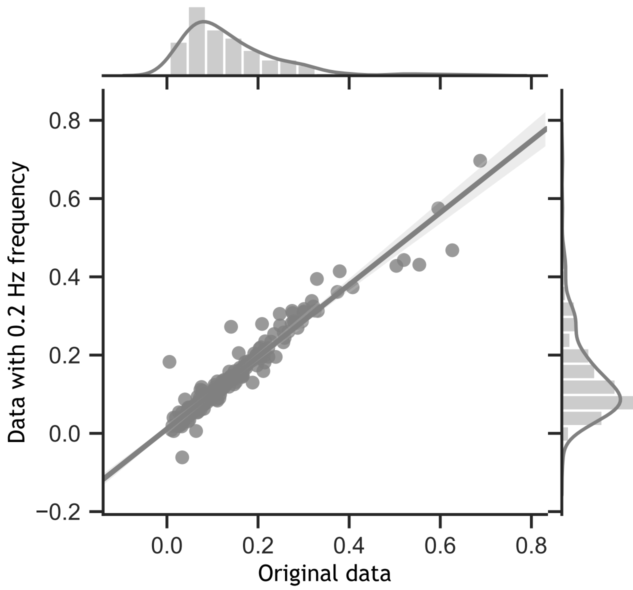


Figure 5 – Joint plot of original data and data with 0.2 Hz cut-off frequency

$$w(n) = 0.42 - 5\cos\left(\frac{2\pi n}{N-1}\right) + 0.88\cos\left(\frac{4\pi n}{N-1}\right) \qquad \forall\, 0 \le n \le (M-1) \tag{9}$$

$$M = \begin{cases} N/2 & \text{if } N \text{ is even} \\ (N+1)/2 & \text{if } N \text{ is odd} \end{cases}$$

The data processing or filtering is implemented to separate the noise content from the data-set. The next section covers the detection of wind ramp events using RBA. The original data in frequency domain is filtered using Blackman filter with 0.2 Hz cut-off frequency. To demonstrate the relationship between the two variables, a linear regression model $y\,x$ is invoked with 95% confidence interval. Fig. 5 presents a joint plot where the variables are drawn through a scatter plot, followed by a linear regression and the corresponding distributions. It is evident from the picture that in frequency domain with 0.2 Hz cut-off frequency the original property of the data is preserved. Moreover, in frequency domain the data is filtered using Blackman window to separate the noise content and smooth the data while reataining the original pattern. More investigation is required to explore various other filtering methods such as low-pass and band-pass.

1.2 Model for Wind Power Ramp Events Detection

RBA (Ramping Behvaiour Analysis) is a model to detect wind ramp events. Consider the wind power generation w_t at discrete time t and the consecutive measurement as w_{t+1} where the balance is $w_{\Delta t}$. A finite variation $\Delta\, w$ therefore denoted as a ramp. Positive value of $\Delta\, w$ becomes a positive ramp and otherwise is denoted as a negative ramp. Thus a ramp event $\Delta\, w_s$ is defined as an event where a significant change in power generation takes place in a time period $\Delta\, t$. The significance is determined through the parameter T

that stands for an adjustable threshold to neglect $\Delta\, w$ values that are smaller than a set threshold T as in $\Delta\, w_s = \Delta\, w \quad |\Delta\, w > T$. A detailed explanation of RBA is presented in the [22, 24]. The mathematical expression for identification of peak points is presented in (10a).

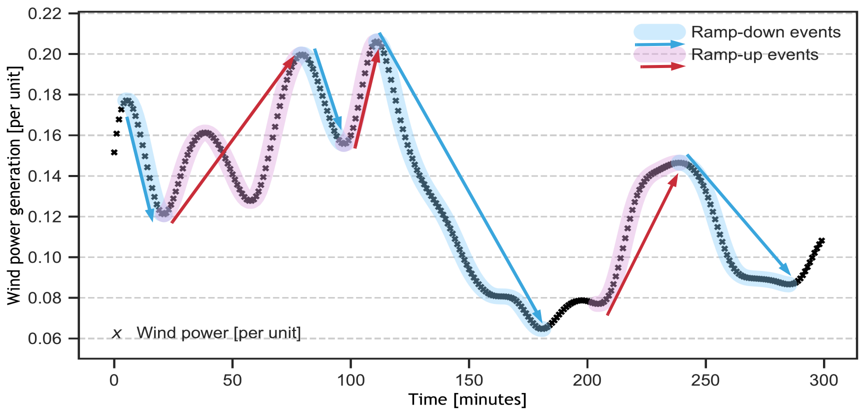


Figure 6 – Wind ramp detection with 10% threshold

The angle between the interval and the change in amplitude is denoted as θ. It can be formulated as in (10). An algorithmic description of the RBA_θ is presented in algorithm 1. The algorithm takes the input of time varying wind data (vector), the total length of the time period (scalar). Then the events are identified with respect to peak points and event characteristics: ramp-up, ramp-down, rise-time, fall-time, angle of the peak with respect to adjacent minima and persistence values are the results. The angle θ between two significant changes $\Delta\, w_s$ and $\Delta\, t$ is presented in (10b).

$$mean\ (\Delta w_s) = \left(\frac{w_{s,t} + w_{s,(t+\Delta\, t)}}{2}\right) \tag{10a}$$

$$\theta(\Delta w_s) = \arctan\,(\Delta w_s, \Delta\, t) \tag{10b}$$

The threshold T was set to 0.01 or 10% of the nominal capacity of the wind turbine. Fig. 6 presents the resulting ramp event detection. The power production are marked with x marks and ramp-up and ramp-down events are highlighted through blue and purple. Arrows present the scheme used to classify the events, for instance a long red arrow describes a long ramp-up event wherein there are two peak points considered as one significant up event. Note that the first ramp up event persisted for longer period in comparison to the second and third. Similarly the third ramp event persisted for longer period than the other. Clearly persistence values explain the length of an event and thereby indicates the severity. The frequency of occurrence is a measure to count these ramp events in the time horizon of the wind power data. Since this method identifies local peaks in the data with respect to the adjacent minimum point, the first ramp-up event ignores the first peak (at 0.16) and considers the peak at 0.20 in y-axis. Consequently the adjacent minimum point 0.13 is ignored and 0.12 is recorded. There is an overlap of the events at the meeting point this is because the event length includes both the starting and end point of the wind power data. In this way, events are aligned end-to-end. With that, RBA_θ identifies and counts the events while recording the event details at all stages.

Algorithm 1: RBA_θ algorithm pseudo code

Result: $w_{s,t}, w_{s,(t+\Delta t)}, (t+\Delta t), \Delta w_s, \theta(\Delta w_s)$, mean($\Delta w_s$)

function RBA_θ ;
Input : $T, w_t \quad \in \mathrm{R}^+$
$\Delta w = w_{t+\Delta t} - w_t$;
for $i \leftarrow 1$ **to** *length*(Δw) **do**
  if sign($\Delta w[i]$) == sign($\Delta w[i+1]$) **then**
    concatenate(Δw)
  end
end
for $i \leftarrow 1$ **to** *length*(Δw) **do**
  if $\Delta w > T$ **then**
    $\Delta w_s \leftarrow \Delta w$
  end
end
for $i \leftarrow$ **to** *length*(Δw_s) **do**
  if sign($\Delta w_s[i]$) == sign ($\Delta w_s[i+1]$) **then**
    concatenate(Δw_s)
  end
end

A genetic algorithm (GA) is a local search optimization technique primarily applied to optimization problems that are highly non-linear, non-differential or discontinuous. Conceptually GA is based on the biological evolution process. The process randomly selects a pair out of the total population and crossover takes place to produce successors for next generation. This process of evolving toward an optimal solution classifies GA as an evolutionary technique [25]. The fitness function ff of the GA is (11). The ff considers the time horizon t, the turbine w, month m and y as the current data point.

$$ff = \sum_{w=1}^{|w|} \sum_{j=1}^{|t|} y_j^w * \left(y_j^w - m_j \right)^2 \tag{11}$$

Through application of GA the total volume of data to be processed is reduced to four folds. The results are elaborated in the publication [24].

The K-means clustering algorithm, also known as Lloyd's algorithm is used for classification of objects into K number of groups based on attributes. In case of wind ramps, the ramping behaviour analysis uses 7 attributes for each event, which can be used for classifying the wind ramps to groups. The algorithm is based on the minimization of squared Euclidean distance between the objects and centers of the assigned clusters [26]:

$$min\ (E) = min \left(\sum_{i=1}^{K} \sum_{x \in C_i} d(x, z_i) \right) \tag{12}$$

where z_i is the center of cluster C_i and $d(x,z_i)$ is the Euclidean squared distance between point x and cluster center z_i. Clustering is applied to group the events with an objective to distinguish the significant attributes in ramp events. Optimal weight is allocated to RBA attributes using optimal cluster formulation. Clustering of identified ramp events based on different combinations of RBA parameters was conducted are the results are described in the publication [24].

Summary

Wind is an intermittent energy resource. The power system planner need to ensure the power balance and system reliability. Better understanding of the wind ramp events increases the preparedness for the power swings. The wind power ramp events are classified and significant events are extracted using developed ramping behaviour analysis (RBA) model. RBA is modelled as a heuristic technique. When it comes to heuristic techniques, an educated guess significantly increases accuracy and solution time. In case of RBA, determining an optimal threshold value is the starting point. The events extracted as further classified using clustering method to understand the significance of events. An evolutionary technique is applied to combine the time-series data. In particular RBA applications to wind power swings is discussed and presented in more details publications [22, 24]. During the investigation, the authors have used the nominal capacity of wind turbine as the threshold value. The setting of threshold value for wind ramp event extraction remains a challenge and more studies are needed in this direction. Prediction of RBA events in place of time-series data is a future scope of the work. A good research direction is to apply RBA in to identify events in current and voltage variations, rate of charging and discharging with respect to heat generation on cables is to be explored. Apart from that RBA can also be applied to other research disciplines such as growth of cancer cells.

2 Modelling Generation and Transmission Expansion Planning

"Power concedes nothing without demand. It never did and it never will."

Frederick Douglass [27]

This chapter provides a broad and concise overview in order to address the main concepts behind coordinated decision making that are thoroughly illustrated in a forthcoming publication. Generation and transmission expansion planning models are inherently intertwined. For example, the sizing of generation units or the siting of new units depends upon the existing power network capacities for energy transmission. As a further example, the renewable resources such as wind are located far from the point of consumption. Moreover the decisions of aforementioned independent planning models are therefore sub-optimal. Thus GTEP (otherwise referred as G&TEP) has attracted more research in the literature. GTEP offers a significant advantage for optimal system control through congestion management addressing the bottle-necks in the transmission infrastructure and energy curtailment from non-dispatchable resources. For instance the unit commitment for hydro power units for prolific usage in terms of wind power generation incurs significant economic and environmental benefits. Typically this utilizes a node-arc formulation with a mass balance constraint which concerns the total power generated being equal to the energy being consumed. There are multiple versions of the model namely: considering the circuit theory (an Alternating Current (AC) version, Direct Current (DC)) version, a higher granularity version which considers the hourly resolution and the details inherent in the network topology, and a low granularity based version which considers a larger network [28–30].

2.1 Model for Multivariate and Stochastic Scenario Generation

In mathematical optimization uncertainty is expressed through scenarios. The auto-regressive integrated moving average (ARIMA) is one established practise when it comes to generate scenarios.The ARIMA Model is a widely used model for modeling in stochastic programming for generating scenarios for uncertain variables like wind, prices [31].

ARIMA(ϕ, φ) a quasi-contemporaneous stochastic process price ($y^a_{s,t}$), demand ($y^b_{s,t}$) and wind ($y^c_{s,t}$) as in 13(a-c). The residuals $\varepsilon^a_{t,s}, \varepsilon^b_{t,s}, \varepsilon^c_{t,s}$ are statistically dependent. Thus the dependency structure of the stochastic processes can be stated as $\varepsilon\{\varepsilon^a_{t,s} \cdot \varepsilon^b_{t-j,s} \cdot \varepsilon^c_{t-j,s}\} \neq 0$. $\varepsilon^a_s, \varepsilon^b_s, \varepsilon^c_s$ are the series of errors simulated to produce residual cross-correlogram of stochastic process. In 13(d) the error correlation between stochastic process a & b, a & c are presented and finally reduced to a product of an orthogonal matrix B and identity matrix $\psi(E[\psi \cdot \psi^T] = I)$. The cross correlation between $\varepsilon^a_{t,s}$ and $\varepsilon^b_{t,s}$ can be represented through variance-covariance matrix G. G is essentially a positive semi-definite and symmetric matrix. This matrix is further decomposed using Cholesky decomposition $(G = LL^T)$ [32–34]. L is the upper triangular matrix that is also the orthogonal matrix ($B = L$).

$$y^a_{s,t} = \sum_{j=1}^{\eta^a} \phi^a_j \cdot y^a_{t-j,s} + \varepsilon^a_{s,t} - \sum_{j=1}^{\tau^a} \varphi^a_j \cdot \varepsilon^a_{t-j,s} \tag{13a}$$

$$y^b_{s,t} = \sum_{j=1}^{\eta^b} \phi^b_j \cdot y^b_{t-j,s} + \varepsilon^b_{s,t} - \sum_{j=1}^{\tau^b} \varphi^b_j \cdot \varepsilon^b_{t-j,s} \tag{13b}$$

$$y^c_{s,t} = \sum_{j=1}^{\eta^c} \phi^c_j \cdot y^c_{t-j,s} + \varepsilon^c_{s,t} - \sum_{j=1}^{\tau^c} \varphi^c_j \cdot \varepsilon^c_{t-j,s} \tag{13c}$$

$$\varepsilon^1_{s,t} = \begin{pmatrix} \varepsilon^a_{s,t} \\ \varepsilon^b_{s,t} \end{pmatrix} \quad \varepsilon^2_{s,t} = \begin{pmatrix} \varepsilon^a_{s,t} \\ \varepsilon^c_{s,t} \end{pmatrix} \implies \varepsilon = \begin{pmatrix} \varepsilon^1_{s,t} \\ \varepsilon^2_{s,t} \end{pmatrix} \implies \varepsilon = B\psi \tag{13d}$$

$$G = cov(\varepsilon, \varepsilon^T) = BB^T \tag{13e}$$

$$G = LL^T = BB^T \tag{13f}$$

The residuals of the ARIMA Model are fitted to a R-vine Copula Model in order to capture time varying dependence of the data.

The general theory for copulas is Skalars Theorem (1959), based on this Theorem, Skalar shows that a every multivariate distribution can be written as a multivariate copula function. Equation (2) shows Skalars Theorem applied to a three dimensional dataset.

Variables with joint density function:

$$f(a,b,c) = f(a) \cdot f(b|a) \cdot f(c|b,a) \cdot ... \cdot f(a|b,c) \tag{14}$$

Following Skalar (1959) this density function is uniquely represented by the following form, if it is continuous.

$$F(a,b,c) = C(F_a(a), F_b(b), F_c(c)) \tag{15}$$

The R-vine (regular vine) model is chosen to model the multivariate dependence in this empirical application. Fitting multivariate data to a copula is a challenging task, since commonly used copula models, like the normal copula, the t copula or the gumbel copula are either symmetric or have only one parameter to estimate the entire copula, which decreases the flexibility of the distribution. Bivariate copulas have a wider variety of choices. Thus R-vine copula models that fit multiple bi-variate copulas to the multivariate dataset. Subsequently able to capture the dependence structure of the multivariate dataset. R-vine's are represented by a hierarchical tree structure, where the first tree is estimated by n-1 bivariate copulae and the second by n-2 conditional on a single variable. For a three dimensional dataset two copulae need to be estimated directly and one conditional copula. In order to estimate the R-vine a sequential search approach, they first estimate the family and parameters of the first tree via the AIC criterion. Then they use this result to estimate the second tree. Additionally they employ a maximum spanning tree algorithm to choose an appropriate edge weight. the proposed multivariate scenario generation technique ARIMA is used for forecasting and copula for adjusting the residuals. The model and results are elaborated in the publication [35].

2.2 Coordinated Decision Making for strategic expansion planning

Microgrids (MG) are small geographical areas with self-sufficient local production to mitigate consumption of energy. MG are becoming smart, smart-Microgrid (SMG), due to the integration of emerging smart information and communication technologies such as smart metering, lighting and thermostats. Meanwhile, the one of the biggest retail utility infrastructure, energy utility, is undergoing a reformation to accommodate alternative energy sources.

The coordinated microgrid (CoMG) is a novel GTEP formulation that is implemented for investments under uncertainty in network capacity expansion decisions. In addition to that, the value of coordination and interactive decision making is reported. The tests are performed on data from islands near Trondheim (Norway) and illustrate that the CoMG is a time efficient, tractable and scalable model for optimal grid expansion planning. CoMG is a math-heuristic based two-stage stochastic mixed integer linear programming (MILP) model. The inter-operation of heuristic algorithms with mathematical programming can be defined as a math-heuristic technique [36, 37]. The heuristic offers an advantage on computational time by means of compromising the exact solution. Contrary to this, the mathematical model offers an exact solution. Combining them results in "best of both worlds". CoMG is a collaborative strategy based math-heuristic model: in a collaborative strategy algorithms exchange information but are not part of each other (i.e, they can be executed independently). The advantage is that both models can be executed sequentially or in parallel At the top layer of CoMG, a novel evolutionary heuristic algorithm called evolutionary vertical sequencing algorithm (EVS) is implemented. Meanwhile, at the bottom layer a two-stage multi-period stochastic investment model is developed.

Transmission and generation expansion problems are one of the fundamental yet interesting research areas [38, 39]. The contemporary power system primarily operates in a top-down hierarchy. This means that the power flows in a unidirectional way from the producer to the consumer. Meanwhile, distributed energy resources (DER) are increasingly gaining importance and changing the paradigm of top-down and unidirectional flow of power. In fact when DER are involved, individual energy communities will feed in as and when they produce creating a bottom-up system. DER and smart grid help balance and enable the existing central grid. Stochastic optimization models for MG expansion through investment decisions appear in literature [28, 40–46]. A comprehensive analysis of multi-MG system is performed and the permeability of the individual MG found to be improved [47, 48]. However, the efficient and reliable co-ordination of the multi-MG is expressed as a concern. In [49] a mesh of networked smart grids are presented to be the future power system. Multi-Agent-System (MAS) is extensively studied as the solution to automatize the grid operation [50, 51]. Furthermore MAS based control schemes for MG are found to be effective [52, 53]. A real time framework for SMG control using MAS is presented in [54]. These studies clearly outlines that the synchronization and coordination among agents are important issues in the shared and connected information power infrastructure.

In the previous paragraph some of the existing GTEP formulations are discussed. Nevertheless, distribution systems are different from transmission systems. For example in one country there is one transmission network operator while there are several distribution system operators. While existing GTEP formulations can be applied to transmission sector, they do not fit the contemporary distribution sector. The investment planning for a distribution system operator very much depends on the surrounding environment and thus MG coordination. Besides that, there is a growing interest in demand side participation (DSM) that increases the number of stakeholders. In contrast a centralized decision making no longer serves such a problem with conflicting interests. The CoMG introduces for the first time the CDM (coordinated decision making) approach in a GTEP context for distribution systems under uncertainties. The proposed CDM is a methodology to take optimal decision considering the surrounding environment. The key contributions of this investigation on expansion planning is the inclusion of a two-stage stochastic math-heuristic model within a CDM methodology. This results in a coordinated microgrid (CoMG) expansion planning formulation. Furthermore, the value of coordination and interactive deci-

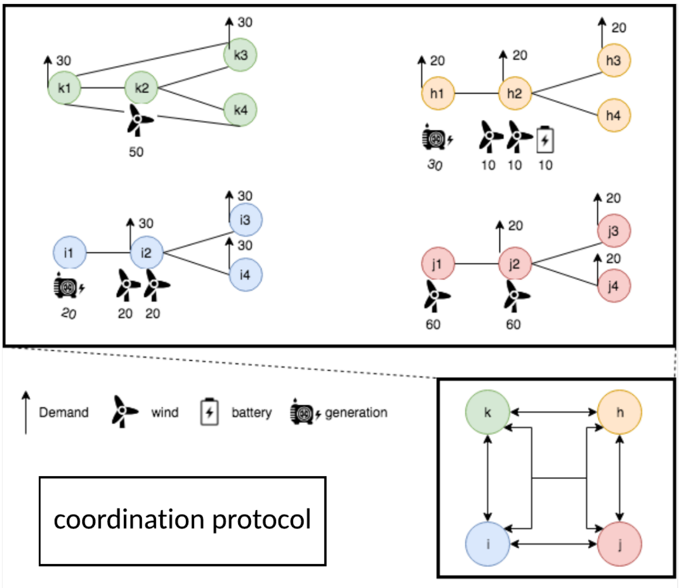


Figure 7 – An instance demonstrating the coordination protocol for four microgrids- k,h,i,j

sion making is reported. CoMG has been tested with real-world data from islands near Trondheim (Norway) and turned out to be a tractable and scalable model for optimal grid expansion planning. To the best of the knowledge of authors, this is the first time that the MG GTEP problem is addressed by integrating a two-stage stochastic MILP model with a heuristic CDM methodology. The inter-operation of heuristic algorithms with mathematical programming can be defined as a math-heuristic technique [36,37]. The heuristic offers an advantage on computational time by means of compromising the exact solution. Contrary to this, the mathematical model offers an exact solution. Combining them results in "best of both worlds". CoMG is a collaborative strategy based math-heuristic model: in a collaborative strategy, algorithms exchange information but are not part of each other (i.e, they can be executed independently). The advantage is that both models can be executed sequentially or in parallel. At the top layer of CoMG, a novel evolutionary heuristic algorithm called evolutionary vertical sequencing protocol is implemented. At the bottom layer a two-stage multi-period stochastic investment model is implemented. In fig. 7 a coordination protocol is presented. The protocol follows sequential enumeration meaning the models are ordered in a sequence and solved. The solutions are then compared and the best solution is determined through comparison. There are four microgrids namely k,i,h,j with different capacities and demand are presented in fig. 7. Individual microgrid is solved to optimality before the final decisions on capacities are passed on to the subsequent microgrid in the order.

The manuscript describing the detailed mathematical model formulation and the results are presented in the manuscript that is currently under processing for publication. In this section the expansion planning model for distribution network is discussed. The following chapter introduces a decision support system (DSS) that facilitates the decision making process.

Decision Support System
The decision support system (DSS) is merged with cloud computing to avoid the barrier of software installation and maintenance inside the companies computers. Moreover, a cloud-based service can perform software updates seamlessly and can deliver a generic product that can be used by everyone independently of the knowledge in data-science. Therefore, the DSS proposed in this work is designed as a Software-as-a-Service (SAAS). The client is the visual part of the application; where the decision-makers interact with uploading the parameters, performing executions and exploring results. The client has been implemented using a trendy technology, Angular 5 [55]. Then, the server or the business logic is implemented using different technologies such as Node.js [56] to interact with the client, Python 3 [57] scripts to automate the integration tasks, the resolution of the model and also the interaction with the mathematical model.

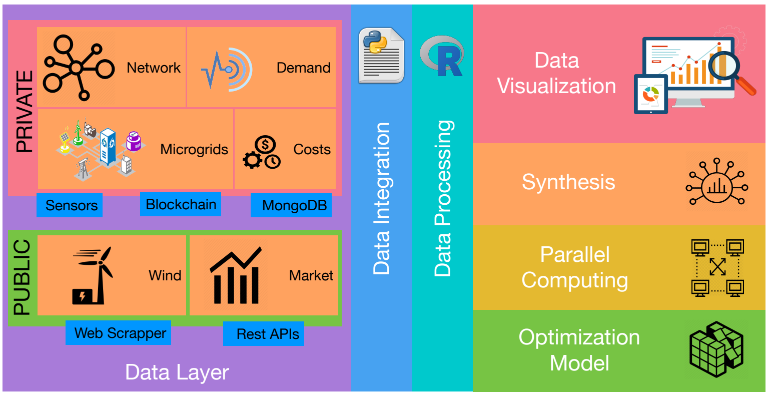


Figure 8 – Decision Support System architecture.

To begin with, the DSS has a data layer to integrate public and private data. These data are merged and integrated to obtain a model very close to the real processes of the company and also very similar to the current context. The data integration is made using python scripts to extract data and adapt to the Pyomo mathematical model. Next, the data processing stage is about data cleaning, filtering and data transformation to make the prescriptions feasible. After that, a parallel framework runs the method and solve the GTEP model. Fig. 8 depicts the main parts of the architecture described. To keep the focus of the project, the application offers the main results obtained and required to support decision making using charts and an interactive network to explore the solution.

This information is of great interest to power industry either power suppliers and customers. The optimal network allows power suppliers to know which are the best clients to satisfy. Otherwise, clients can evaluate if they want to make a connection to a supplier or invest in installing new generators. This information has a clear strategical focus for both companies and clients. Furthermore, knowing the status of the batteries and generators can help distribution operator centre in their daily tasks in the maintenance of the microgrid. Finally, from a tactical point of view, the power supplier can evaluate how many energy the should produce at each period to maintain the service available.

The main features of the decision support system designed can be summarized as follows:

- **Data integration:** The DSS proposed can be fed with public and private data coming from sensors, databases, or indeed current software tool.

- **Optimal prescriptions:** One of the novel features of the DSS presented in this work is the capability of making recommendations. These recommendations are based on the execution of a complex mathematical model that is able to obtain the optimal network taking into account real conditions.
- **Scenario Analysis:** Another important contribution of this work is the capability of simulating and analyzing different scenarios. The distribution system operator can evaluate what is going to happen if new links or nodes are introduced to the system and how the system evolves from this point. Besides, strategical decisions of choosing the right place to build new generators can be evaluated using this function. This feature helps to anticipate and mitigate the undesired effects of some decisions.

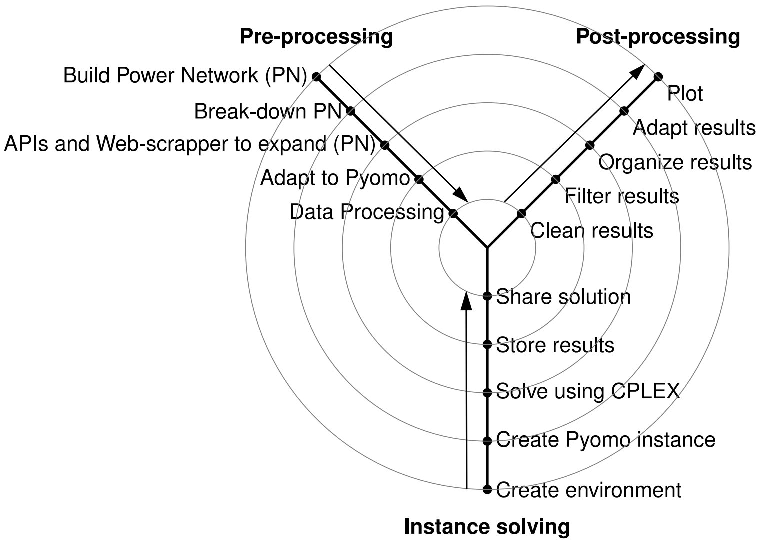


Figure 9 – Decision support system work flow

Fig. 9 illustrates the three stages of the decision making process. The arrow marks shows the sequence of actions in order. The preliminary step is pre-processing where the input data is gathered from different sources and process to feed the mathematical model. First of all, the information is collected to build the power network. Once the power network is built, is time to smartly break-down to multiple sub-networks. At the same time, public real-time data is obtained from APIs and Web Scrappers to model the current context. For instance, weather data or data related to renewable production. After that, all this data is integrated into the mathematical model. In addition, the data is also transformed and manipulated.

Summary

A novel multivariate scenario generation model using ARIMA and copula is developed and applied to three uncertain variables: wind, demand and energy market prices. These scenarios are used for uncertain variables in the stochastic programming expansion planning model discussed in this chapter. The generated scenarios are used as input for the stochastic CoMG model. A novel math-heuristic model with a two-stage stochastic MILP on bottom and heuristic coordination protocol on top, coordinated microgrid (CoMG), is formulated. Evolutionary vertical sequencing protocol is developed and implemented simulating the coordinated decision making process in a multi-agent environment. The coordinated decision making process bears multiple advantages over conventional modelling techniques as it enhances the power security and grid resiliency through:

a. improvement in minimization of cost for by avoiding an upfront investment cost

b. investment and operational synchronization through information exchange

c. reduced power generator and reserve requirements through nested generators

d. reduced power threats due to system status synchronization

e. optimal utilization of renewable power production (i.e, avoid curtailment)

f. A "bottom-up" approach considering higher autonomy of small scale power systems (microgrids) with peer to peer interaction (local market clearing)

The presented CoMG formulation is employed to solve a future power distribution network model in Norwegian islands. The data-set are procured in collaboration with the local distribution company in Trondheim through NTNU. RBA, presented in chapter 1 of the thesis, is used to generate the scenarios for the decision variables. With introduction of the real data, the model grew in exponential scale, thus a high-power cluster computer at NTNU is used to solve the instances and the results are illustrated in the manuscript under review.

A further investigation about the value of coordination would follow based on the co-operative game theory approach. A direction for future investigation could be to record the energy transactions and facilitate smart contracts among agents in a blockchain mechanism. A possible improvement to CoMG can be synchronization of responsive loads and power apparatus within a microgrid. Specifically, the information shared across is also about the responsive load signals. EVS can be improved in a future work by incorporating weather, geographical and network topology factors to be even more accurate in selection of permutations. A further investigation could to be lead to parallelize the CDM to investigate the tractability.

3 Power Distribution System Expansion planning and Network Restructuring

> "The task is, not so much to see what no one has yet seen; but to think what nobody has yet thought, about that which everybody sees."
>
> *Erwin Schrödinger [58]*

Power distribution system operator has to ensure the power flow on demand and maintain stability of power system. To avoid interruptions or power quality issues, the DSO monitors and plans maintenance of the network. The plan includes system reliability and adequacy. Power system reliability refers to the state of network to sustain flow of energy from point of generation to demand at any point in time. Power system adequacy refers to the condition of a power network considering generation, transmission and distribution units. In the modern power distribution network, an optimal investment decision has to include both the reliability and adequacy aspects. An integrated decision making process needs a framework to classify the network into zones based on operating conditions from critical to normal. In this chapter the reliability framework and reliability oriented network restructuring model are presented. The former generates inputs in form of weights for the later model. The developed model utilizes the weights to derive near optimal decisions for PDS expansion with reconfiguration.

The relationship between network reliability and cost is presented in fig. 10. The figure outlines that with increase in number of outages or otherwise power interruptions the reliability reduces. Consequently, the investment increases along with total investment cost. The cross section of investment and outage is the optimal point of operation. Moving left increases the cost, so does moving right. Moving left refers to post-outage scenario and moving right refers to pre-outage scenario where a significant investment is needed to avoid an outage. Considering the node-arc representation for the electrical power distribution network where nodes represent substations and arcs represent transmission lines the framework is formulated. The proposed reliability framework has five categories of reliability parameters: life cycle of power apparatus, environmental and sociological, node reliability, arc reliability and node reliability concerning losses. The table 1 presents the reliability parameters and the individual components or indicators that indicate the condition of a given network. The reliability parameters directly or indirectly influence the network performance. The network performance is directly related to the cost of maintenance and new investments. For this investigation, the maintenance cost values are derived from distribution company. Note that these values are specific to a power network and therefore sensitive in nature. All the maintenance cost values presented in this work are normalized. The investment costs for new installations are subject to the market price during the analysis.

Life cycle assessment (LCA) is a method which helps to determine the environmental impacts that can be suffered due to the use of a product, a process, or an activity. It is also used to assess remaining utilization life. Throughout the product's lifetime any impact should mainly originate from power losses suffered during the usage phase, although installation, maintenance, and dismantling also contribute to it. Transmission and distribution assets have comprised power lines, cables, transformers, substations, and other electrical components in order to generate a wide range of environmental impacts, such as equipment emissions and material weight value. The life cycle stages being viewed are

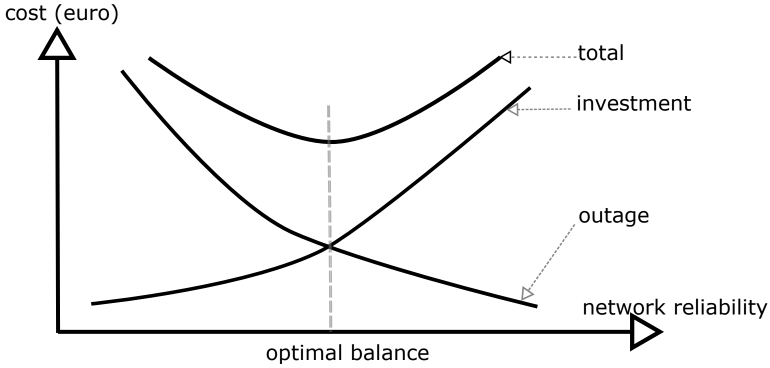


Figure 10 – Relationship between power network reliability and cost

Table 1 – RNR Parameters for power distribution network components

Parameters / Components	LCA	Environmental & Sociological	Node reliability	Arc reliability	Node reliability concerning losses
1	Cable	Terrain	SAIFI	AIT	LOLE
2	OHL	Weather	SAIDI	AIF	LOEE
3	Transformer	Consumption	ENS	AID	EIR
4	Switchgear	Generation	Availability	–	–
5	Emissions	–	–	–	–

interpreted as the production or manufacturing phase of a product or its usage phase. The usage life cycle inventory consists of material requirement for grid components and their environmental impact. For all of the components, the functional unit is one piece of equipment which is operational during the lifetime of that piece of equipment.

The supplied reliability indices combine factors that are related to outage duration or response time, the frequency of any outage, the number of customers who are involved in an interruption or the power and energy that has been lost. The system average interruption frequency index (SAIFI), system average interruption duration index (SAIDI), energy not supplied (ENS), average service availability index (in terms of availability), average interruption time (AIT), average interruption frequency (AIF), and average interruption duration (AID) are among the various measures which hep in evaluating interruptions and their potential impacts.

The formulation for any probability of failure Ω_n for bus n can be presented as in (16). In (17) frequency of failure Ω_n^f is presented.

$$\Omega_n = \sum_i [P(O_j)(P_{i,g}P_{l,i} - P_{g,i}P_{l,i})] \tag{16}$$

$$\Omega_n^f = \sum_i [O_j^f(P_{i,g}P_{l,i} - P_{g,i}P_{l,i})] \tag{17}$$

Where O_j is the condition of outage in the power transmission network. $P_{i,g}$ is the probability of occurrence of capacity outage beyond reserves. And probability of uninterrupted power supply. The availability (γ) is calculated as $\gamma = 1 - \frac{60*ENS}{\sum_i P_i}$ where P_i is average power supplied by the total system and ENS (Energy not supplied because of interruption) and P_i stands for power interruption for incident i. The total cost is a product of component capital cost times availability. The repair cost is calculated as a product of repair time and the total cost. Similarly the maintenance cost is cost of the fault times the repair costs.

3.1 Framework for Determining Power System Reliability and Adequacy

In this subsection, the power system reliability parameters and corresponding indicators are listed and subsequently elaborated. These indicators are used to evaluate the condition of power distribution network. The distribution network operator monitors the condition and plans accordingly to ensure an uninterrupted power supply and maintain power quality. The consumer satisfaction indicators are as follows: SAIDI, SAIFI, ENS and availability. SAIDI depicts how often an average customer experiences a sustained interruption over defined period of time (year). SAIFI Shows the total duration of the interruption for the average customer during a predefined period of time. (per minute;hour). ENS explains the total amount of energy that would have been supplied to the interrupted customers if there would not have been any interruption. Availability refers to the time period that a customer has received power during the defined reporting period (duration of the interruption) AIT address the total time the supply is interrupted AIF explain the number of times when the supply is interrupted annually. AID indicates the duration of an interruption. LOLE is defined as the expected number of days in the specified period in which the daily peak load will exceed the available capacity. LOEE represents the ratio between the probable load energy curtailed due to deficiencies in available generating capacity and the total load energy required to serve the system demand. EIR represents the time period that a customer has not received the energy load during a defined reporting period (expected loss of energy). Terrain-type influence is assumed to have an impact on the probability of a fault occurring. For this, nine terrain types are listed to give a weight evaluation based on the land-use. The impact of weather to the probability of fault is based on the annual normal and adverse weather conditions. Furthermore, probability of fault due to weather effects are assumed to be based on the component failure rate, repair rate and the forced outage rate due to weather.

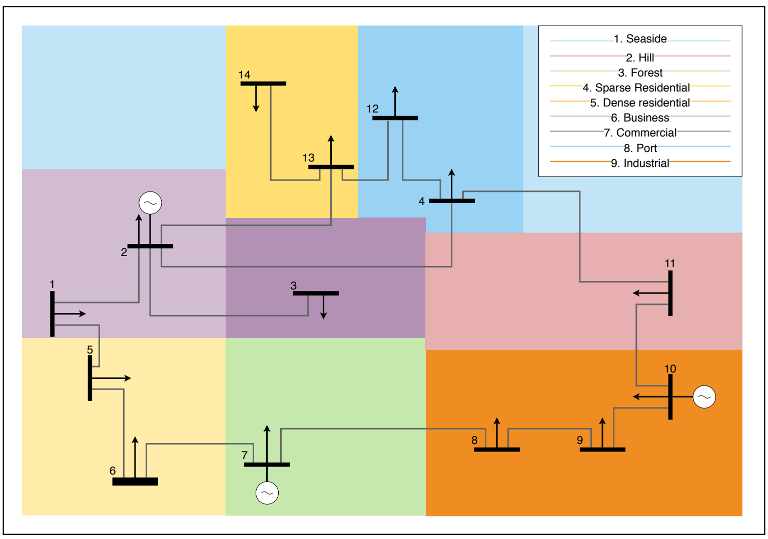


Figure 11 – Proportional evaluation of zones

A coefficient or weight is assigned to each transmission line considering the reliability factors: RUL, consumer satisfaction factors and terrain. In fig. 11 a bus-line diagram for IEEE 14-bus system with zones for weight evaluation are demonstrated. These zones or districts in a microgrid are evaluated based on the buses and transmission line types, for instance residential, forest, dense residential. The zones account for failure rates, frequency of faults and duration of faults. These factors are summarized as maintenance and repair costs for the zones or districts in the existing network. If a new transmission line is to be installed between two zones an average value for costs is considered. Thus the total cost for capacity expansion of a transmission line becomes C^{tot} = (capital cost of equipment * availability* reliability factor) + (maintenance cost*availability*LCA factor * reliability factor). Reliability factor is derived from the LCA, remaining utilization time(efficiency of the equipment), terrain, consumer type, weather pattern, repair time. The costs are assumed to be increasing linearly with respect to time. The ELNS (expected load not served) measures the average amount of energy that is not supplied to loads as a result of load-shedding events. As its name indicates, the expected load not served is a weighted average energy value which accounts both for the probability of contingencies and the damage that these contingencies cause to the system in terms of lost load. The LOLP (loss-of-load probability) is computed as the probability that failure events will lead to load-shedding. In opposition to the ELNS, however, is the fact that the loss-of-load probability is a dimensionless number that does not provide any information on the severity of the disturbance - on the energy not supplied. This lack of a clear physical meaning makes the LOLP a less intuitive metric to work with on the part of system operators. The LOLE (loss-of-load expectation) assesses the expected number of hours during which loss-of-load events could happen. As with the LOLP, the loss-of-load expectation fails to provide an estimation of the damage done to the system by contingencies. From a mathematical viewpoint, both the LOLE and the LOLP require the use of binary variables to be considered within a mixed-integer linear programming problem.

This subsection has presented the RNR framework and various indicators for evaluating the condition of the power distribution network. The evaluated weights segregates the network into zones which are based on the parameter values as in fig. 11. The framework is further explained by a case study on generation outage in the paper listed in appendix 4. Subsequent chapter presents non-linear AC-OPF model with the objective of expanding the power distribution network with reconfiguration.

3.2 Model for Power System Expansion with Reconfiguration

The traditional power network expansion problem is centralized, and it often pays no consideration to network reconfiguration. In other words, any investment decisions often come without considering network restructuring. However, distribution system expansion or, in this case, MG expansion, is an optimum investment which considers both internal and external expansion. The work being presented here considers the reliability aspects of a power network with some network planning insight being added. Reliability-orientated network restructuring, abbreviated as RNR, considers the remaining utilization life of power apparatus, the consumer satisfaction index, and environmental factors. In this study the distribution network has been divided into five main components: aerial lines, cables, transformers, and switches. For each component the main reason behind permanent faults and auto re-closings are determined. Separate failure rates for each component type are based on the reasons for any such failure, such as the transformer overall failure rate being dependent upon lightning, animal interference, or other fault causes. For all of those reasons, the main stress factors which affect the failure rate have been determined. All of the stress factors are classified into appropriate classes such as, for instance the location being a forest location which is near a road or a field. For all classes a weight has been deemed, one which represents the effect of a certain class on the failure rate. For the total failure rate, permanent and temporary faults can be calculated. A practical approach in component modelling is to use the idea that it should be possible to affect the parameters being used in failure rate modelling, with selected planning strategies being added to the process. The weather pattern is not considered directly in failure rate evaluation, but is included in the apparatus condition such as, for instance, in terms of stress tolerance. The age factor is included in condition weight information. Voltage dip analysis is also used for examining short interruptions, where each component is classified based on permanent or temporary short circuit failures. Dip rates are used to define the number and depth of dips in the network. A voltage dip can be analyzed by adding information regarding the total short circuit ratio to every separate failure rate. Failure rate parameters must be determined before modelling methods can be used.

The general failure rate for components was calculated as a weighted mean from the failure rates that had been logged by individual companies. The deemed parameter groups are used to calculate the separate failure rates. The basic input data set is the component information, specifically the type, failure rate, and network topology, as well as some other areas of information which are needed and which can affect the results of the analysis, such as repair times and automation equipment that has been installed. In the enhanced radial reliability analysis, the network is analyzed in terms of feeders and zones, where the use of 'zone' refers to part of a feeder. In the given analysis, the expected number of permanent and temporary failures and voltage dips within a zone are calculated as a sum of the individual network component failures. A determination of repair time is made by analyzing the options when it comes to isolating load points from the faulted component and then restoring the load points with dis-connectors. For a temporary fault, the whole feeder will be experiencing the same short interruption. In the given analysis, permanent and temporary faults that are experienced, along with voltage dips, are calculated for each load point. Cost-related information is based on total interruption times in certain area, permanent and temporary faults, and voltage dip occurrences which are deemed to have taken place when using the radial network reliability analysis [59]. Utility outage cost is based on the value of non-distributed energy and fault repair costs. Other costs, such as losses in production, are considered from the point of view of defining inconvenience costs for the customer. The expected annual costs for permanent outages are the result of a fault occurring in the zone that is being studied. Therefore the RNR framework can be expressed as an asset management model when considering the LCA for power system equipment. When combined with OPF, this forms a complete one-stop solution network management and planning platform. The reliability of reconfiguration when replacing overhead lines and underground cables is evaluated by considering environmental, consumer preference, n-1 contingency, and DSO objectives while minimizing the investment cost.

The optimization models in power system can be broadly classified into operational and planning. The optimal power flow is an example for operation problem, while capacity expansion planning is a planning problem. The distribution system operator need to ensure uninterrupted supply of power on demand while maintaining the quality of power. Therefore, determining optimal capacity expansion considering network restructuring is problem that corresponds to both the operation and planning. Power flow problems can be either DC or AC. The optimal power flow models has an objective of minimizing the operating costs while maintaining the operational variables voltage level and power generation. The DC-OPF models are usually linear or MILP type, therefore can be solved using classic solution strategies. However, the AC-OPF models formulated as MINLP are non-convex and therefore heuristic techniques are used to solve these models. The distribution network has a AC power flow. This leads to the modelling choice of a non-linear AC-OPF model over a DC-OPF.

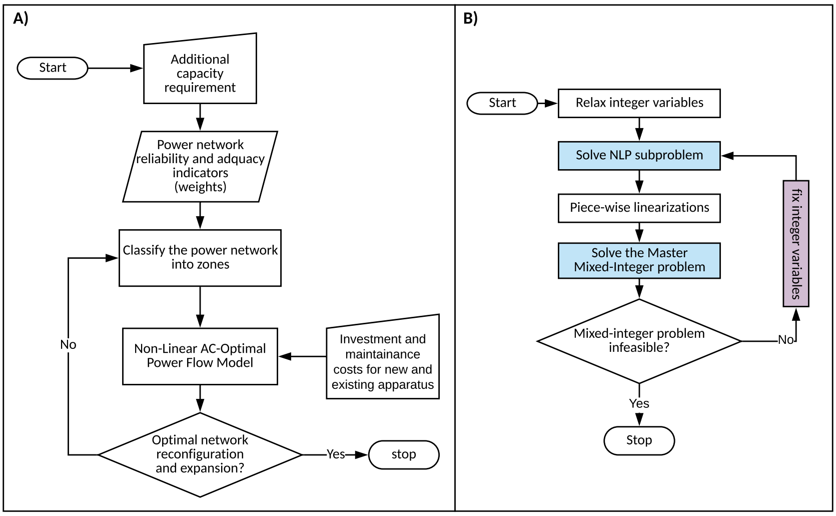


Figure 12 – Flow chart demonstrating the process for (a) expansion planning considering reconfiguration (b) Outer Approximation algorithm

The RNR model presented in appendix 5 is a MINLP formulation and an outer approximation technique is used to solve the model. The work flow of the model is presented through a flowchart in fig. 12(A). The RNR formulation considers both active and reactive power flow and associated losses to ensure the power quality in the distribution network. The MILP formulation is solved using Outer Approximation algorithm [12]. The steps taken by the algorithm to solve the problem is presented in form of a flow-chart in fig. 12(B). The outer approximation is a iterative method. The iteration limit is set as a termination criteria in instances where the problem takes very long time to solve. The restructuring is described by four terms: the cost of installation of new potential cables where a connection still do not exist, the cost of replacing existing obsolete cables with new ones, a representative cost of keeping existing cables as they are and the cost of installing SVC devices in certain nodes. The cost of existing cables is a representative cost that incorporates all the costs that a company should face to keep a cable as it is: this cost is calculated according to the history of the cable, its maintenance requirements, failures and issues and represented by the parameter maintenance cost. The detailed mathematical model is presented in the manuscript at appendix 5.

Summary

In this chapter, A methodology has been presented for analyzing how the process of connecting emerging districts to existing microgrids can affect the reliability of the whole system. The technical aspects of AC-OPF have been thoroughly taken into account and the reliability-orientated Network Restructuring RNR framework has been developed and implemented. The results showed that reliability aspects are crucial when evaluating new investments in grid expansion: new connections should always be coupled with a more holistic evaluation of the conditions of the existing networks as they may require further investment in terms of upgrades so that they can properly fulfill their new requirements. When the system operator is considering the required investment levels for a power network expansion, it should also consider the restructuring of the existing network at the same time. The RNR model being presented is able to address both decisions holistically and, therefore, more investigation is required in this area. In summary the results are elaborated in the publication. Moreover, the RNR expands the dimension of decision making for an power distribution network expansion model.

Conclusions and Outlook for Further Exploration

This thesis is a collection of research works carried out in developing state-of-the-art models for improving modern power distribution system planning during the doctoral studies. This work contributes to two areas- modern power distribution system planning, mathematical optimization modelling. The motivation behind the formulation of research objectives and background for model building in mathematical modelling is presented in the introductory chapter.

In introduction the modern power distribution system and associated planning problems are presented. A brief overview of power system planning problems and an intuitive illustration is provided to better understand the problems. Following that optimal investment planning for capacity expansion in modern power distribution network is elaborated. An introduction to mathematical model building is presented providing an overview on different types of models. Next to that the tools used to develop the models are presented.

In first chapter a novel model for classification of power variations into significant events is presented. The ramping behaviour analysis model is discussed in detail and the results are elucidated. The model is tested with real world data from the wind farm. A further detailed investigation of ramp events and implications are outlined in the associated publications [22, 24]. Identifying ramp events is an important issue specifically for sudden power swings. A better understanding of these events would be an advantage in planning capacity expansion. Further investigation is needed to determine an optimal threshold value for ramp events identification. In a further investigation on prediction of ramp events would be studied.

In second chapter a novel two-stage stochastic, chance constrained math-heuristic model for generation and transmission expansion planning in power distribution network with multiple network operators is presented. The strategic investment planning for expansion in modern power distribution network with multiple system operators can be solved with the coordinated decision making protocol. Coordination among different entities is essential to render an optimal decision. In this chapter a brief discussion about coordinated decision making protocol with the model is discussed. For illustration an instance demonstrating the coordination among four microgrids is presented. The detailed model and the results are in the related manuscript is under review. The scenarios for the stochastic model are generated through a novel stochastic multi-variate ARIMA and copula model. The scenario generation and the model is presented in details in the publication [35]. The correlation between the stochastic parameters- wind, demand and price is captured. In real world these parameters influence each other, thus the correlation among them is essential for optimal decision making. Following that, the decision support system that translates the model results into an interactive visual platform is discussed. Further investigation is required on the classification of microgrids from the power network.

In third chapter a framework for monitoring the condition of power network based on five primary factors: life cycle, environmental and sociological, node reliability, arc reliability, node reliability with losses is presented. The framework is applied to modified IEEE 14 bus network. The resulting weights leads to input parameter for the reliability oriented restructuring model. This is a Non-Linear AC-optimal power flow formulation for determining optimal expansion with reconfiguration. The reliability and adequacy framework is presented in details in the publication [60]. The reliability oriented network restructuring (RNR) model formulation is presented with the analysis on modified IEEE 14 bus system is illustrated in publication [61]. For a distribution network planner determining optimal investment strategy considering the condition of the existing network is an important problem. Expansion in one part of the network may have a significant effect in rest of the network. Such a scenario can arise when upgrading the capacity of one transmission line would create a bottleneck if the adjacent capacities are not upgraded accordingly. Thus expansion with reconfiguration is practical and optimal decision making process. The model presented addresses this issue and further research is needed to test the framework and the model in different network conditions to obtain a balance between reliable and adequate power supply in the modern power distribution network. In addition whether an expansion is still optimal and practical considering both the current condition and required additional capacity.

Abbreviations

RES	Renewable Energy Resources
PS	Power system
PDS	Power Distribution System
TSO	Transmission System Operator
DSO	Distribution system operator
PDN	Power Distribution Network
GEP	Generation Expansion Planning
TEP	Transmission Expansion Planning
GTEP	Generation and Transmission Expansion Planning
OR	Operations Research
OPF	Optimal Power Flow
AC	Alternating Current
MG	Microgrid(s)
LP	Linear Programming
MILP	Mixed Integer Linear Programming
NLP	Non-Linear Programming
MINLP	Mixed Integer Non-Linear Programming
SP	Stochastic Programming
FFT	Fast Fourier Transformation
GA	Genetic Algorithm
RBA	Ramping Behaviour Analysis
ARIMA	Auto-Regressive Integrated Moving Average
CDM	Coordinated Decision Making
CoMG	Coordinated Microgrid
EVS	Evolutionary Vertical Sequencing
DSS	Decision Support System
RNR	Reliability Oriented Network Restructuring

References

[1] Hans-Wilhelm Schiffer, Tom Kober, and Evangelos Panos. World energy council's global energy scenarios to 2060. Zeitschrift für Energiewirtschaft, 42(2):91–102, 2018.

[2] Birol Fatih. World energy outlook 2018. Technical report, International Energy Agency (iea), 2018.

[3] Pavlos S Georgilakis and Nikos D Hatziargyriou. A review of power distribution planning in the modern power systems era: Models, methods and future research. Electric Power Systems Research, 121:89–100, 2015.

[4] Peter Burgherr and Stefan Hirschberg. Comparative risk assessment of severe accidents in the energy sector. Energy Policy, 74:S45–S56, 2014.

[5] Klaus Skytte. The regulating power market on the nordic power exchange nord pool: an econometric analysis. Energy Economics, 21(4):295–308, 1999.

[6] Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. Designing microgrid energy markets: A case study: The brooklyn microgrid. Applied Energy, 210:870–880, 2018.

[7] Wayne L Winston, Munirpallam Venkataramanan, and Jeffrey B Goldberg. Introduction to mathematical programming, volume 1. Thomson/Brooks/Cole Duxbury; Pacific Grove, CA, 2003.

[8] H Paul Williams. Model building in mathematical programming. John Wiley and Sons, 2013.

[9] A Taha Hamdy and A Taha Hamdy. Operations research: an introduction. Printice-Hall of India Private Limited, 2003.

[10] Philip Wolfe. A duality theorem for non-linear programming. Quarterly of applied mathematics, 19(3):239–244, 1961.

[11] John R Birge and Francois Louveaux. Introduction to stochastic programming. Springer Science & Business Media, 2011.

[12] Marco A Duran and Ignacio E Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Mathematical programming, 36(3):307–339, 1986.

[13] Marco A Duran and Ignacio E Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Mathematical Programming, 39(3):337–337, 1987.

[14] Stefan Vigerske and Ambros Gleixner. Scip: Global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. Optimization Methods and Software, 33(3):563–593, 2018.

[15] Johannes Bisschop. AIMMS optimization modeling. Lulu. com, 2006.

[16] William E Hart, Carl Laird, Jean-Paul Watson, and David L Woodruff. Pyomo-optimization modeling in python, volume 67. Springer, 2012.

[17] William E Hart, Jean-Paul Watson, and David L Woodruff. Pyomo: modeling and solving mathematical programs in python. Mathematical Programming Computation, 3(3):219–260, 2011.

[18] Priyadaranjan Ray, Hirendra Nath Gupta, and Mira Roy. Susruta Samhita:(a Scientific Synopsis). Indian National Science Academy, 1980.

[19] Mingjian Cui, Deping Ke, Yuanzhang Sun, Di Gan, Jie Zhang, and Bri-Mathias Hodge. Wind power ramp event forecasting using a stochastic scenario generation method. IEEE Transactions on sustainable energy, 6(2):422–433, 2015.

[20] Xi-Yuan Ma, Yuan-Zhang Sun, and Hua-Liang Fang. Scenario generation of wind power based on statistical uncertainty and variability. IEEE Transactions on Sustainable Energy, 4(4):894–904, 2013.

[21] J Wang, A Botterud, R Bessa, H Keko, L Carvalho, D Issicaba, J Sumaili, and V Miranda. Wind power forecasting uncertainty and unit commitment. Applied Energy, 88(11):4014–4023, 2011.

[22] S. Mishra, M. Leinakse, and I. Palu. Wind power variation identification using ramping behavior analysis. Energy Procedia, 141:565–571, 2017.

[23] James H McClellan, Ronald W Schafer, and Mark A Yoder. Signal processing first. Pearson education Upper Saddle River, NJ, 2003.

[24] S. Mishra, M. Leinakse, I. Palu, and J. Kilter. Ramping behaviour analysis of wind farms. EEEIC2018, 1(1):1–7, 2018.

[25] David E Goldberg and John H Holland. Genetic algorithms and machine learning. Machine learning, 3(2):95–99, 1988.

[26] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1):100–108, 1979.

[27] Frederick Douglass and Rayford Whittingham Logan. The life and times of Frederick Douglass. Courier Corporation, 2003.

[28] Antonio J Conejo, LB Morales, S Jalal Kazempour, and Afzal S Siddiqui. Investment in electricity generation and transmission. Decision Making Under Uncertainty. Springer, New York, 2016.

[29] C Munoz, E Sauma, J Contreras, J Aguado, and S de La Torre. Impact of high wind power penetration on transmission network expansion planning. IET Generation, Transmission & Distribution, 6(12):1281–1291, 2012.

[30] Tohid Akbari and Mohammad Tavakoli Bina. A linearized formulation of ac multi-year transmission expansion planning: A mixed-integer linear programming approach. Electric Power Systems Research, 114:93–100, 2014.

[31] Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. Arima models to predict next-day electricity prices. IEEE transactions on power systems, 18(3):1014–1020, 2003.

[32] Martin Haugh. Generating random variables and stochastic processes. Monte Carlo Simulation: IEOR EA703, 2004.

[33] Alex Schwarzenberg-Czerny. On matrix factorization and efficient least squares solution. Astronomy and Astrophysics Supplement Series, 110:405, 1995.

[34] Doug Hakkarinen and Zizhong Chen. Algorithmic cholesky factorization fault recovery. In Parallel & Distributed Processing (IPDPS), 2010 IEEE International Symposium on, pages 1–10. IEEE, 2010.

[35] S. Mishra, C. Würsig, and I. Palu. Multivariate scenario generation an arima and copula approach. International Journal of Modeling and Optimization, 2018.

[36] Marco Caserta and Stefan Voß. A math-heuristic algorithm for the dna sequencing problem. In International Conference on Learning and Intelligent Optimization, pages 25–36. Springer, 2010.

[37] Fady Youssef Melhem, Olivier Grunder, Zakaria Hammoudan, and Nazih Moubayed. Energy management in electrical smart grid environment using robust optimization algorithm. IEEE Transactions on Industry Applications, 2018.

[38] Antonio J Conejo and Luis Baringo. Power systems. In Power System Operations, pages 1–15. Springer, 2018.

[39] Antonio J Conejo, Luis Baringo, S Jalal Kazempour, and Afzal S Siddiqui. Transmission expansion planning. In Investment in Electricity Generation and Transmission, pages 21–59. Springer, 2016.

[40] Sérgio Haffner, Luís Fernando Alves Pereira, Luís Alberto Pereira, and Lucio Sangio Barreto. Multistage model for distribution expansion planning with distributed generation—part i: Problem formulation. IEEE Transactions on Power Delivery, 23(2):915–923, 2008.

[41] Amin Khodaei and Mohammad Shahidehpour. Microgrid-based co-optimization of generation and transmission planning in power systems. IEEE transactions on power systems, 28(2):1582–1590, 2013.

[42] Ehsan Hajipour, Mokhtar Bozorg, and Mahmud Fotuhi-Firuzabad. Stochastic capacity expansion planning of remote microgrids with wind farms and energy storage. IEEE Transactions on Sustainable Energy, 6(2):491–498, 2015.

[43] Manijeh Alipour, Behnam Mohammadi-Ivatloo, and Kazem Zare. Stochastic scheduling of renewable and chp-based microgrids. IEEE Transactions on Industrial Informatics, 11(5):1049–1058, 2015.

[44] Iman Goroohi Sardou and Ehsan Azad-Farsani. Network expansion planning with microgrid aggregators under uncertainty. IET Generation, Transmission & Distribution, 2018.

[45] Luis Baringo and Ana Baringo. A stochastic adaptive robust optimization approach for the generation and transmission expansion planning. IEEE Transactions on Power Systems, 33(1):792–802, 2018.

[46] Miguel Asensio, Pilar Meneses de Quevedo, Gregorio Munoz-Delgado, and Javier Contreras. Joint distribution network and renewable energy expansion planning considering demand response and energy storage —part i: Stochastic programming model. IEEE Transactions on Smart Grid, 2016.

[47] Zhirong Xu, Ping Yang, Chengli Zheng, Yujia Zhang, Jiajun Peng, and Zhiji Zeng. Analysis on the organization and development of multi-microgrids. Renewable and Sustainable Energy Reviews, 2017.

[48] Pan Wu, Wentao Huang, Nengling Tai, and Shuo Liang. A novel design of architecture and control for multiple microgrids with hybrid ac/dc connection. Applied Energy, 2017.

[49] Bo Zhao, Meidong Xue, Xuesong Zhang, Caisheng Wang, and Junhui Zhao. An mas based energy management system for a stand-alone microgrid at high altitude. Applied Energy, 143:251–261, 2015.

[50] Thillainathan Logenthiran, Dipti Srinivasan, and Ashwin M Khambadkone. Multi-agent system for energy resource scheduling of integrated microgrids in a distributed system. Electric Power Systems Research, 81(1):138–148, 2011.

[51] Vitor N Coelho, Miri Weiss Cohen, Igor M Coelho, Nian Liu, and Frederico Gadelha Guimarães. Multi-agent systems applied for energy systems integration: State-of-the-art applications and trends in microgrids. Applied Energy, 187:820–832, 2017.

[52] Chun-Xia Dou and Bin Liu. Multi-agent based hierarchical hybrid control for smart microgrid. IEEE transactions on smart grid, 4(2):771–778, 2013.

[53] Runfan Zhang and Branislav Hredzak. Distributed finite-time multi-agent control for dc microgrids with time delays. IEEE Transactions on Smart Grid, 2018.

[54] Roberto S Netto, Guilherme R Ramalho, Benedito D Bonatto, Otavio AS Carpinteiro, AC Zambroni de Souza, Denisson Q Oliveira, and Rodrigo AS Braga. Real-time framework for energy management system of a smart microgrid using multiagent systems. Energies, 11(3):656, 2018.

[55] AngularJS — Superheroic JavaScript MVW Framework. `https://angular.io//`. Online; Accessed: September, 2018.

[56] Node.js. `https://nodejs.org/en/`. Online; Accessed: September, 2018.

[57] Guido Van Rossum and Fred L Drake Jr. Python tutorial. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995.

[58] Ludwig Von Bertalanffy. Problems of life; an evaluation of modern biological thought. Wiley, 1952.

[59] Rajesh Karki, Roy Billinton, and Ajit Kumar Verma. Reliability Modeling and Analysis of Smart Power Systems. Springer, 2014.

[60] S. Mishra, C. Bordin, J. Fornes, and I. Palu. Reliability framework for power network assessment. E3S Web of Conferences, 2018.

[61] S. Mishra, C. Bordin, and I. Palu. Rnr: Reliability oriented network restructuring. International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 2018.

Acknowledgements

"To everyone who, ∃ in my life, truly loves me and fills my life with friendship, love and kindness. To everyone who will come into my life."

Abstract
Models for Modern Power Distribution System Planning

With an objective of cleaner and sustainable power production the electrical power sector is experiencing a shift from centralized power flow to a decentralized and distributed one. Consequently Non-dispatchable (renewable) energy resources have gained an increased share in the total energy mix. In addition, the grid ancillary services like demand side management where the consumer interacts with the energy network is reshaping the power distribution system. This transition introduces many technical and economic challenges in maintaining the balance between supply and demand of power. Apart from that the challenges in reserve capacity allocation to meet the demand at all times. Moreover, the strategic planning for power system security and balance is the focal point for a smooth transition. The context therefore is from the distribution system planner's strategic planning problem for expansion. Expansion refers to capacity expansion of existing power production units and transmission lines to meet the growing demand or better utilize the production from renewable (non-dispatchable) resources. This thesis presents a collection of research works focused on optimal power system expansion planning for modern power distribution system. The key challenges addressed are ranging from classification of wind power variations to optimal power network expansion and power system adequacy and reliability. The mathematical models are developed and presented with applications for each of the aforementioned challenges.

The significant contributions and practical implications of this thesis can be summarized as follows: (a) Ramping behavior analysis (RBA) model for classification of wind power variations (b) multi-variate scenario generation (c) Coordinated microgrid (CoMG) model for optimal investment in power generation and transmission expansion planning considering uncertainty (d) Reliability framework to evaluate condition of power network using indicators (e) reliability oriented network restructuring (RNR) model to determine optimal investment for power distribution system expansion and network reconfiguration.

The thesis has three chapters that discuss various aspects of the power network expansion problem. For strategic planning an accurate and precise classification of stochastic variables such as wind is important. The thesis begins with the chapter on ramping behaviour analysis (RBA) of wind power production. RBA is a novel algorithm that characterizes the time-series wind data to a series of ramp events. RBA is applied to time-series data from wind farms to extract the ramp events. Each ramp event has a peak and valley that shows the trajectory of each event. RBA improves the operational planning of wind farms with the mechanism of event detection. Successively a multi-variate scenario generation model is developed capturing the correlation between the uncertain variables. The variables are wind, demand and price that are correlated to each other. These scenarios are input to the stochastic programming model(CoMG) presented in chapter two.

The second chapter emphasizes on optimal investment decision making strategies for generation and transmission expansion planning. An innovative coordinated decision making process is introduced through a multi-agent-system on top level. And a stochastic multi-period math-heuristic model at the bottom for optimal expansion under uncertainty in demand, wind and prices. The novel math-heuristic model, coordinated microgrid expansion planning, is applied to real-world planning for expansion.

The third chapter further enhances the scope of planning problem by including power system adequacy and reliability. A power system reliability framework is proposed considering the life-cycle of the power apparatus, consumer satisfaction index and geographical terrain. Subsequently, reliability oriented power system restructuring model is formulated as a Non-Linear deterministic multi-period model. The model contains an AC-

optimal power flow formulation with a grid reconfiguration based on the weights allocated through the proposed framework. The model evidently improves the expansion decision through consideration of existing network condition.

The models developed are applied to real-world planning problems for distribution network operator. The models are validated with real data from industrial partners. As an open-source initiative, a modified version of two models developed are made available via Github repository.

keywords: *wind power swings*, *mathematical optimization*, *generation and transmission expansion planning*, *decision support system*, *power system reliability and adequacy*, *network restructuring*, *energy informatics*

Lühikokkuvõte
Kaasaegsete jaotusvõrkude planeerimise mudelid

Puhtama ning jätkusuutlikuma elektrienergia toomise eesmärgi nimel toimub elektrisüsteemi üleminek tsentraliseeritud tootmiselt hajatootmisele. Eelnevast tulenevalt on taastuvenergiaallikate kasutamine ning taastuvenergia osakaal lõpptarbimises kasvanud. Lisaks on hakatud tegelema tarbimise juhtmisega ning see avaldab mõju elektrijaotussüsteemi ümberkujunemisele.

Hajatootmisele üleminek toob endaga kaasa mitmeid tehnoloogilisi ja majanduslikke väljakutseid, mis on seotud elektrienergia tarbimise ja tootmise tasakaaluga. Selleks, et tagada sujuv üleminek hajatootmisele ning samas säilitada süsteemis tasakaal ja varustuskindlus, on väga oluline tegeleda süsteemi strateegilise planeerimisega. Arvestada tuleb nii tarbimise kasvu kui ka taastuvenergiaallikate kasutamise suurenemisega.

Käesolevas töös vaadeldakse optimaalset elektrisüsteemi laienemise planeerimist kaasaegsetes jaotusvõrkudes. Peamised väljakutsed on tuuleenergia muutlikkus, optimaalne elektrivõrgu laienemine ning elektrisüsteemi töökindlus ja piisavus. Iga eeltoodud väljakutse jaoks koostati matemaatiline mudel.

Töö olulised panused ja praktilised tulemused on järgnevad: (a) võimsusmuutuste analüüsi (Ramping behavior analysis - RBA) mudel tuuleenergia muutuste klassifitseerimiseks (b) mitmemõõtmelise stsenaariumi koostamine (c) koordineeritud mikrovõrgu (CoMG) mudel optimaalsete investeeringute jaoks elektrienergia tootmises ja jaotamises võttes arvesse ebamäärasust (d) töökindluse raamistik elektrisüsteemi kasutamise indikaatorite hindamiseks (e) Töökindlusele orienteeritud võrgu restruktureerimise (Reliability oriented network restructuring – RNR) mudel jaotusvõrgu rekonfigureerimise optimaalse inevsteeringu tegemiseks.

Töö on jaotatud kolmeks peatükiks, mis käsitlevad elektrisüsteemi laienemise probleemi erinevaid aspekte. Töö algab peatükiga, mis käsitleb tuuleenergia tootmise võimsusmuutuste analüüsi. RBA on uudne algoritm, mis teisendab tuuleandmete aegrea sündmuste reaks. Igal sündmusel on tipp ja põhi, mis näitavad sündmuse trajektoori. RBA parandab sündmuse tuvastamise abil tuuleparkide planeerimist. Mitmemõõtmelise stsenaariumi koostamise mudel on arendatud ebamääraste muutujate vahel seose määramiseks. Muutujad on tuule tugevus, nõudlus ja hind ning kõik muutujad on omavahel soetud. Need stsenaariumid on sisendiks stohhastilisele programmeerimise mudelile, mida käsitletakse teises peatükis.

Teine peatükk keskendub tootmise ja ülekandmise edasiarenduse planeerimise optimaalse investeerimisotsuse tegemise strateegiatele. Innovatiivset koordineeritud otsuse tegemise protsessi tutvustatakse korrapärase kõrgema tasandi iseseisva arvutisüsteemi abil. Uudset matemaatilis-heuristilist mudelit ja koordineeritud mikrovõrgu edasiarenduse planeerimist rakendatakse reaalsete arenduste planeerimiseks.

Kolmas peatükk tõstab esile planeerimisprobleemi ulatuse, mis keskendub energiasüsteemi töökindlusele ja piisavusele. Pakutakse välja energiasüsteemi usaldusväärsuse raamistik, võttes arvesse energiaseadmete elutsüklit, tarbija rahulolu indeksit ja geograafilist maastikku. Järgnevalt on RNR mudel formuleeritud nagu mittelineaarne ette määratud mitmeperioodiline mudel. Mudel sisaldab optimaalset vahelduvvoolu energiavoo formuleerimist koos võrgu rekonfiguratsiooniga, mis põhineb kaalude jaotamisel läbi esitatud raamistiku. Mudel täiustab silmnähtavalt edasiarenduse otsust olemasoleva võrgustiku tingimuse abil.

Arendatud mudelid on kohaldatud reaalsete jaotuvõrgu planeerimise probleemide lahendamiseks. Mudelid on valideeritud reaalsete partneritelt saadud andmetega. Kahe mudeli muudetud versioonid on tehtud kättesaadavaks GitHub'i keskkonnas.

Võtmesõnad: *tuuleenergia juhuslikus*, *matemaatiline optimeerimine*, *tootmise ja jaotamise laienemise planeerimine*, *otsuste toetamise süsteem*, *elektrisüsteemi töökindlus ja piisavus*, *võrgu restruktueerimine*, *energia informaatika*

Appendix 1

S. Mishra, M. Leinakse, and I. Palu "Wind Power Variation Identification using Ramping Behavior Analysis" Elsevier Energy Procedia, pages - 141:565–571, 2017

Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 141 (2017) 565–571

Energy Procedia

www.elsevier.com/locate/procedia

4th International Conference on Power and Energy Systems Engineering, CPESE 2017, 25-29 September 2017, Berlin, Germany

Wind power variation identification using ramping behavior analysis

Sambeet Mishra*, Madis Leinakse, Ivo Palu

Tallinn University of Technology, Department of Electrical Power Engineering and Mechatronics, Ehitajate tee 5, Tallinn 19086, Estonia

Abstract

Harvesting energy from renewable sources has become prominent since the use of fossil fuels became unsustainable. Traditional practice for mitigating the energy demand around globe majorly consists of utilizing conventional sources and injection of renewables as and when available. The continuous and exponential growth in consumption alongside the need to reduce the carbon footprint and to counter the climate change has paved the way for Renewable Energy Sources (RES). Availability and maturity in technology made wind and PV (photo-voltaic) the most prominent among others. Per contra, the inherent variations in the weather in form of wind speed, solar irradiance act as a barrier in utilizing the full potential. The variations, ramp events, in case of wind energy have adverse effects on determining the reliability, economical profitability, and flexibility. Accurate recognition of the wind ramp events can improve energy management, forecasting and causality. This paper proposes a data analysis oriented approach exploring the pre-processing technique of wind power variations using moving average filter, followed by noise extraction and separating the power swings. Further clustering the power swings utilizing K-means clustering technique. The proposed technique improves the power swings identification process by reducing the noise content.

© 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the 4th International Conference on Power and Energy Systems Engineering.

Keywords: Power swings; Renewable energy resources; data clustering; ramp event detection;

1. Introduction

The systems are static, quasi-static or dynamic with respect to time. An event in a system can be a continuous or discrete time process. Ramp is used to explain an event where a sudden positive or negative swing occurs within a period. Harvesting energy from every possible source before the conventional sources run out of reserve pushed us for the change in regime. Now that the energy requirement is foreseeing a rapid growth over time, the challenge became multi objective – manage consumption, increase production from alternative sources and ICT (Information and Communication Technologies) integrated smart controlling of the overall system. Most renewable sources of energy are non-deterministic owing to the factor of random availability as in wind speed, solar irradiance. This unreliability acts as the major hindrance apart from the economic standpoint in wide scale implementation. Today's

* Sambeet Mishra. Tel.: +372-620-3759.
E-mail address: sambeet.mishra@ttu.ee

1876-6102 © 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the 4th International Conference on Power and Energy Systems Engineering.
10.1016/j.egypro.2017.11.075

practice consists of majority of production and reserve from conventional sources and penetrative renewable energy. The ramp events in power conversion from RES have adverse effects on reliability, economical profitability and flexibility. Characterizing the ramp behavior of renewable sources fosters the chances for better management and thus improving the system operation. Inherently RES have multi-level time varying uncertainty. In wind energy, ramp-ups are usually caused by intense low pressure systems, low level jets, thunder storms, wind gusts, climatic variations [1]. Again, the statistical model presented providing insights to the ramping events- frequency of occurrences and severity levels. The study of ramp events are utilized in system control and day-ahead forecast. Sevlian and Rajgopal et. al. used a dynamic programming recursion to analyze ramp events by virtue of a statistical model [2]. Wherein Florita et. al. used swinging door algorithm and indicated the fine tune required to improve the ramping event analysis in solar and wind energy [3]. Ouyang et. al. illustrated the current forecast models based on ramping events. Then extended suggestions in contrast with numerical weather prediction [4]. Bianco et al. presented wind ramp detection, time forecast, observed ramp and impact of the up-down events on the grid operation [5]. In [6] ramping behavior analysis technique was elaborated for the ramp detection in RES. In [7] a time series simulation for the large wind farm in turbulent scenario was described. In [8] a spacio-temporal model for the short-tern wind power forecast model was developed. However, the wind power variations are highly dependent on the precise peaks identification and a setting a proper threshold. The noise in the dataset often lead to misclassification or over-estimation of the ramp events. In this paper, the focus was on the pre-processing of wind power data that give rise to precise time-series data removing the noise content and preserving the swing property of the original data. Consecutively identifying the peaks and clustering the peaks into groups classified the variations accurately. This article is divided into four sections- first the wind power variations explaining the ramp events, second wind power data filtering method, followed by clustering the data sets to emphasize the significance and finally the conclusion.

2. Wind power variations

The Ramping Behavior Analysis (RBA) is a relatively recent field of study in the domain of RES. The causality of the ramp events are not clearly traced. In wind energy, sudden oscillation of output power from wind turbine and high input power injection with notable pace is an identifier of ramping. Performing the ramping analysis on both the input and output is a key point. There exists no particular increment in magnitude or oscillation ranges in literature for the RES to characterize the events. As a result, there are multiple ramp events in various sets of thresholds. The second fold of problem is based on reliability statement in agreement with magnitude and times of occurrences in the period of analysis. Various levels of system with multiple time steps, intermediate delays, and the instantaneous weather changes make the system highly random. Statistical model requires iterative investigations with multiple thresholds, data sets and time stamps to make an inference, setting aside the Hybrid Renewable Energy Sources (HRES). The traditional definitions for the wind ramp events are distinguished by the pre-determined threshold values. The four equations of threshold values below are widely used definitions in literature. [4].

$$\left|P(t+\Delta t)-P(t)\right| > P_{thr} \tag{1}$$

$$\left|\max\left(P\left[t,t+\Delta t\right]\right)-\min\left(P\left[t,t+\Delta t\right]\right)\right| > P_{thr} \tag{2}$$

$$\frac{\sum_{n=1}^{m}\left(P_{t+h}-P_{t+h-N}\right)}{N} > P_{thr} \tag{3}$$

$$\frac{\left|P(t+\Delta t)-P(t)\right|}{\Delta t} > P_{thr} \tag{4}$$

Where P, t, Δt, P_{thr} stand for power generated, time, time interval and power threshold respectively. The yardstick of the analysis depends on the threshold, and change in magnitude of wind power production over a period. Considering only the ends or difference between maximum or minimum power productions including end approach has a disadvantage in the form of special case inclusion or exclusion [3].

Setting up a threshold depends on multiple factors as in grid topology, size of turbine, placement and region. Ramp refers to significant increase or decrease in wind power within a set time period. A swing can at times be a special case, as in the wind speed drops below the limit or sudden increase due to untraceable factors. RBA consists of a) ramp-up b) ramp-down c) rise-time d) fall-time e) ramp-up/down rate [6]. The objective is to identify the set of significant ramps considering the time as a reference and the difference between two consecutive high and low values give rise to peak points in both ascending and descending order. It is to be considered that the minute changes have to be ignored to catalogue only the significant variations. Fig. 1(a) explains the RBA in terms of a power swing above threshold. A threshold value is often set, above which the ramp events are significant for the system depending on the scenario. Ramp-up is an event of rise in positive number, where the ramp-down in opposite direction. Rise and fall time are successive time steps taken to reach the corresponding peak. Ramp up rate goes by $(peak - lowest) / Timesteps$.

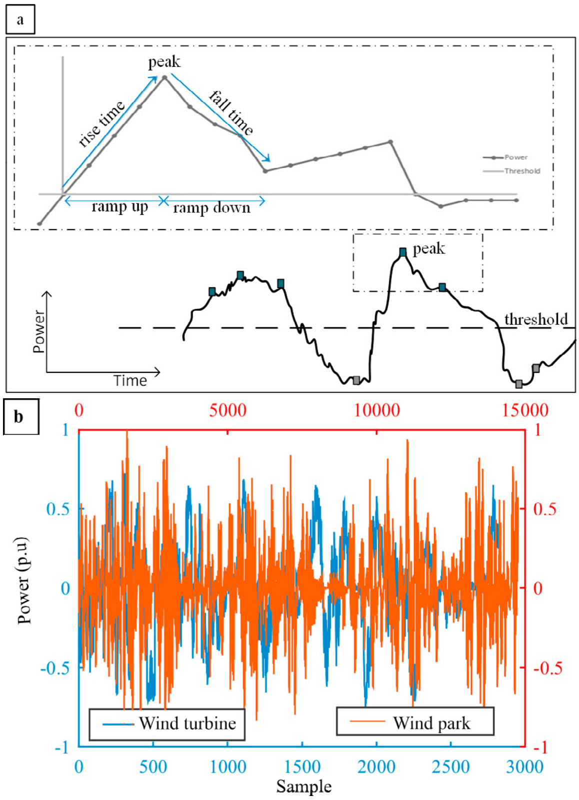


Fig. 1.(a) ramp event (b) wind power swings

3. Wind Power data filtering

The data sets considered for this analysis have 10 min time step and were measured in a wind park consisting of 17 wind turbines, 2.5 MW each. The average wind speed in the region is 6 m/sec. Fig. 2 represents the wind park power generation of four months. It was observed that the aggregated output for four months shows continuous power swings and number of swings reduces upon stretching the data. Fig. 1(b) represents both power data from the wind turbine and aggregated wind park data. It can be observed from fig. 2 that the individual turbine data is more erratic than that of the wind park aggregated data that is causing the excess noise content. Exponential moving average filter technique is used for smoothing the data and presented in fig. 3(a) and 3(b) along with the extracted noise content. The equation 5 represents the calculation for moving average.

$$f(c) = [(c - f(p))\omega] + f(p) \tag{5}$$

Where f, c, p, ω stands for exponential moving average, current value, previous value and $\omega=2/(N+1)$ weight factor wherein N is number of periods respectively.

Filters are data processing techniques that can smooth out high-frequency fluctuations in data or remove periodic trends of a specific frequency. Using the aforementioned filtering technique the noise is extracted from the original data and presented in fig. 3(a) and 3(b). Wherein A- Original wind power, B- smoothed power curve and C - noise content in the data. Fig. 1(b) represent the wind power swing data evaluated from differentiation of the derived smooth data that refers to the peak points or local maxima.

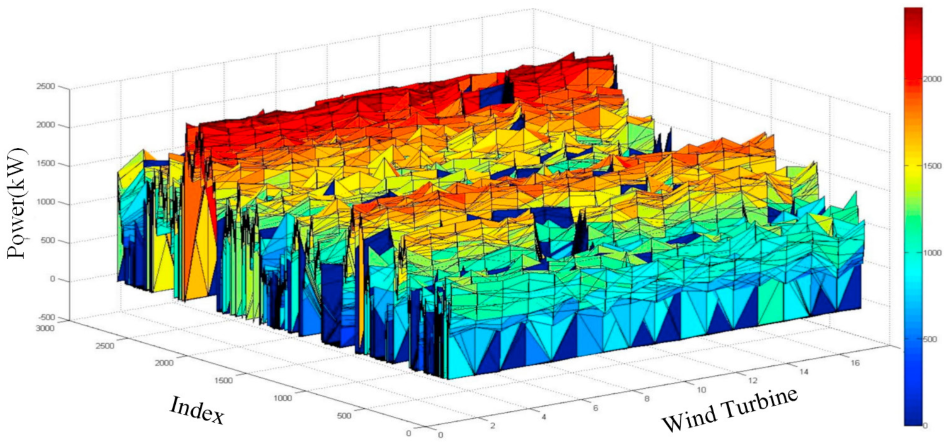


Fig. 2. Wind park power data.

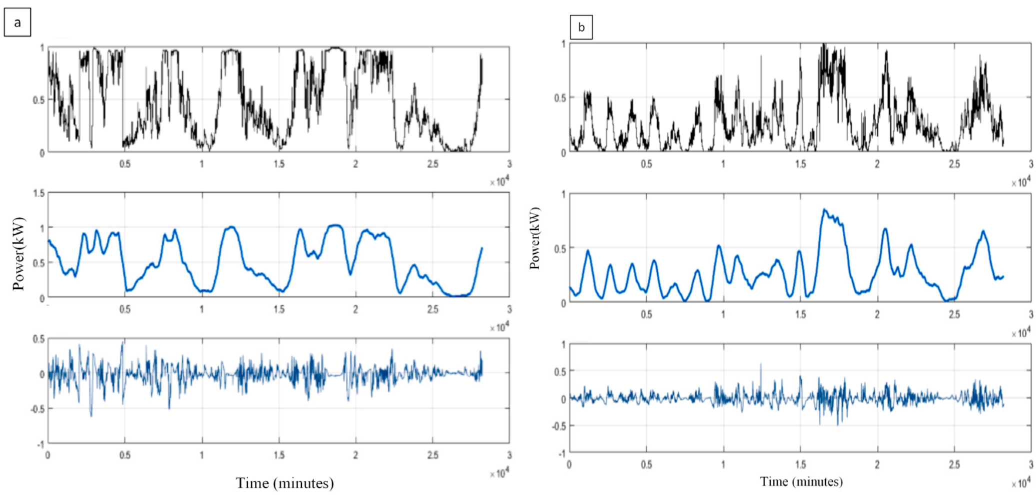


Fig. 3. (a) wind turbine power data; (b) wind park aggregated data.

4. Clustering power swings

In this section, the derived data is clustered to categorize the power swings. Clustering the data indicates the pattern in the swings by aggregating them into appropriate number of groups. K-means clustering technique was selected to perform the analysis due to simplicity and reasonable computational time. To determine optimal number of clusters, two different approaches have been combined. The number of clusters has been analyzed by the knee-point of sum of distances and by using silhouette coefficient of the clustered data.

4.1. K-means clustering algorithm

The K-means clustering algorithm also known as Lloyd's algorithm [9] is applied for classification of objects into K number of groups based on attributes. The algorithm is based on the minimization of squared Euclidean distance between the objects and centers of the assigned clusters [10]:

$$\min(E) = \min(\sum_{i=1}^{K} \sum_{x \in C_i} d(x, z_i)) \tag{6}$$

where z_i, is the center of cluster C_i and $d(x, z_i)$ is the Euclidean squared distance between point x and cluster center z_i.

4.2. Data sets used for clustering

Two data-sets were used for clustering: data set 1 includes single wind turbine data and data set 2 wind park data. Both data-sets include observations with four attributes, namely: original data, smoothed data, swing data, noise content.

4.3. Optimal number of clusters

Fig. 4(a) displays the total sum of Euclidean square distances as a function of K, number of clusters. The figure displays the results for both data-sets, which behave similarly. Based on the curve, the decrease rate of the sum of distances decreases rapidly until around 10 clusters. Thus, it is possible to conclude that for clustering, values above 10 offer limited gains with respect to the sum of distances.

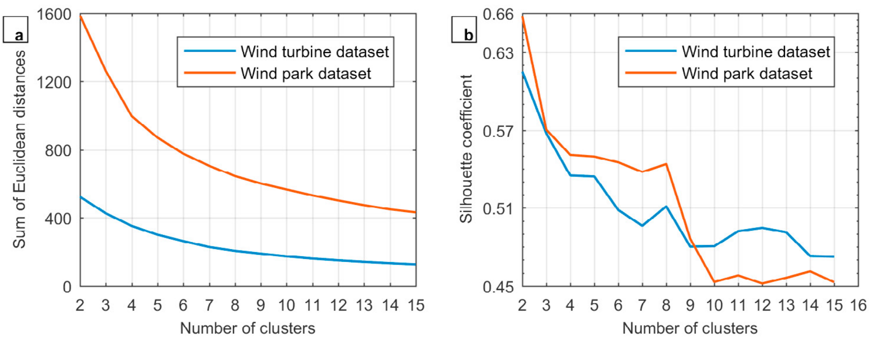


Fig. 4. Number of clusters and (a) the sum of Euclidean squared distances; (b) silhouette coefficient.

A second way to analyze the suitability of the value of K is to use the values of silhouette coefficient [10] – [11]:

$$SC = \frac{1}{K} \sum_{j=1}^{K} S_j \tag{7}$$

where S_j is the local silhouette coefficient:

$$S_j = \frac{1}{n_j} \sum_{i=1}^{n_j} S_i \tag{8}$$

and the silhouette values of points [11] S_i are defined as:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)} \tag{9}$$

where S_i is the silhouette value for point i, b_i is the minimum average distance from point i to points of other clusters; a_i is the distance of point i to the other points of the same cluster.

According to fig. 4(b), the silhouette coefficient value decreases with the increase of the number of clusters, indicating that in case of larger number of clusters the dissimilarity between the clusters decreases. An interpretation for the values of the silhouette coefficient was given in [12]. Values equal to and below 0.25 indicate that no substantial structure was found. The found structure can be considered to be weak and possibly artificial if the value of the silhouette coefficient is 0.50 or lower. Values above 0.50 and up to 0.70 indicate a reasonable structure. This value range is also achievable with the used datasets by using 8 or lower value for K, number of clusters.

Considering both aspects of the clustering, the sum of squared Euclidean distances and silhouette coefficient value, 8 can be chosen as a suitable number of clusters. This value offers a compromise between achieving minimal total distance of objects from cluster centers and describing a reasonable structure of the data. A lower value would strengthen the structure of the clusters, but would also increase the sum of Euclidian distance squares between the clustered objects and the centers of the clusters. This means that the ramps classified to the same clusters would have higher variance.

4.4. Clustering Results

Previously, the number of clusters were selected to be eight. With this number of clusters, the first set of data, wind turbine data was clustered according to the silhouette plot shown in fig. 5(a). The created clusters seem to be close to each other, indicated by the negative silhouette values and silhouettes with sharp tips. For comparison, the result of two clusters are presented in fig. 5(b). The clusters have flatter tips and lower number of objects have negative values, thus two clusters offer in this case better separation of created clusters. The second set of data, wind park data, clusters into two clusters according to fig. 5(c) and to eight according to fig. 5(d). Again, clustering the dataset into two clusters leads to more distinguishable clusters.

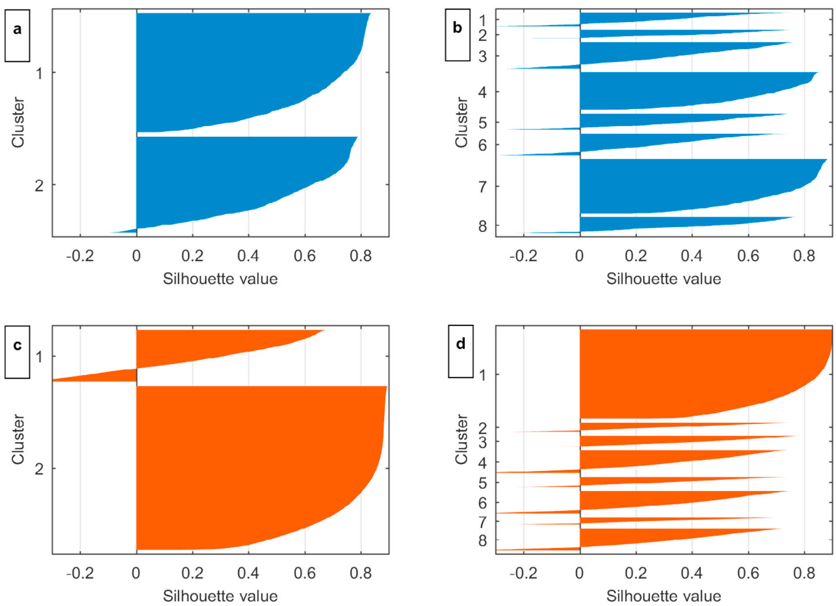


Fig. 5. Silhouette plots of aggregated data: (a) wind turbine data and 2 clusters (b) wind turbine data and 8 clusters (c) wind park data and 2 clusters (d) wind park data and 8 clusters.

5. Conclusion

The proposed analysis sheds light into identification of power swings from wind parks and individual turbines. Extracting the noise content by exponential moving average from the data prior to applying RBA provides accurate measure to the variations and computationally efficient. Clustering the variations to catalogue them provides required order of filtering. Optimal number of cluster selection through the total Euclidean distance measure and silhouette coefficients led to successful aggregation of the data. In continuation of the current work the event classification in multiple orders with relations, causality and novel clustering methods are being prepared.

References

[1] C. Kamath, Understanding. *Wind Ramp Events Through Analysis of Historical Data.* Transmission and Distribution Conference and Exposition IEEE PES pp. 1-6 April 2010

[2] R. Sevlian and R. Rajagopal. *Detection and Statistics of Wind Power Ramps.* IEEE Transactions on power system Vol. 28 No. 4 November 2013

[3] A. Florita, B.-M. Hodge, and K. Orwig. *Identifying Wind and Solar Ramping Events.* IEEE Green Technologies Conference 2013

[4] T. Ouyang, X. Zha and L. Qin. *A Survey of Wind Power Ramp Forecasting.* Energy and Power Engineering Vol. 5 No. 4B, pp. 368-372, 2013
[5] L. Bianco, I. V. Djalalova, J. M. Wilczak, J. Cline, S. Calvert, E. Konopleva-Akish, C. Finley, and J. Freedman. *A Wind Energy Ramp Tool and Metric for Measuring the Skill of Numerical Weather Prediction Models.* Wea. Forecasting, 31, 11371156 2016
[6] S. Mishra and I. Palu. *A user interface tool for ramping behaviour analysis of renewable energy.* IEEE Electric Power Quality and Supply Reliability August 2016
[7] P. Sorensen, N. A. Cutululis, A. Vigueras-Rodriguez, L. E. Jensen. *Power Fluctuations From Large Wind Farms.* IEEE Transactions on Power Systems, Vol. 22, No. 3, pp. 958 965 Aug 2007
[8] M. He, L. Yang, J. Zhang, V. Vittal. *A Spatio-Temporal Analysis Approach for Short-Term Forecast of Wind Farm Generation.* IEEE Transactions on power systems, Vol. 29, No. 4 July 2014
[9] S. P. Lloyd. *Least Squares Quantization in PCM* IEEE Transactions on Information Theory. Vol. 28, 1982, pp. 129-137
[10] G. Grigoras, F. Scarlatache and G. Cartina. *Load estimation for distribution systems using clustering techniques.* Optimization of Electrical and Electronic Equipment (OPTIM), 2012 13th International Conference on, Brasov, 2012 pp. 301-306
[11] P. J. Rousseeuw. *Silhouettes: a Graphical Aid to the Interpretation and Validation Cluster Analysis.* Journal of Computational Applied Mathematics, Vol. 20, pp. 53-65, 1987
[12] L. Kaufman and P. J. Rouseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis.* Hoboken, NJ: John Wiley & Sons, Inc., 1990

Appendix 2

S. Mishra, M. Leinakse, I. Palu, and J. Kilter "Ramping Behaviour Analysis of Wind Farms" EEEIC2018, pages - 1(1):1–7, 2018

Ramping Behaviour Analysis of Wind Farms

Sambeet Mishra, Madis Leinakse, Ivo Palu, Jako Kilter
Department of Electrical Power Engineering and Mechatronics
Tallinn University of Technology
Tallinn, Estonia

***Abstract*—Wind power ramps are power swings induced by the variations of wind speed. The ramp pattern of wind farms and turbines differs. In this paper, the pseudo code of ramping behaviour analysis with primary functionalities is presented. Genetic algorithm is applied to aggregate turbine-wise wind power data from the wind farm. This pre-processing is done to shrink the volume of the data through a aggregated representative data. The ramping behaviour analysis results are analysed by using K-means clustering to identify patterns in the ramping behaviour of the wind farm. The results indicate that for certain wind ramp parameter combinations, the identified patterns are stronger. In this paper, it is shown that the ramp energy values calculated based on ramping behaviour analysis parameters provide a strong cluster structure with low number of type events.**

***Keywords*—wind energy; wind ramp; ramping behaviour analysis; genetic algorithm; clustering; renewable energy resources;**

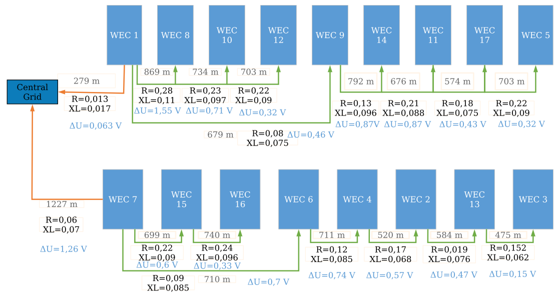


Fig. 1. Layout of the analysed wind farm.

I. INTRODUCTION

Transition from conventional energy sources to renewable energy sources (RES) has become one of the biggest transitions in the last decade. Climate change and utilization of renewable resources are one of the primary reasons. Solar and wind are by contrast the most prominent RES. Technologies to extract energy from RES are mature enough and fast growth in ICT has worked as the catalyst for the grid integration. This transition of unidirectional flow of power from source to demand is changing with distributed energy resources (DER). DER in turn gives rise to micro and smart grids with small and medium capacity of generation.

Wind, being a natural resource, is variable and unpredictable by nature. As one goes further from the Mediterranean region, the wind energy becomes more prevalent as the availability of solar energy diminishes. A sudden wind power swing is called a ramp. It could be up ramp or down ramp based on the movement. Though the ramps are often swings that settle down quite fast. Although near precise weather forecasts are extremely helpful for operating the wind farms, frequent variations in the wind speed and intervals make the operation more complex and motivate the use of energy storage technologies. Nevertheless, battery banks are used as an intermittent source to maintain uninterrupted energy supply.

A wind ramp of an individual wind turbine differs from the ramp of the wind farm [1]. In [2] frequency regulation for the wind power variation are discussed. In [3] ramp event identification is presented. In [4] a tool to characterize ramp events is presented. A smoothing technique is applied to reduce the noise content in the data [5], [6]. However, processing big-data from a wind farm and identifying ramp events require high computing power. Averaging the wind time series data looses the time significance. For instance a significant ramp event at a small window gets lost while looking at a bigger time window. In this paper, a wind farm consisting of 17 wind turbines is considered for the analysis. The layout of wind turbines in the farm are presented in Fig. 1. The genetic algorithm (GA) is used to optimally aggregate the data from 17 wind turbines. The granularity of the precise ramp event identification is preserved across hourly time resolution. Wind ramp events are identified with the aggregated data. In order to identify type events of the wind farm, ramping behaviour analysis results are clustered. Different combinations of ramping behaviour analysis parameters are used for detecting patterns.

The rest of this paper is organized as follows. In section II-A, a method for aggregation of power data is presented. Section II-B provides an overview of ramping behaviour analysis (RBA) and presents an algorithm for identifying the values of RBA parameters. Section II-C ends with a description of K-means clustering, which is used in section III for identifying patterns in RBA results. The clustering was done with different parameter weights and combinations to detect the patterns. The results of the paper are summarized in section IV.

This paper was funded by the European Union via the European Regional Development Fund.

978-1-5386-5186-5/18/$31.00 © 2018 IEEE

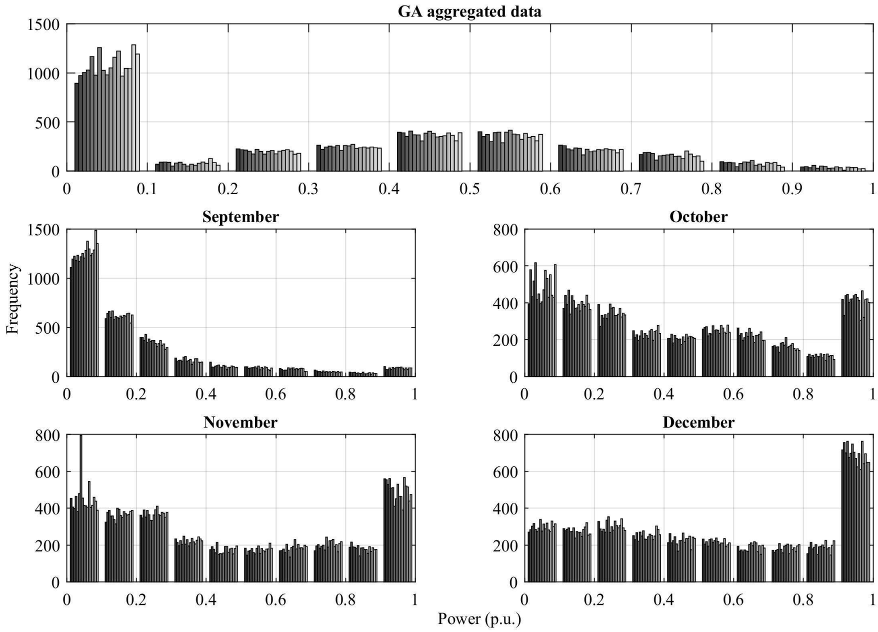


Fig. 2. Histograms of normalized and aggregated wind farm data.

II. Wind Data Aggregation and Ramp Event Identification

A. Data Aggregation

A genetic algorithm (GA) is a local search optimization technique primarily applied to optimization problems that are highly non-linear, non-differential or discontinuous. Conceptually GA is based on the biological evolution process. The process randomly selects a pair out of the total population and crossover takes place to produce successors for next generation. This process of evolving toward an optimal solution classifies GA as an evolutionary technique [10]–[14].

The fitness function ff of the GA is (1). The ff considers the time horizon t, the turbine w, month m and y as the current data point. In the present context m is the winter period: September through December. Also w is 17 and t is 2920 hours. The power data is normalized by using the maximum value, to obtain normalised values between 0 and 1 p.u. Fig. 2 presents the histogram of the month wise power data from the wind farm and the optimized data. Typical month-wise power data from wind farm is combined and using GA one optimal representative array is generated for each turbine.

$$ff = \sum_{w=1}^{|w|} \sum_{j=1}^{|t|} y_j^w * (y_j^w - m_j)^2 \tag{1}$$

The data aggregation is applicable to identify and forecast significant ramp events. Specifically in long-term forecasting of the wind ramp events or building an event matrix. Wherein the volume of data can be reduced using GA. The average number of peaks identified individually from the raw data is 566 for wind turbine. Processing through GA the number of events identified is 618. Thus, the number of events identified is close to the original data. It means the frequency of ramp events in the data is preserved while the data is reduced four folds.

B. Ramping Behaviour Analysis

In paper [3], ramping behaviour analysis (RBA) was described to identify the ramp events. The Algorithm 1 depicts the pseudo code for the RBA algorithm. RBA consist of four functions: *peakval, risetime, falltime* and *ramprate*. Function *peakval* identifies the peak point through index values and stores them in i^+. Successively the valley points are stored in i^-. An example of peak value identification process in a data array is presented in Fig. 3. The x stands for the input array and D stands for the difference. The arrow sign denotes the peak values: 8 and 5. The rectangle outlines the pair of values compared to find the peak values: (1, -6) and (4, -2).

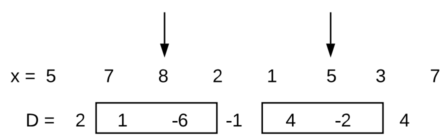


Fig. 3. Example of peak value identification algorithm

Functions *risetime* and *falltime* evaluate the time taken to reach the peak point and valley point and stores the values in $trise$, $tfall$ successively. $P(x)$ refers to the index value x of input power data P. Finally, *ramprate* function calculates the power change rate at which the peak or valley points were reached. The calculated ramp-up and ramp-down rates are stored in $rrate^+$, $rrate^-$, respectively.

Algorithm 1: Ramping Behaviour Analysis (RBA)

input wind power data array x ;
Function *peakval*
  $n = |x|$;
  while $i \leq n$ **do**
    $D_i = x_{i+1} - x_i \quad \forall \quad i$;
    if $D_i < 0 \quad \forall \quad i$ **then**
      $i^+ = i \in x$ [index value of D_{i+1}];
    else
      $i^- = i \in x$ [index value of D_i];
    end
  end
  Result: peak points identified
Function *risetime*
  $x_{max} = i^+$; $x_{peak} = i^+$;
  $x_{min} = (x_{max} - 1)$;
  if $P(x_{min}) \leq P(x_{max})$;
  then
    $P(x_{max}) = P(x_{min})$;
    $P(x_{min}) = P(x_{min} - 1)$;
  else
    $P(x_{min}) = P(x_{max})$;
    $trise = tscale * (P(x_{peak}) - P(x_{min}))$
  end
  Result: rise time identified for each peak
Function *falltime*
  if $P(x_{max}) \leq P(x_{min})$;
  then
    $P(x_{min}) = P(x_{max})$;
    $P(x_{max}) = P(x_{max} - 1)$;
  else
    $P(x_{max}) = P(x_{min})$;
    $tfall = tscale * (P(x_{max}) - P(x_{peak}))$
  end
  Result: fall time identified for each peak
Function *ramprate*
  $rrate^+ = \Big(P(x_{max}) - P(x_{min})\Big)/trise$;
  $rrate^- = \Big(P(x_{max}) - P(x_{min})\Big)/tfall$;
  Result: ramp-up and ramp-down rate identified for each peak

C. K-means Clustering Algorithm

The K-means clustering algorithm, also known as Lloyd's algorithm [15] is used for classification of objects into K number of groups based on attributes. In case of wind ramps, the ramping behaviour analysis uses 7 attributes for each event, which can be used for classifying the wind ramps to groups. The algorithm is based on the minimization of squared Euclidean distance between the objects and centers of the assigned clusters [7]:

$$min(E) = min(\sum_{i=1}^{K} \sum_{x \in C_i} d(x, z_i)) \tag{2}$$

where z_i is the center of cluster C_i and $d(x, z_i)$ is the Euclidean squared distance between point x and cluster center z_i.

III. Classification of Wind Ramps

A. RBA Parameter Subsets and Normalization

In ramping behaviour analysis, the wind ramp events were described by start P_{min1}, peak P_{max} and end P_{min2} value of power variation; rise speed and time; fall speed and time. Thus, for every significant variation, 7 parameter values are available. Start, peak and end value are in kW, rise and fall speed in kW/min, rise and fall time in minutes.

The clustering was applied to combinations of ramping behaviour analysis (RBA) parameters to determine the attributes of wind ramps that are suitable for event classification. Firstly, the RBA parameters were divided by the physical quantity, this led to formation of data subsets 1..3, shown in Table I. These subsets were not normalised, as within the subsets, the same units were used for all values. Subsets 4 to 6 were formed to test the impact of RBA parameter normalisation on clustering results. Subset 4 was not normalised. The values in subset 5 were normalised using median value of P_{max}, $w1$, and median value of rise and fall times $w2$. For subset 6, the maximum value of P_{max}, $w3$, and maximum value of time were used. In case of subset 6, all values in the subset were in the range of 0 to 1. The ramp energy parameters used for subset 7 are explained in section III-B.

TABLE I
Normalisation Bases Used for Subsets of Ramping Behaviour Analysis (RBA) Attributes

RBA Attribute	Data Subset						
	1	*2*	*3*	*4*	*5*	*6*	*7*
Pmin1	1	-	-	1	w1	w3	-
Pmax	1	-	-	1	w1	w3	-
Pmin2	1	-	-	1	w1	w3	-
Rise time	-	1	-	1	w2	w4	-
Fall time	-	1	-	1	w2	w4	-
Rise speed	-	-	1	-	-	-	-
Fall speed	-	-	1	-	-	-	-
Ramp energy	-	-	-	-	-	-	1

B. Evaluating the Strength of Clusters

The clustering results were evaluated in this study based on the value of silhouette coefficient SC (3) [7], [8].

$$SC = 1/K \sum_{j=1} KS_j \tag{3}$$

where S_j is the local silhouette coefficient

$$S_j = (1/n_j) \sum_{i=1} n_j S_i \quad (4)$$

and the silhouette values of points s_i are defined as

$$S_i = (b_i - a_i)/max(a_i, b_i) \quad (5)$$

where s_i is the silhouette value for point i, b_i is the minimum average distance from point i to points of other clusters; a_i is the distance of point i to the other points of the same cluster. The following interpretation of silhouette coefficient SC values from [9] was used to evaluate the cluster strength and choose the optimal number of clusters.

- ≤ 0.25: No substantial structure has been found
- 0.26 — 0.50: The structure is weak and could be artificial; please try additional methods on this data set
- 0.51 — 0.70: A reasonable structure has been found
- 0.71 — 1.0: A strong structure has been found

C. Clustering Based on Subsets of Ramping Behaviour Parameters

Clustering results in Fig. 4 indicate that the ramping time parameters (subset 2) offer better clustering results than peak and valley points (subset 1) and ramp rates (subset 3). A strong structure is obtained for subset 2 with at least 15 clusters. The marginal decrease of sum of squared Euclidean distances, Fig. 5, is largest for small number of clusters. Thus, with the increase of number of clusters, the benefit of additional clusters decreases. Subset 2, which displayed good

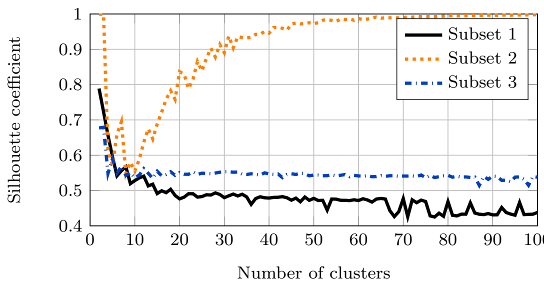


Fig. 4. Subset 1, 2 & 3: silhouette coefficient.

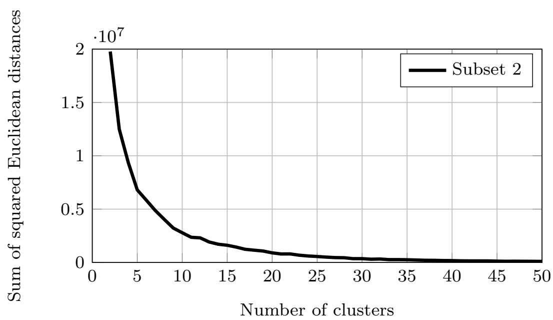


Fig. 5. Subset 8: sum of squared Euclidean distances.

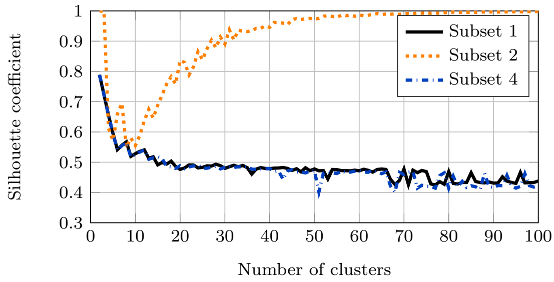


Fig. 6. Subset 1, 2 & 4: silhouette coefficient.

clustering characteristics, describes only the ramping duration of the wind ramps. The combination of subset 1 and subset 2 would describe the key points of the wind ramp (start, peak and end power) and also the duration of the event. Thus, a combination of the two subsets would have significantly higher value than subset 2. If subset 1 and 2 are combined without any normalisation of values, subset 4 is acquired. The clustering characteristics of subset 4 are similar to subset 1, as shown in Fig. 6. Reason is the data normalisation, which was not applied and due to large subset 1 values, subset 2 parameter values had small impact on clustering results.

Comparing the non-normalised subset 4 results with the normalised subsets 5 and 6, in Fig. 7, it is clear that subset 5 has different clustering characteristic. Reason is again the obtained set of values. In case of ramping times, the use of median value (subset 5) leads to value range 0 to 225, while the power values normalised with median of peak power are in range 0 to 2. In case of normalisation with maximum values, the power dimension has larger impact due to larger spread of values. Ramping times are mostly near 0 if the maximum value is used for normalisation, thus the ramping time differences are mostly smaller than the differences of normalised power values.

D. Clustering Based on Ramp Energy

During the rise ramp, additional power, with maximum value $P_{peak} - P_{min1}$, in comparison to the initial power P_{min1} is generated. Using the ramp time t_{rise}, it is possible to derive equation for rise ramp energy U_{rise} (6). Similarly, during the fall ramp, the wind turbine power output increases up to

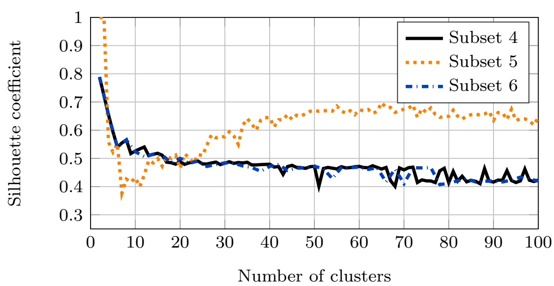


Fig. 7. Subset 4, 5 & 6: silhouette coefficient.

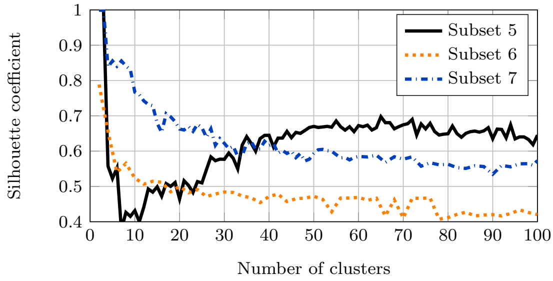


Fig. 8. Subset 5, 6 & 7: silhouette coefficient.

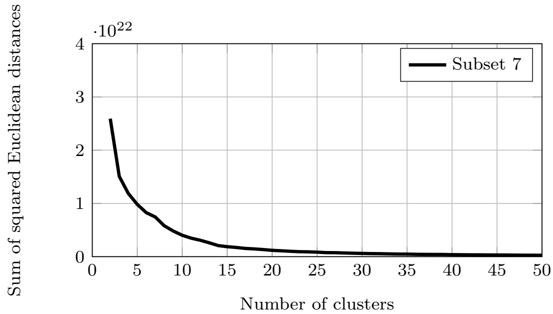


Fig. 9. Subset 7: sum of squared Euclidean distances.

$P_{min2} - P_{peak}$, compared to the value of P_{peak}. The energy of the fall ramp can be described by (7).

$$U_{rise} = \frac{t_{rise} \cdot (P_{peak} - P_{min1})}{2} \tag{6}$$

$$U_{fall} = \frac{t_{rise} \cdot (P_{min2} - P_{peak})}{2} \tag{7}$$

Elements of subset 7 were calculated using (6) and (7). The clustering results in Fig. 8 indicate that a strong structure can be acquired with up to 17 clusters and with higher number of clusters, the found structures have a reasonable strength.

IV. Conclusion

In this paper GA is applied to aggregate and shrink the original data volume. Ramping behaviour analysis (RBA) is used to identify the ramp events in the wind power data from wind farm. The pseudo-code of RBA is presented with individual functionalities. Optimal weight is allocated to RBA attributes using optimal cluster formulation. Clustering of identified ramp events based on different combinations of RBA parameters was conducted.

It was shown that in case of the analysed data, the strongest structure was found using the subset of ramping times. The clustering of RBA parameters with different units is a challenging task due to the need of parameter normalisation. It was shown in the paper, that one solution is to cluster the wind ramps based on rise and fall energy.

References

[1] C. Kamath, "Understanding wind ramp events through analysis of historical data," in *IEEE PES Transmission and Distribution Conference and Exposition 2010*, New Orleans, LA, USA, 2010, pp. 1-6.

[2] G. Martínez-Lucas, J. I. Sarasúa, and J. Á. Sánchez-Fernández, "Frequency Regulation of a Hybrid Wind–Hydro Power Plant in an Isolated Power System," *Energies*, vol. 11 no. 1, pp. 239, 2018.

[3] S. Mishra, M. Leinakse, and I. Palu, "Wind power variation identification using ramping behavior analysis," *Energy Procedia*, vol. 141, pp. 565-571, 2017.

[4] S. Mishra and I. Palu, "A user interface tool for ramping behavior analysis of renewable energy," in *Electric Power Quality and Supply Reliability (PQ) 2016*, Tallinn, Estonia, 2016, pp. 175-178.

[5] C. R. Dohrmann, H. R. Busby, and D. M. Trujillo, "Smoothing noisy data using dynamic programming and generalized cross-validation," *Journal of biomechanical engineering*, vol. 110, no. 1, pp. 37-41, 1988.

[6] M. A. Chowdhury, N. Hosseinzadeh, and W. X. Shen, "Smoothing wind power fluctuations by fuzzy logic pitch angle controller," *Renewable Energy*, vol. 38, no. 1, pp. 224-233, 2012.

[7] G. Grigoraş, F. Scarlatache, and G. Cârţină, "Load estimation for distribution systems using clustering techniques," in *13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, Brasov, Romania, 2012, pp. 301-306.

[8] P. J. Rousseeuw, "Silhouettes: a Graphical Aid to the Interpretation and Validation Cluster Analysis," *Journal of Computational Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[9] L. Kaufman and P. J. Rouseeuw, Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, NJ: John Wiley & Sons, Inc., 1990.

[10] D. E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Boston, MA, USA: Addison-Wesley Professional, 1989.

[11] M. Mitchell, An introduction to genetic algorithms, Massachusetts, MA, USA: MIT press, 1998.

[12] B. Coppin, Artificial intelligence illuminated, Sudbury, MA, USA: Jones & Bartlett Learning, 2004.

[13] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257-271, 1999.

[14] G. Mitsuo and R. Cheng, Genetic Algorithms and Engineering Design, New York City, NY, USA: John Wiley & Sons, 1997.

[15] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, pp. 129–137, 1982.

Appendix 3

S. Mishra, C. Würsig and I. Palu "Multivariate Scenario Generation - an ARIMA and Copula Approach, International Journal of Modeling and Optimization, 2018

Multivariate scenario generation

an ARIMA and copula approach

Sambeet Mishra[1,*], *Christoph* Würsig[2,**], and *Ivo* Palu[1]

[1]Department of Electrical Power Engineering and mechatronices, Tallinn University of Technology, Tallinn, Estonia
[2]Institute for Financial Market Theory, Leibniz Universität Hannover, Germany

Abstract. In mathematical optimization uncertainty is expressed through scenarios. auto-regressive integrated moving average (ARIMA) is one of the known practise to generate scenarios. This paper is about scenario generation using multivariate data: electrical power demand, wind power generation and energy market price. An ARIMA model along with Copula is implemented for scenario generation. The results are presented and discussed.

1 Introduction

This work is about scenario generation using copula. Scenario generation is an important part of stochastic programming. The generated scenarios however should retain the original statistical properties of the data. Auto regressive inter grated moving average (ARIMA) has been used extensively in literature [1] to generate scenarios. One of the drawbacks of ARIMA is the applicability to multivariate distributions. To overcome this Copula is used to generate scenarios as presented here [2, 3]. A regular vine copula and the goodness of fit measures are discussed here [4]. A Bayes theory based copula is presented here. [5]. A comprehensive study of various copula models with real world data is presented here [6]. A multivariate copula based forecasting method is explained here [7].

Multivariate copula is gaining more importance due to the nature and availability of data and relations among them. In this paper wind, demand and price data are considered as the multivariate data. Copula is used to generate multivariate distributions. These are sampled using ARIMA and the results are presented. The rest of the paper is organized in four sections: scenario generation, computational experiments, discussion and conclusion.

2 scenario generation

This section describes the mathematical model for the scenario generation using copula. This section is further divided into two subsections: ARIMA model and copula. The former presents a multivariate ARIMA formulation considering three variables. The later states the copula to sample the residuals.

[*]e-mail: sambeet.mishra@ttu.ee
[**]e-mail: wuersig@fmt.uni-hannover.de

2.1 ARIMA model

The ARIMA Model is a widely used model for modeling [8]. We use the model, to capture the time series behavior of the series.

The authors consider the statistically correlated scenarios because the stochastic variables: wind, demand, and price are co-related. Thus formulating the ARIMA(ϕ, φ) a quasi-contemporaneous stochastic process price ($y^a_{s,t}$), demand ($y^b_{s,t}$) and wind ($y^c_{s,t}$) as in 1(a-c). The residuals $\varepsilon^a_{t,s}, \varepsilon^b_{t,s}, \varepsilon^c_{t,s}$ are statistically dependent. Thus the dependency structure of the stochastic processes can be stated as $\varepsilon\{\varepsilon^a_{t,s} \cdot \varepsilon^b_{t-j,s} \cdot \varepsilon^c_{t-j,s}\} \neq 0$. $\varepsilon^a_s, \varepsilon^b_s, \varepsilon^c_s$ are the series of errors simulated to produce residual cross-correlogram of stochastic process. In 1(d) the error correlation between stochastic process a & b, a & c are presented and finally reduced to a product of an orthogonal matrix B and identity matrix $\psi(E[\psi \cdot \psi^T] = I)$. The cross correlation between $\varepsilon^a_{t,s}$ and $\varepsilon^b_{t,s}$ can be represented through variance-covariance matrix G. G is essentially a positive semi-definite and symmetric matrix. This matrix is further decomposed using Cholesky decomposition ($G = LL^T$) [9–11]. L is the upper triangular matrix that is also the orthogonal matrix ($B = L$).

$$y^a_{s,t} = \sum_{j=1}^{\eta^a} \phi^a_j \cdot y^a_{t-j,s} + \varepsilon^a_{s,t} - \sum_{j=1}^{\tau^a} \varphi^a_j \cdot \varepsilon^a_{t-j,s} \tag{1a}$$

$$y^b_{s,t} = \sum_{j=1}^{\eta^b} \phi^b_j \cdot y^b_{t-j,s} + \varepsilon^b_{s,t} - \sum_{j=1}^{\tau^b} \varphi^b_j \cdot \varepsilon^b_{t-j,s} \tag{1b}$$

$$y^c_{s,t} = \sum_{j=1}^{\eta^c} \phi^c_j \cdot y^c_{t-j,s} + \varepsilon^c_{s,t} - \sum_{j=1}^{\tau^c} \varphi^c_j \cdot \varepsilon^c_{t-j,s} \tag{1c}$$

$$\varepsilon^1_{s,t} = \begin{pmatrix} \varepsilon^a_{s,t} \\ \varepsilon^b_{s,t} \end{pmatrix} \quad \varepsilon^2_{s,t} = \begin{pmatrix} \varepsilon^a_{s,t} \\ \varepsilon^c_{s,t} \end{pmatrix} \implies \varepsilon = \begin{pmatrix} \varepsilon^1_{s,t} \\ \varepsilon^2_{s,t} \end{pmatrix} \implies \varepsilon = B\psi \tag{1d}$$

$$G = cov(\varepsilon, \varepsilon^T) = BB^T \tag{1e}$$

$$G = LL^T = BB^T \tag{1f}$$

2.2 Copula

The residuals of the ARIMA Model are fitted to a Copula Model in order to capture time varying dependence of the data. The authors use for this purpose R-vine copulas introduced by Bedford and Cooke (2001b,2002).

The general theory for copulas is Skalars Theorem (1959), based on this Theorem, Skalar shows that a every multivariate distribution can be written as a multivariate copula function. Equation (2) shows Skalars Theorem applied to a three dimensional dataset.

Variables with joint density function:

$$f(a,b,c) = f(a) \cdot f(b|a) \cdot f(c|b,a) \cdot ... \cdot f(a|b,c) \tag{2}$$

Following Skalar (1959) this density function is uniquely represented by the following form, if it is continuous.

$$F(a,b,c) = C(F_a(a), F_b(b), F_c(c)) \tag{3}$$

Joe (1996) makes this theorem usable for Vine Copulas, since he showed that Skalars Theorem can be decomposed to bivariate copulas. For a multivariate distribution with three variables it thus follows that this decomposition can uniquely identify the density function.

$$f(a|b,c) = c_{ac|b}(F_{a|b}(a|b), F_{c|b}(c|b)) \cdot f(a|b) \tag{4}$$

,where

$$f(a|b) = c_{ab} \cdot (F_a(a), F_b(b)) \cdot f_a(a)$$

The R-vine (regular vine) model is chosen to model the multivariate dependence in this empirical application. Fitting multivariate data to a copula is a challenging task, since commonly used copula models, like the normal copula, the t copula or the gumbel copula are either symmetric or have only one parameter to estimate the entire copula, which decreases the flexibility of the distribution. Bivariate copulas have a wider variety of choices, thus Kurowicka and Cooke (2006) developed the R-vine copula models that fit multiple bivariate copulas to the multivariate dataset and are thus able to capture the dependence structure of the multivariate dataset. "The modeling scheme is based on a decomposition of a multivariate density into a cascade of pair copulae" (Aas et al. p.1). R-vine's are represented by a hierarchical tree structure, where the first tree is estimated by n-1 bivariate copulae and the second by n-2 conditional on a single variable. For a three dimensional dataset two copulae need to be estimated directly and one conditional copula. In order to estimate the R-vine, Dissmann et al. (2012) developed a sequential search approach, they first estimate the family and parameters of the first tree via the AIC criterion. Then they use this result to estimate the second tree. Additionally they employ a maximum spanning tree algorithm to choose an appropriate edge weight. This paper implements their method and estimation technique, in order to take advantage of the benefits of the diversity of bivariate copulae.

2.3 ARIMA forecasting using Copula

The approach used in this paper is reminiscent of the Copula GARCH model, introduced by Jondenau and Rockinger (2006). First the ARIMA model is estimated, with the standardized residuals of the ARIMA model the R-vine copula model is estimated. The R-vine Model is then estimated using the remaining errors terms from the ARIMA model to capture dependencies between the variables that the time series model ARIMA cannot capture. The Copula model is fitted to uniform [0,1] margins. Afterwards following Dissmann et al. (2012) we simulate from the copula model and transform the thereby obtained data using the not standardized residuals from the empirical ARIMA model as an empirical density function. To model the time series behavior, the simulation result is obtained using the sampled residuals and the fitted ARIMA model.

3 Implementation

In the following we present the implementation of our method, this simulation is conducted on the logarithm of wind, price and demand variables for 100 times. The scripts are written in *R* programming language.

We estimate missing data, via linear interpolation for single missing values. For wind we estimate the last month via an ARIMA forecast due to the unaccounted data for December. The ARIMA model is fitted based on the conditional sum of squares to find the starting values. Following that maximum likelihood to find the optimal parameter estimates with respect to the AIC criterion.

We use the residuals and standardize them in order to fit an R-Vine Copula onto the residuals. The tree structure is determined via pair-copula families and estimated sequentially. For the model families the AIC criterion is used, parameter values are estimated using maximum likelihood estimation.

Following [12] methods we simulate the uniform estimates from the R-Vine Copula model.

We transform the uniform values using the trimmed empirical quantile distribution of our residuals into simulated observations.

We enter the simulated estimates into the ARIMA model and obtain the results after taking the exponential function of the values.

4 Computational experiments

The provided sample is hourly data for the year 2017, with the Price in €/MWh, Wind in MWH and Demand in MWH. The data contains two missing observations, they are interpolated, additionally the last 263 observations for Wind data are missing, in order to model this data an ARIMA model is fitted on the observed sample and the 263 missing values are estimated. The approach used is close to the GARCH Copula estimation, in place for a ARMA(p,q)-GARCH model and ARIMA model is used, since the data is unlikely heteroskedastic and it is unnecessary to model GARCH effects for this time series. This method enables us to fit the a copula approach easily to the data and to model the time series behavior.

First the data is fitted to an ARIMA model, that is optimally chosen based on the AIC criterion. The ARIMA process is required to be stationary and seasonal, this is necessary because of the limited amount of data, we are forecasting a year using only a year of data, trends cannot be captured reliably. It might be a substantial increase in wind production, but it is not clear if it is due to a windy year or additional wind farms, that would increase next year's production as well. The seasonality is assumed because of the nature of the data, wind is seasonal, as well as the demand, the price is seasonal as well. In order to ensure positivity of the data, we are fitting the natural logarithm of the data and transform them for analysis later on. In order to minimize extreme observations in our data set, considering the large time frame we are trying to model, we trim the residuals at 3% (we remove the 3% lowest and the 3% highest values). With this value we have a near normal kurtosis, before the kurtosis for the price and the wind reached over 40. In order to ensure that our results remain robust for different cutoff values, we used multiple values, the results are not inconsistent, the variation of the data increases as expected.

The estimated coefficients of the ARIMA model are presented in table (2), the standard errors for the coefficients are low and the model fit seems to be reasonable. In order to model serial dependence the innovations need to be modeled, in order to model them we are standardizing the residuals and transforming them into uniform [0,1] margins. The best R-Vine copula model is chosen by optimizing the bi-variate copula models and choosing the best fit with the AIC criterion. We sample the residuals from the trimmed series, we draw them based on their assigned uniform [0,1] margins provided by the random sampling from the copula. In the next step we find the best R-Vine Copula using maximum likelihood estimation and the AIC criterion. Simulations are conducted from this R-Vine structure. The result are uniform [0,1] simulation results of correlated seasonal innovations for wind, price and demand. To transform the uniform margins into realistic values, we use the quantiles of the trimmed residual series.

Using this series and the ARIMA model the simulation is conducted using the simulated innovations. The exponent of this result is combined with the new series to generate the plots (a)-(c), left from the red line is the original series and right from it the simulated series. The model clearly outperforms an ARIMA model with standard normal errors, that is not capturing any correlation between demand, wind and price, that the copula innovations are able to capture.

In table two the estimated ARIMA coefficients are shown, the best ARIMA Model is chosen according to its Aikaike estimation criteria. The model is assumed to be seasonal and we allow for models with non-zero mean. In order to achieve a positive simulation, we add the absolute minimum to the series, this does not change the character of the time series modeled but ensures consistent positive values.

5 Discussion

Table one shows the kendall correlation of the empirical sample. Demand and price is positively correlated as well as demand and wind, we see a small negative relationship between wind and price, likely because the wind barely has influence on the price, outside of extremely windy circumstances. From the correlations themselves we cannot make conclusions about the endogeneity. Surprising is the large correlation between demand and wind and the lack thereof in terms of prices. But maybe when it is windy it is more likely cloudy, thus more energy is consumed for heat and light.

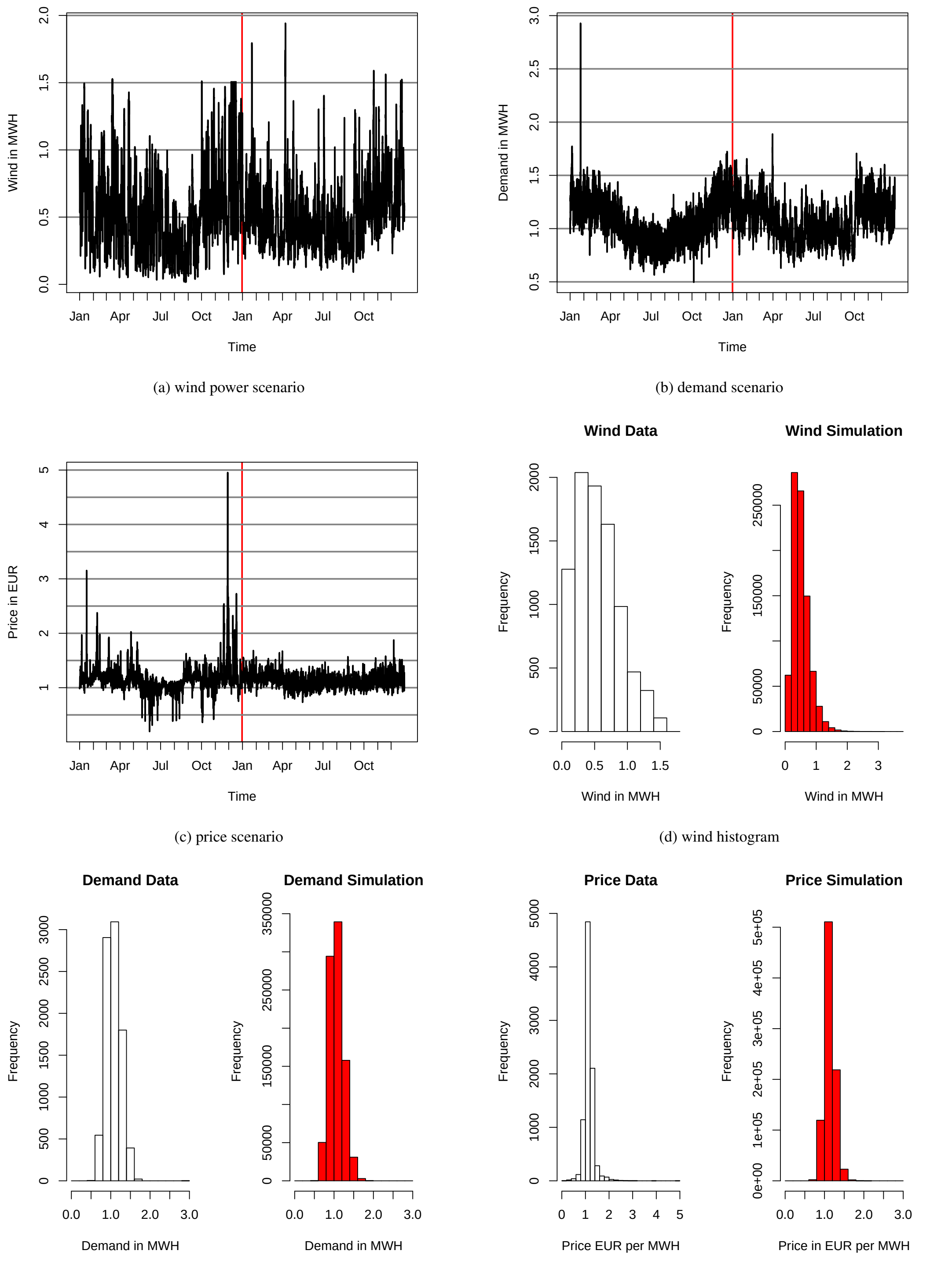
The proposed model with copulas can model dependencies, this benefit can be seen in table one, this table displays the range of the kendall correlation for all simulations. The range is wide but it is reasonably close to the sample and is capturing a large portion of the observed sample correlation. The coefficients cannot be the same, because there is likely a higher correlation for extreme observations, which we omit for the simulation in order o receive more realistic simulations.

Table 1: Data correlation

correlation sample			
	Demand	Price	Wind
Demand	1		
Price	0.4	1	
Wind	0.2	-0.07	1
correlation simulation			
Demand	1		
Price	0.14-0.31	1	
Wind	0.08-0.30	0-0.14	1

Table two shows the estimated ARIMA coefficients, since we required the model to be stationary, a mean is always estimated. This is reasonable here, because we attempt to forecast a year of data, because we just have a sample of one year length, assuming there is a trend in the wind production would be likely overfitting the model in sample. The model is fitted on logarithms, in order to ensure positive values after the simulation. Below the values the standard errors are displayed.

The fig. 1 shows each time series, on the left side of the vertical line is the original time series, on the right side the simulation. The time series is standardized to 1 and for the simulation we trim the values at 3%, this reduces the kurtosis of the residuals substantially and thus produces more reliable simulations over such a long time frame. We tried different ranges and it produces still reasonable results. The histograms display that the sample properties are conserved, we can see more outliers, because we have more observations in 100 simulations. The



(a) wind power scenario

(b) demand scenario

(c) price scenario

(d) wind histogram

(e) demand histogram

(f) price histogram

Figure 1: Original data and generated scenarios for (a) wind power (MW/H) (b) demand (MW/H) (c) price (€) followed by subsequent distributions

Table 2: ARIMA coefficients

ARIMA coefficient estimates Wind					
ar1	ar2	ma1			mean
1.5544	-0.5705	0.1052			7.3758
0.0137	0.0136	0.0165			0.0558
ARIMA coefficient estimates Demand					
ar1	ar2	ma1			mean
1.1107	-0.1455	-0.3716			6.9896
0.0681	0.0637	0.0652			0.0128
ARIMA estimates Price					
ar1	ar2	ma1	ma2	ma3	mean
0.3117	0.493	1.1201	0.5782	0.2114	3.3671
0.0693	0.0624	0.0689	0.0381	0.0149	0.0073

histograms show that the distribution of the year in sample and the simulations is reasonably close.

The model is able to capture correlation structures in the data that traditional approaches, like an ARIMA simulation with standard normal errors are not able to capture.

6 Conclusion

In this paper multivariate scenario generation based on three variables: demand, wind and price is presented. In the proposed multivariate scenario generation technique ARIMA is used for forecasting and copula for adjusting the residuals. The tail adjustment of the distribution and the impact is also discussed. In future works a comparative analysis of different statistical scenario generation technique for multivariate data would be conducted.

References

[1] P. Chen, T. Pedersen, B. Bak-Jensen, Z. Chen, IEEE transactions on power systems **25**, 667 (2010)

[2] P. Schütz, A. Tomasgard, S. Ahmed, European Journal of Operational Research **199**, 409 (2009)

[3] M. Kaut, Computational Management Science **11**, 503 (2014)

[4] U. Schepsmeier, Econometric Reviews pp. 1–22 (2016)

[5] G. Elidan, *Copula bayesian networks*, in *Advances in neural information processing systems* (2010), pp. 559–567

[6] T. Nagler, C. Schellhase, C. Czado, Dependence Modeling **5**, 99 (2017)

[7] C. Simard, B. Rémillard, Dependence Modeling **3** (2015)

[8] H. Hoeltgebaum, A. Street, C. Fernandes, IEEE Transactions on Power Systems (2018)

[9] M. Haugh, Monte Carlo Simulation: IEOR EA703 (2004)

[10] A. Schwarzenberg-Czerny, Astronomy and Astrophysics Supplement Series **110**, 405 (1995)

[11] D. Hakkarinen, Z. Chen, *Algorithmic Cholesky factorization fault recovery*, in *Parallel & Distributed Processing (IPDPS), 2010 IEEE International Symposium on* (IEEE, 2010), pp. 1–10

[12] J. Dissmann, E.C. Brechmann, C. Czado, D. Kurowicka, Computational Statistics & Data Analysis **59**, 52 (2013)

Appendix 4

S. Mishra, C. Bordin,J. Fornes, and I. Palu "Reliability Framework for Power Network Assessment" E3S Web of Conferences, 2018

Reliability framework for power network assessment

Sambeet Mishra[1,*], *Chiara* Bordin[2], *Jordi Mateo* Fornes[3,**], and *Ivo* Palu[1]

[1]Tallinn University of Technology, Estonia
[2]SINTEF Energy research
[3]Department of Computer Science, University of Lleida, Jaume II 69, 25001 Lleida, Spain.

Abstract. Reliability of power system in terms of investments in network maintenance and restructuring for power distribution network has gained importance due to increase in distributed generation. To determine the reliability of the power distribution network, the state of power apparatus, losses in the network and consumer satisfaction indices are key factors. Considering the aforementioned, this paper proposes a holistic reliability framework for power distribution networks. The framework lists the following factors: life cycle of power apparatus, environmental and sociological, node reliability, arc reliability. A case study for reliability evaluation is performed on a modified IEEE 14 bus network. Furthermore, multiple scenarios of generation fault or outage are studied and results are presented. The key contribution of this paper is to present a novel and holistic reliability framework to model distribution network.

1 Introduction

The aim of this paper is to develop an evaluation framework for reconfiguration of an electrical power system. Given study is influenced by the economical aspect of a network operator which implies to have the needed electrical, environmental, and economical indicators of an electrical power system to assess the optimal possibilities for reconstructing it. Many power system management utilities try to rationalize their network and optimize the total life-cycle costs of the components, due to the regulation of power quality and the reliability issues. For the given problem, many methods of assessment modeling have been introduced in the literature. System well-being method for power networks adequacy assessment using monte carlo simulation is presented, specifically for capacity reserve allocation [1]. This study expanded with renewable energy resources in [2]. This subject is further explored in the book series about reliable and sustainable power systems management [3]. Role of outage management and strategies are discussed here [4]. Reliability evaluation of power systems establishes measures to identify power interruptions and their implications [5]. However the approach is limited to power system events. For instance the state of the power apparatus is an important indicator to observe the system status. Detail concepts of power system reliability are presented in [6]. Reliability and connected parameters are introduced and and event tree analysis for event prediction is described here [7].

Many regionally owned utilities have been privatized, where the main concern are optimal investment planning for expansion and maintenance while maintaining the system reliability. Therefore, to improve and maintain power quality various regulation models have been introduced. These models enforce network companies to optimize their operations and schedule their maintenance activities without having to compromise the reliability nor the safety of the power system network. Reliability analysis is one of the key ways to inspect optimal asset management. A study has been conducted in [8–10] for evaluating the failure rates and reliability modeling for power distribution network. The reliability analysis has been utilized for optimal dis-connector allocation. In the analysis, the failure rates are constant for similar components, which are influenced by many different mechanical, environmental and electrical stresses. Usually for component failure models reliability calculations are based on the exponent distribution and failure rates are considered to be constant, which may be an inadequate approach. In some scenarios a scalar value is used for estimating the component failure rates while planning. Monte Carlo simulations are implemented to take into account effects of the surrounding, or enhanced component failure models which are based on constant component failure rates, to evaluate component related and the environmental aspects in reliability analysis. Another modeling approach is proportional hazard method, where age and other various additional information have been considered. Further information may include weather and environmental factors that affect the components. These models require significant data to be analyzed to investigate the essential dependencies affecting the component failure and malfunction. At the same time, Markov models are more commonly used, where the component failure rates are modeled by estimating the effects of the component faults for the system. Moreover, there are many factors that may lead to component failure

*e-mail: sambeet.mishra@ttu.ee
**e-mail: jmateo@diei.udl.cat

such as weather condition and power surges. Therefore, predicting the component failure accurately is a challenge. As a consequence, the readiness of the power distribution network to optimally resolve the fault while ensuring uninterrupted power supply is the research objective.

1.1 Motivation for Reliability-based Network Decision-making

The main demand for reliability oriented network restructuring (RNR) is to have the estimates of failure rates which consider the main stress factors and the possibility to have first estimates from incomplete data to update values when more improved data is available. Components of distribution networks (DN) are modeled separately since the component failure rate is independent of factors that affect the power network. In this study, DN has been divided into four main components: overhead-lines (OHL), cables, transformers, and groups of switch-gear. For each component the main reasons for permanent faults and auto re-closings is determined. Separate failure rates for each component types are based on the reasons for a fault, e.g. aerial line overall failure rate is dependent on lightning, different weather conditions, and other fault related causes. The main stress factors that influence the failure rate, have been determined for all the reasons mentioned above. All the stress factors are classified into appropriate classes, for instance, the location could be by the seaside, near the forest or on a field. For all these classes a weight has been defined to represent the effects of a certain class to failure rate. For a total failure rate, temporary and permanent faults can be calculated. A practical approach on component modeling is to use the idea that parameters used in failure rate modeling should be possible to be affected by the selected planning strategies. Therefore, effects of the weather are included as an own parameter to failure rate modeling, however they are still included in condition information of the affect to the stress tolerance. Furthermore, the effects of component aging can be explained with other component related factors, therefore, life expectancy is included in condition weight information. Failure rate parameters must be determined before modeling methods can be used [11].

General failure rate of components were calculated as a weighted mean from failure rates of separate companies. Defined parameter groups are used to calculate the separate failure rates. Basic input data is the component information, i.e. component type, failure rate, and the network topology. Additionally, other information are needed regarding those factors that affect the analyses results, such as maintenance costs. In the enhanced radial reliability analysis, the network is with feeders and zones, which refer to a part of a feeder. In the given analysis, expected amount of permanent and temporary failures in a zone is calculated as an entirety of the individual network component failures. For a temporary fault, the whole feeder is experiencing the same short interruption. In the given analysis, experienced faults are defined for each load point. Determination of maintenance costs is done by analyzing the possible terrains where the components are located related to the total interruptions in the certain area. Therefore, the RNR framework can be expressed as an asset management model considering the Life-Cycle Assessment of power system components. To replace OHL's and underground cables, reliability of its reconfiguration is based on environmental and consumer preferences, N-1 criteria, and the objectives for minimizing the investment costs. Reconfiguration of networks is primarily done to accommodate new consumers, which is achieved by extending the existing node through a new arc, and to replace some existing out-of-date lines. Network operators can adjust the failure rate and reliability parameters with their own network data. Smart grid components can identify the fault region of the feeder and update it with secure supply of energy from the same power network. Reliability indicators mainly include measures of outage duration and its frequency, the amount of power or energy not supplied, and the number of customers involved in outages. These indicators are determined over predefined period of time, such as SAIFI, SAIDI, CAIDI, ENS.

The rest of the paper is organized as the following sections: reliability framework, power network weight evaluation. In reliability framework each element of the framework concerning the power network is presented. An example case of IEEE 14-bus network is presented and discussed. Subsequently we evaluate generation loss scenarios and the results are presented in the power network evaluation section.

2 Power System Reliability and Adequacy

Power system adequacy refers to the condition of a power network considering generation, transmission and distribution units. Power system reliability refers to the state of network to sustain flow of energy from point of generation to demand at any point in time. The relationship between reliability and investment cost is presented in fig. 1. The figure signifies with outages or otherwise power interruptions the reliability reduces and the investment increases. Therefore, the planner need to pay attention so as to optimally plan for the expansions or maintenance of the distribution network. The indicators to determine the condition of power system can be broadly classified in to five categories: life cycle of power apparatus, environmental and sociological, node reliability, arc reliability and node reliability concerning losses. The table 1 outlines the associated indicators. These parameters organize the indicators taking into account the network as a node-arc formulation.

2.1 Life cycle assessment (LCA)

Life cycle assessment (LCA) is a method to determine the environmental impacts from a product, a process, or an activity. It is also used to assess the remaining utilization life. Throughout the product lifetime the impacts mainly originate from the power losses during the use phase, although installation, maintenance, and dismantling also contribute

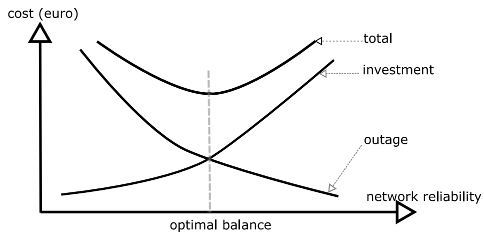


Figure 1: Relationship between reliability and cost

Table 1: Power system reliability and adequacy indicators

LCA	Environmental & Sociological	Node reliability indicators
Cable	Terrain	SAIFI value
OHL	Weather	SAIDI value
Transformer	Consumption	ENS value
Switchgear	Generation	Availability
Emissions		
Arc reliability indicators	**Node reliability indicators concerning losses**	
AIT	LOLE	
AIF	LOEE	
AID	EIR	

to it. Transmission and distribution assets have been comprised of power lines, cables, transformers, substations, and other electrical components to generate wide range of environmental impacts, such as the equipment emissions and material weight value. The life cycle stages viewed are interpreted as the production or the manufacturing phase of a product or its use phase. The used life cycle inventory consists of material requirement for grid components and their environmental impact. For all the components, the functional unit is one equipment operational during the lifetime.

Given network is a high and medium voltage network, with three distributed generation inputs. Biomass and oil shale produce the distributed generated energy, with an additional input from a submarine cable. The network nodes are depicted as substations with each bus-line having an ideal energy consumer as residential or commercial. The transmission line is either an underground cable or an overhead-line. The normalized weights are allocated to arcs and nodes based on the historical data and expert's opinion. For instance, an indicator depicting transmission line weight is valued at 4000.5 kg/km, although Switchgear emissions are valued at 185.38 kg CO_2 per transformer. Weights are normalized and translated to the cost of maintenance for the distribution system operator. In this method, dissimilar components can be compared based on the cost of investment.

2.1.1 Cable

Medium voltage power cables are characterized in [12]. In this study we concentrate on three different voltage levels, with each having one or two different types of cables, therefore five different cables are evaluated. Cables are chosen based on their suitable voltage level, cable diameter and the conductor and the insulation type.

A transmission line of 6 or 10 kV spanning over 9 km, a three-core cable is proposed with diameter of 3*70 mm^2. This cable weighs 5400 kg/km. Cable indicators in a network reconstruction evaluation are the conductor and insulation. In the example, conductor weight is calculated as the diameter of the conductor (i.e. Al conductor diameter = 9.8 mm) multiplied with conduction material density (Al density = 2.7 kg/km). i.e. Cable, which weighs 5400 kg/km has an aluminum conductor which weighs 124.7 kg/km and insulation of 5275.3 kg/km [13]. Different kinds of lines vary mainly in sizing of the cross section of the conductor, however the mentioned cross section is not the actual measurement of the wire, but rather the area of the conductor gives the needed cross section, such as aluminum, copper or steel. Data from cable producers is used to identify the correct amount of material input for each of the line types. For the power system network, at three different voltage levels, each voltage level is applied with power lines with the following conductors: aluminum (2.7 $\frac{gm}{cm^3}$), copper (8.89 $\frac{gm}{cm^3}$) and steel (7.83 $\frac{gm}{cm^3}$). The power line data is derived from manufacturer's data provided by ABB [14]. From the catalogue, the most suitable power lines are chosen, based on their voltage level and the conductor material. Considering the length of a given power line and the cross section provided by ABB, it is possible to find the weight of a conductor [14].

Given the total length of underground cables and overhead lines, and additionally it contains the calculated values of the power line conductor and insulation weight relative to their material. The insulation material composition is also derived from the manufacturer specifications [14]. Typically, the installation of cables requires extensive underground pathways. Due to insufficient data, these parameters are not included. However, some general construction processes are included, such as the weight difference and the relation to the actual line weight are applied. These electrical masts are calculated by their tension and sag related the length between each pole. Also, different height of masts is assumed for each voltage level. Poles are described by their suitable voltage level, material use, height, span, tension, and sag.

Table 2: Reliability assessment of power conductor

Voltage level (kV)	Cable length (km)	Cable diameter (mm^2)
6+10	9	3*70
Cable line weight (kg/km)	**Conductor diameter (mm)**	**Weight (kg/km)**
5400	9.8	646.5
Conductor material (Aluminum (Al))	**Insulation type**	**Weight (kg/km)**
Al	XLPE	5275.3

2.1.2 Transformer and switch-gear

Transformers used in this study are assumed to be ideal, without having to relate to the criterion N-1 (in case one

Node	1	2	3	4	5	6	7
Voltage level (kV)	110/10	110/35/10/6	110/35/6	110/35/6	110/10	35/6	110/35/6
Generation [0/1]	0	1	0	0	0	0	1
Generation, kWh	0	925	0	0	0	0	1000
Transformer [0,1,2]	1	2	1	1	1	1	1
Transformer type[1,2,3]	1	1+3	2	2	1	3	2
Life expectancy	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Conductor weight (kg/Trfo)	0.61	1.00	0.57	0.57	0.61	0.00	0.57
Conductor weight (kg/kVA)	1.00	1.00	0.001	0.001	1.00	0.00	0.001
Trfo_oil weight (kg/Trfo)	0.56	1.00	0.22	0.22	0.56	0.00	0.22
Trfo_oil weight (kg/kVA)	1.00	1.00	0.00	0.00	1.00	0.0002	0.00
Energy demand (kWh)	250	275	275	325	300	350	350
Energy demand	0.00	0.25	0.25	0.75	0.50	1.00	1.00
SG type	1	4	2	2	1	3	2
Life expectancy	0	0	0	0	0	0	0
GWP	0.14	1.00	0.02	0.02	0.14	0.00	0.02
AP	0.22	1.00	0.03	0.03	0.22	0.00	0.03
NP	0.44	1.00	0.05	0.05	0.44	0.00	0.05
SF6 (% of all emissions)	0.00	1.00	0.11	0.11	0.00	0.02	0.11
Total value:	0.15	1.00	0.05	0.05	0.15	0.01	0.05
Terrain	4.00	9.00	5.00	7.00	4.00	5.00	7.00
Terrain coefficients	0.38	1.00	0.50	0.75	0.38	0.50	0.75

Table 3: Reliability indicators and weights for modified IEEE 14 bus network [1/2]

transformer is interrupted, the power flow will continue). Therefore, a substation is only constructed with 1 transformer, with one exception of a substation with four different voltages. For that substation, the assumption of having two transformers gives a more effective distribution. The data for this table is provide by ABB [14]. Transformer weight is based on the conductor material (copper) weight in the wiring and the profile during the manufacturing phase. Furthermore, the transformers under evaluation are assumed to be almost equal to the ones from the manufacturer, e.g. for a 110/10 kV substation, 220/15.6 kV transformer is used [110 −220]. For a network substation, the different combinations of switch-gear are assumed based on their operating voltage level. These combinations are needed for environmental impact assessment, which will be talked about later on. Switch-gear values are based on the environmental impact per transformer.

2.2 Node reliability indicators

Given reliability indicators combine the factors related to outage duration or the response time, frequency of outage, number of customers involved in interruption or their lost power and energy. System average interruption frequency index (SAIFI), system average interruption duration index (SAIDI), energy not supplied (ENS), average service availability index (Availability), average interruption time (AIT), average interruption frequency (AIF), average interruption duration (AID) are among the measures to evaluate the interruptions and its impacts.

2.2.1 Terrain and weather effects

The failure rate of any component in a power system network is assumed to be influenced by different internal (quality and type of transmission line) and external factors (terrain, topology and weather conditions). Nine different terrain types are considered for this study that outlines the network environments. For example, the hill-like location has a lower failure rate compared to a commercial location, because in the commercial surrounding, there are more internal factors related to the consumption, and seaside has a lower failure rate than the forest, because forest is prone to fail e.g, trees falling. For the outage of a generation unit a maintenance cost incurs for the system operator. Factors like replacement or repair or maintenance cost for a corridor or line segment of the power network is considered. Additionally the environmental impact of an outage in terms of fuel consumption during the servicing and resulting environmental impact is also taken into consideration [15].

3 Evaluation of Power Network Condition under Outages

The fig. 2 presents the IEEE-14 bus network with zones. Each zone presents a geographical location. Tables 3 and 4 presents the indicators and the corresponding values (normalized) for measuring the system reliability based on nodes under the proposed framework.

The formulation for probability of failure Ω_n for bus n can be presented as in (1). In (2) frequency of failure Ω_n^f is presented [6].

$$\Omega_n = \sum_i [P(O_j)(P_{i,g}P_{l,i} - P_{g,i}P_{l,i})] \quad (1)$$

$$\Omega_n^f = \sum_i [O_j^f(P_{i,g}P_{l,i} - P_{g,i}P_{l,i})] \quad (2)$$

Nodes	8	9	10	11	12	13	14
Voltage level (kV)	110/35/6	110/10/6	110/35/10	110/10	110/10	110/35/6	35/6
Generation [0/1]	0.00	0.00	1.00	1.00	0.00	0.00	0.00
Generation, kWh	0.00	0.00	2000.00	1000.00	0.00	0.00	0.00
Transformer [0,1,2]	1.00	1.00	2.00	1.00	1.00	1.00	1.00
Transformer type[1,2,3]	2.00	2.00	1+3	1.00	1.00	2.00	3.00
Life expectancy	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Conductor weight (kg/Trfo)	0.57	0.57	1.00	0.61	0.61	0.57	0.00
Conductor weight (kg/kVA)	0.00	0.00	1.00	1.00	1.00	0.00	0.00
Trfo_oil weight (kg/Trfo)	0.22	0.22	1.00	0.56	0.56	0.22	0.00
Trfo_oil weight (kg/kVA)	0.00	0.00	1.00	1.00	1.00	0.00	0.00
Energy demand (kWh)	325.00	350.00	275.00	300.00	300.00	325.00	300.00
Energy demand	0.75	1.00	0.25	0.50	0.50	0.75	0.50
Switch-gear type	2.00	2.00	4.00	1.00	1.00	2.00	3.00
Life expectancy	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GWP	0.02	0.02	1.00	0.14	0.14	0.02	0.00
AP	0.03	0.03	1.00	0.22	0.22	0.03	0.00
NP	0.05	0.05	1.00	0.44	0.44	0.05	0.00
SF6 (% of all emissions)	0.11	0.11	1.00	0.00	0.00	0.11	0.02
Total value:	0.05	0.05	1.00	0.15	0.15	0.05	0.01
Terrain	7.00	5.00	9.00	5.00	1.00	4.00	5.00
Terrain coefficients	0.75	0.50	1.00	0.50	1.00	0.38	0.50

Table 4: Reliability indicators and values for modified IEEE 14 bus network [2/2]

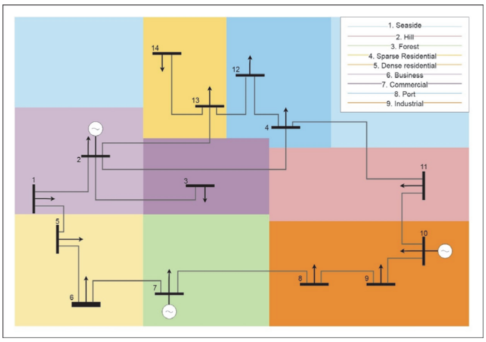


Figure 2: Classification of modified IEEE-14 bus network into zones based on weights

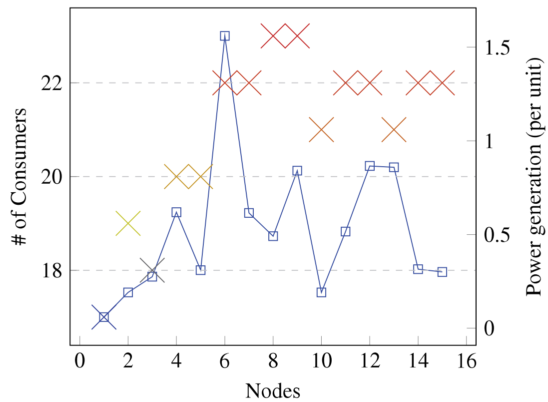


Figure 3: Node wise number of consumers with respective consumption in per unit

Where O_j is the condition of outage in the power transmission network. $P_{i,g}$ is the probability of occurrence of capacity outage beyond reserves. And probability of uninterrupted power supply. The availability (γ) is calculated as $\gamma = 1 - \frac{60*ENS}{\sum_i P_i}$ where P_i is average power supplied by the total system and ENS (Energy not supplied because of interruption) and P_i stands for power interruption for incident i. The total cost is product of component capital cost times availability. And the repair cost is the cost of repair times the total cost. Similarly the maintenance cost is value of fault times the repair costs.

Considering the four generation units 15 generation outage states are assessed. Figure 3 demonstrates the network topology and number of consumers with respective per unit energy consumption. There are 314 numbers of consumers in the 14 bus network with total consumption of 8.008. Table 5 states the power system indicators for evaluating the faults and losses with respect to line for interrupted load. The sum total SAIFI, SAIDI, ENS for the network are 0.00426, 8.925 and 0.196 respectively. The relative cost in € for investment, repair and maintenance are presented in table 6. The net overall investment is 29.29215 followed by 0.124837 for repair and 0.0005320 for maintenance cost.

Line	**Interrupted load**	**SAIFI**	**SAIDI**	**ENS**	**Availability**	**CAIDI**	**CAIFI**
1	0.0002514	0.000231	0.48	0.001442673	0.999999849	0.226468312	0.000230734
2	0.0008140	0.000258	0.54	0.004670347	0.999999831	0.253111642	0.000257879
3	0.0011720	0.000244	0.51	0.006724321	0.99999984	0.239789977	0.000244307
4	0.0026381	0.000271	0.57	0.015135836	0.999999822	0.266433308	0.000271452
5	0.0013212	0.000271	0.57	0.007580144	0.999999822	0.266433308	0.000271452
6	0.0066484	0.000299	0.63	0.038145242	0.999999804	0.293076639	0.000298597
7	0.0026210	0.000299	0.63	0.015038028	0.999999804	0.293076639	0.000298597
8	0.0020925	0.000312	0.65	0.012005970	0.999999796	0.306398304	0.00031217
9	0.0035842	0.000312	0.65	0.020564198	0.999999796	0.306398304	0.00031217
10	0.0008140	0.000285	0.60	0.004670347	0.999999813	0.279754973	0.000285025
11	0.0021991	0.000299	0.63	0.012617272	0.999999804	0.293076639	0.000298597
12	0.0036907	0.000299	0.63	0.021175500	0.999999804	0.293076639	0.000298597
13	0.0036609	0.000285	0.60	0.021004335	0.999999813	0.279754973	0.000285025
14	0.0013425	0.000299	0.63	0.007702405	0.999999804	0.293076639	0.000298597
15	0.0012785	0.000299	0.63	0.007335623	0.999999804	0.293076639	0.000298597

Table 5: indicators regarding system interruptions and losses10^{-3}

Line	**Loss**			**Cost**		
	LOLE	*LOEE*	*EIR*	*Investment*	*Repair*	*Maintenance*
1	0.001051799	1.85256E-06	0.999998147	3.49999	0.014916260	0.000063570
2	0.003404977	1.94149E-05	0.999980585	3.49993	0.014915998	0.000063569
3	0.004902454	4.0247E-05	0.999959753	0.59998	0.002556975	0.000010897
4	0.011034979	0.000203915	0.999796085	0.59988	0.002556556	0.000010896
5	0.005526403	5.11437E-05	0.999948856	0.59997	0.002556947	0.000010897
6	0.027810287	0.001295143	0.998704857	3.49547	0.014896969	0.000063488
7	0.010963671	0.000201288	0.999798712	0.59988	0.002556563	0.000010896
8	0.00875311	0.000128301	0.999871699	3.49955	0.014914374	0.000063562
9	0.014992597	0.00037641	0.99962359	3.49868	0.014910673	0.000063546
10	0.003404977	1.94149E-05	0.999980585	0.59999	0.002557028	0.000010898
11	0.009198787	0.000141699	0.999858301	3.49950	0.014914174	0.000063561
12	0.015438275	0.000399121	0.999600879	0.59976	0.002556057	0.000010893
13	0.015313485	0.000392695	0.999607305	0.59976	0.002556074	0.000010893
14	0.005615539	5.28068E-05	0.999947193	0.59997	0.002556943	0.000010897
15	0.005348132	4.78973E-05	0.999952103	3.49983	0.014915573	0.000063567

Table 6: Relative costs for lines (€)

4 Conclusion

This study is about reliability of power distribution network. The presented reliability framework includes interdisciplinary aspects of life cycle analysis, consumer satisfaction index and power interruptions. These measures are summarized in terms of investment, repair and maintenance costs in the presented framework. The case study conducted on modified IEEE-14 bus network for generation outage scenarios demonstrate the relation between investments, repairs and maintenance costs to maintain the reliability. Apart from that, the system operator has an incentive in form of reducing direct investment to with marginal repair or maintenance cost. It is evident from the case study that the outages has an adverse impact on the system reliability and therefore results in additional cost. However, the additional cost can be shared with timely maintenance and repairs to avoid an up-front investment. Therefore, in an investment decision for network expansion, reliability analysis of the power network becomes significantly important to make an optimal decision. In future works, investment models with network restructuring will be studied. Application of classification techniques to evaluate weights for each edge in another research avenue, given that there are both empirical and quantitative information is included in the reliability criterion.

References

[1] R. Billinton, R. Karki, IEEE Transactions on Power systems **14**, 1172 (1999)

[2] A.M.L. da Silva, L.A.F. Manso, W.S. Sales, L.C. Resende, M.J. Aguiar, M.A. Matos, J.A.P. Lopes, V. Miranda, European Transactions on Electrical Power **17**, 387 (2007)

[3] R. Karki, R. Billinton, A.K. Verma, *Reliability Modeling and Analysis of Smart Power Systems* (Springer, 2014)

[4] H. Farzin, M. Fotuhi-Firuzabad, M. Moeini-Aghtaie, IEEE Transactions on Power Systems **33**, 2359 (2018)

[5] R.N. Allan et al., *Reliability evaluation of power systems* (Springer Science & Business Media, 2013)

[6] R. Billinton, W. Li, in *Reliability Assessment of Electric Power systems Using Monte Carlo Methods* (Springer, 1994), pp. 9–31

[7] A. Volkanovski, M. Čepin, B. Mavko, Reliability Engineering & System Safety **94**, 1116 (2009)

[8] J. Pylvänäinen, P. Verho, J. Järvinen, S. Kunttu, J. Sarsama, *Advanced failure rate and distribution network reliability modelling as part of network planning software*, in *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on* (IET, 2005), pp. 1–5

[9] P. Jarventausta, J. Partanen, S. Karkkainen (2008)

[10] J. Pylvanainen, J. Jarvinen, P. Verho, *Advanced reliability analysis for distribution network*, in *Electric Utility Deregulation, Restructuring and Power Technologies, 2004.(DRPT 2004). Proceedings of the 2004 IEEE International Conference on* (IEEE, 2004), Vol. 2, pp. 457–462

[11] W. He, X. Shen, J. Guo, Q. Wang, Y. Lu, H. Li, *An improved design of damped AC test system for partial discharge measurement in distribution power cables*, in *IOP Conference Series: Materials Science and Engineering* (IOP Publishing, 2018), Vol. 366, p. 012029

[12] A. Cataliotti, A. Daidone, G. Tinè, IEEE Transactions on Power Delivery **23**, 1896 (2008)

[13] E. Cables, https://www.elandcables.com/electrical-cable-and-accessories/cables-by-type/medium-voltage-cable (2018)

[14] A. Cables, XLPE Land Cable Systems (2018)

[15] N. Ekstedt, P. Hilber, *Categorization and review of failure rate factors used in power systems*, in *Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on* (IEEE, 2014), pp. 1–6

Appendix 5

S. Mishra, C. Bordin, and I. Palu "RNR: Reliability oriented Network Restructuring" International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 2018

RNR: Reliability oriented Network Restructuring

Sambeet Mishra[1], Chiara Bordin [2], Ivo Palu [1]
[1] Tallinn University of Technology, Tallinn, Estonia
[2]SINTEF Energy Research, Trondheim, Norway

***Abstract*—This paper is about an application of optimal power flow calculation for considering how interconnections of microgrids affect the reliability of the system and the need of network reconfiguration. For this purpose reliability indicators for power system restructuring are presented. A reliability oriented network restructuring (RNR) mathematical optimization model is proposed for solving power grid expansion decisions with non-linear AC-OPF. The microgrid structures are derived from the standard IEEE-14 bus system architecture. The proposed reliability framework is implemented with the set of reliability indicators for measuring the system performance. The model was solved using outer approximation algorithm. The analysis is conducted to investigate the importance of restructuring in an investment decision for the expansion. The results with a comparison between investment and investment with restructuring are outlined. Consequently, the expansion considering the restructuring is found to be practical and feasible.**

***Index Terms*—power system expansion, reliability and vulnerability, optimal power flow, microgrids, non-linear system of equations, outer approximation algorithm**

I. Introduction

THERE is an energy transition from "top-to-bottom" to "bottom-to-top" flow of energy. The conventional generator is at the top in the former and multiple renewable energy based generators are at the bottom in the latter. The increased share of renewable energy resources (RES) in the power generation mix is one of the primary reasons for the transition. With this transition the macro-grid is sub-divided in to multiple micro-grids with distributed and renewable energy technologies. However integrated, intermittent and distributed generations have increased the risk of security of supply as their utilization grows in distribution networks. Micro-grid (MG) is more sensitive to power quality issues when it is maintained on local resources. Voltage imbalance, voltage drops, between generation and load are serious issues which are caused by connection of single-phase loads and sources. The objective of the modern network operator is to employ the smart grid technologies to plan, operate and maintain a modern power system economically stable and with an acceptable level of reliability.

Optimal power flow (OPF) is a central operational tool for power systems. The direct current (DC) version is mostly used for the high-voltage networks for transmission of bulk power. The alternating current (AC) version is primarily used in case of distribution networks, especially in the distribution grid problems such as grid planning, optimal controls, reactive power dispatch and unit commitment. Primary objective of OPF is to maintain the system stability while minimizing the cost of operations and maintenance. Investment decisions considering power system constraints are closer to practical. OPF in its original form is a highly non-linear problem. The non-linearity of the system of equation is usually solved using iterative Gauss-Siedel or Newton-Rapson method. Moreover OPF is a non-convex optimization problem. It is also a NP-hard problem (see [1], [2]) to find a solution for radial networks. To solve such a problem literature suggests a) approximation- with relaxed physical properties of OPF b) non-linear optimization methods c) heuristics/meta-heuristics d) convexification. The power system mostly consists of radial networks. In literature a lot of work is done to linearise and relax the constraints. The models can be broadly classified as a) original OPF (O-ACOPF), b) augmented OPF (A-ACOPF), c) augmented-relaxed (AR-ACOPF) OPF [3]. In literature there are exact numerical solutions provided through distributed optimization based on alternative direction method of multipliers and semi-definite relaxations for radial and non-radial networks [3]–[8]. However, only numerical proofs for specific grids are portrayed in place of generalized exact proof of the relaxation of the problem. In this paper we use a feasible and near-optimal outer approximation algorithm to solve the non-li/near ACOPF problem. Keeping the physical properties of the power distribution system intact we focus on the OPF in IEEE 14 bus network with power injection of intermittent and non-dispatch able generations at multiple edges of the distribution network. The model considers the two-port pi network for the transmission line representation. The model is tested over three microgrids with an IEEE-14 bus radial network configuration.

Distribution system expansion or in this case MG expansion is often an optimal investment and operational decision. However most of the investment models lack power system aspect of expansion. Treated problem on a high level prospective. The contribution of this paper is to investigate how expansion decisions affect the reliability of the system, and therefore the importance of restructuring in power network expansion. A reliability oriented network restructuring (RNR) framework is presented.

II. Reliability oriented network decision-making

Due to power system regulation power quality and reliability issues, concerning many management businesses, many utilities try to rationalize their network and optimize the total life cycles costs of the components [9]. Many municipality

This article was partially funded by the European Union via the European Regional Development Fund.

owned utilities have been privatized, where the new owners are mainly considering profitable investments, therefore to avoid power quality's reduction, various regulation models have been issued. Given models enforce network utilities to optimize their operations without compromising the reliability nor the safety of the network. Reliability analysis are one of the ways to inspect the optimal asset management. Similar analysis has been developed in Tampere University of Technology, in the 1980's, where the reliability analysis has been utilized to evaluate optimal dis-connector locations. In the analysis [10], the failure rates are constant for similar components, where those components are influenced by many different mechanical, environmental and electrical stresses. Usually in component failure model's reliability calculations are based on exponent distribution and failure rates are considered as constants. However the constant failure rate is an inadequate approach, therefore many models for estimating component failure rates have to be used. In some cases, Monte Carlo simulation is utilized to take into account effects of the surrounding or enhanced component failure models, which are based on constant component failure rates to evaluate environmental and component related aspects in reliability analysis. Another modelling approach is done as a proportional hazard method, where it can consider age and various additional information, such as weather and the information surrounding the components. These models require lots of data to find essential dependencies affecting the component reliability to fail, therefore these models are not commonly used. Sometimes Markov Models are also used, where the component failure modelling is done by estimating the effects of the component faults for the system [11], [12]. Usually complex system models are needed, because there is a large amount of possible transitions needed for each component, such as for different weather conditions. Main requirement for RNR is to have estimates of failure rates considering main stress factors and the possibility to have first estimates from incomplete data and update values when more improved data are available. Components of distribution networks must be modelled separately, therefore component failure rate is dependent on different factors.

In this study the distribution network has been divided into five main components: aerial lines, cables, transformers and switches. For each component it has been determined the main reasons for permanent faults and auto re-closings. Separate failure rates for each component types are based on the failure reasons, e.g. transformers overall failure rate is dependent on lightning, animals and other fault causes. For all the reasons, the main stress factors which affect the failure rate have been determined. All the stress factors are classified into appropriate classes, for instance the location can be a forest a place near the road or a field. For all classes a weight has been defined, which represents the effect of a certain class to on the failure rate. For total failure rate, permanent and temporary faults can be calculated. A practical approach in component modelling is to use the idea that it should be possible to affect the parameters used in failure rate modelling, with selected planning strategies. The weather pattern is not considered directly in failure rate evaluation but included in the apparatus condition, for instance in the stress tolerance. The age factor is included in condition weight information. Voltage dip analysis is also used for examining short interruption, where each component is defined based on permanent and temporary short circuit failures. Dip rates are used to define number and depth of dips in the network. Voltage dip can be analyzed by adding information of total short circuit ratio to every separate failure rate. Failure rate parameters must be determined before modelling methods can be used.

The statistics, in this paper, have been collected by Finnish network companies, where the used statistics are based on population and outages. The analyzed data consists of 2400 faults, where about 60% of those were aerial line faults. The population covers about 11,000 km of cables and aerial lines and about 12,500 transformers for several years' time period. General failure rate of components were calculated as a weighted mean from failure rates of separate companies. Defined parameter groups are used to calculate the separate failure rates. The basic input data set is the component information, i.e. type, failure rate, and the network topology, also some other information are needed which are affecting results of the analysis, such as repair times and automation devices installed. In the enhanced radial reliability analysis, network is analyzed with feeders and zones, where zone refers to a part of feeder. In the given analysis, the expected amount of permanent and temporary failures and voltage dips in a zone are calculated as a sum of the individual network component failures. Determination of repair time is done by analyzing the possibilities to isolate load points from the faulted component and then restore the load points with dis-connectors. For a temporary fault, the whole feeder is experiencing the same short interruption. In given analysis, experienced permanent and temporary faults and voltage dips are defined for each load point. Cost information is based on total interruption times in certain area, permanent and temporary fault and voltage dip occurrences defined with the radial network reliability analysis [13]. Utility outage costs is based on the value of non-distributed energy and fault repair costs. Other costs, such as losses in production are considered in defining inconvenience costs for the customer. The expected permanent outage annual costs are caused by a fault in the zone under study. Thus RNR framework can be expressed as an asset management model considering the Life Cycle Assessment (LCA) of power system equipment. Combined with OPF, it is a complete one-stop solution network management and planning platform. Reliability of reconfiguration by replacing overhead lines and underground cables, is evaluated considering environmental, consumer preference, n-1 contingency and DSO objectives while minimizing the investment cost.

The reconfiguration of networks is primarily done to accommodate new consumers. This is achieved by extending the connection of an existing node through a new arc. Secondly it is done by replacing some existing lines. Network utilities can adjust the failure rate and reliability parameters with their own network information. A Switch Gear (SG) can identify the fault region of the feeder and update it with secure supply of energy from the same power network. Reliability indices mainly include measures of outage duration and its frequency, the amount of power or energy which is not supplied, and the number of customers involved in outages. IEEE has defined reliability indices, such as System Average Interruption Frequency Index (SAIFI), System Average Interruption Duration Index (SAIDI), Customer Average Interruption Duration Index (CAIDI), Energy Not Supplied (ENS) [14]–[17]. Such index are system and customer average interruption of frequency and duration, and energy-based index, referred to as energy not supplied. Indicators are determined over a predefined period of time.

III. Reliability indices

A. Node reliability indices

1) Expected load not served (ELNS): The ELNS measures the average amount of energy not supplied to loads as a result of load shedding events. As its own name indicates, the expected load not served is a weighted average energy value accounting for both the probability of contingencies and the damage that these contingencies cause to the system in terms of lost load.

2) Loss-of-load probability (LOLP): The LOLP is computed as the probability that failure events lead to load shedding. As opposed to the ELNS, however, the loss-of-load probability is a dimensionless number that does not provide any information on the severity of the disturbance, i.e., on the energy not supplied. This lack of a clear physical meaning makes the LOLP a less intuitive metric to work with by system operators.

$$ENS = \sum_{e} P_e r_e \tag{1}$$

Where e interruption event, r_e- restoration time for interruption event e, and P_e average load interrupted by each event e.

3) loss-of-load expectation (LOLE): The LOLE assesses the expected number of hours during which loss-of-load events could happen. As the LOLP, the loss-of-load expectation fails to provide an estimation of the damage done to the system by contingencies. From a mathematical viewpoint, both the LOLE and the LOLP require the use of binary variables to be considered within a mixed-integer linear programming problem, [16, 57]. On the contrary, the ELNS can be expressed linearly, without binary variables, as follows:

$$LOLE = \sum_{o} P_o t_o \tag{2}$$

Where o is the capacity outage, p_o is individual probability of the capacity outage, t_o is the time interval based on the difference in the capacity outage magnitude due to loss of load.

B. Arc reliability indices

The arc reliability indices are summarized in the table I. This table presents the main properties of cables, overhead lines, transformers, switch-gears, consumption, generation, terrain, probability of fault and maintenance faults. In addition product description with manufacturer references are provided.

IV. OPF with Restructuring Mathematical Model

This section will outline the mathematical model developed for the reliability oriented network restructuring analyses considering AC-OPF for a distribution network. The model has been developed in AIMMS and solved using the Outer Approximation Algorithm [27] that is suitable for solving non linear non convex models like the OPF.

A. Objective Function

$$\min C^{op} + C^{inv} \tag{3}$$

$$C^{op} = \sum_{t,i,g} (P_{g,i,t} + Q_{g,i,t}) * C_g \qquad \forall t, i, g \tag{4}$$

$$\begin{aligned} C^{inv} = & \sum_{i,j,c} CRF_c * Y_{i,j,c} * C_c + \sum_{i,j,c} CRF_c * R_{i,j,c} * C_c \\ & + \sum_{i,j,c} CRF_c * (1 - R_{i,j,c}) * C_{i,j} + C^{SVC} * D_i \end{aligned} \tag{5}$$

The Objective function 3 minimises the total operational costs and investment costs. Operational costs in 4 are related to conventional generator costs due to fuel consumption. The investment costs in 5 are described by four terms: the cost of installation of new potential cables where a connection still do not exist, the cost of replacing existing obsolete cables with new ones, a representative cost of keeping existing cables as they are and the cost of installing Static Var Compensator (SVC) devices in certain nodes. The cost of existing cables is a representative cost that incorporates all the costs that a company should face to keep a cable as it is: this cost is calculated according to the history of the cable, its maintenance requirements, failures and issues and represented by the parameter Maintenance cost listed in Table II.

B. Conventional Generators, Wind Plants and Batteries

$$P_{g,i,t} \leq \overline{P}_{g,t} * W_{g,i,t} \qquad \forall g, i, t \tag{6}$$

$$Q_{g,i,t} \leq \overline{Q}_{g,t} * W_{g,i,t} \qquad \forall g, i, t \tag{7}$$

$$P_{w,i,t} \leq \overline{P}_{w,i,t} \qquad \forall w, i, t \tag{8}$$

$$Q_{w,i,t} \leq \overline{Q}_{w,i,t} \qquad \forall w, i, t \tag{9}$$

$$B^{SOC}_{b,i,t} \leq B^{cap}_{b} \qquad \forall b, i, t \tag{10}$$

$$B^{SOC}_{b,i,t} = B^{SOC}_{b,i,t} - P^{out}_{b,i,t} * \frac{1}{B^{eff}_{b}} + P^{in}_{b,i,t} \qquad \forall b, i, t \tag{11}$$

$$\frac{1}{B^{eff}_{b}} * (P^{out}_{b,i,t})^2 + (Q^{out}_{b,i,t})^2 \leq (B^{rate}_{b} * B^{cap}_{b})^2 \qquad \forall b, i, t \tag{12}$$

$$(P^{in}_{b,i,t})^2 + (Q^{in}_{b,i,t})^2 \leq (B^{rate}_{b} * B^{cap}_{b})^2 \qquad \forall b, i, t \tag{13}$$

TABLE I: Explanation of relibility indicators in RNR

Cable	In this study we concentrate on three different voltage levels, with each having one or two different types of cables, therefore five different cables are evaluated. Cables are picked based on their suitable voltage level, cable diameter, and the conductor and the insulation type. In the [18], [19] cable line weight parameters are evaluated as normalized values between 0. . . 1. The [18], [19] gives the possibility to pick a certain transmission line type [18]–[23] with fixed parameters. There are proposed five types of different cables and overhead-lines. For instance, for a transmission line of 6 kV or 10 kV with a length 9 km, a three-core cable is proposed with diameter of $3 * 70\ mm^2$. This cable weights 5400 kg/km and as tables has proposed we may choose a given conductor and the insulation, although steel conductors do not have insulation in this thesis. Cable indices in a network recuntroction evaluation are the cable's conductor and insulation. In the example, conductor weight is calculated as the diameter of the conductor (i.e, Al conductor diameter = 9.8 mm) multiplied with conduction material density (Al density = 8.89 kg/km). i.e, Cable, which weighs 5400 kg/km has an aluminium conductor which weighs 567 kg/km and insulation of 3699 kg/km [18], [19].
Overhaed lines	As mentioned above, there are different types of transmission lines depicted and some of the named are overhead-lines (OHL). The same evaluation planning in [18], [19] is used as in [18], [19] with the firstly mentioned being dependent also on [18], [19]. OHL indices as before mentioned line conductor and the insulation weight, with additional indices covering the OHL poles. The poles are picked to be suitable for each voltage. For instance, 35 kV OHL usually uses poles which span across 80 m. The number of poles needed are calculated by the tension and the sag of the line. After calculating the needed tension (tension at pole related to tension at the maximum deflection) and sag (tension related to the span of the poles) the number of poles is found with relation to line length (including the sag) and the span length of two poles.
Transformer	Transformers (Trfo) used in this study are ideal and listed in [18]–[23] , without having to relate to the criterium N-1 (in case one transformer is interrupted, the energy flow continues on). Therefore, only one transformer is depicted for a substation, with an exception of two substations which have two transformers because there are four voltage levels, which are distributed. Furthermore, the transformers used in this study are assumed to be almost equal to the ones provided by the companies. i.e. for a 110/10 kV substation, 220/15,6 kV transformer is used. Transformers are depicted as such with transmission line types. The needed indices are conductor (copper wire + profile) and insulation (transformer oil) weight at manufacturing and use phase.
Switch-gear	Switch-gears (SG) used are described in [18]–[23] . In this network, in the node points combination of different switchgears are used, based on their operating voltage levels. For this instance, 4 different combinations are made. SG indices are based on the sum of their emissions per one transformer. Main emissions listed are Climate change (GWP, kg CO2/Trfo), Acidification (AP, molh/Trfo), Eutrophication (NP, kg O2/Trfo), and SF6 % of all emissions. The needed indices' values are calculated with the minimum and the maximum emission values [18], [19].
Consumption	The evaluated network consists of two main types of consumers, residential (0-25 kWh), and commercial (25-50 kWh). It is assumed that around a substation there are 5-10 residential buildings and 1-3 commercial buildings, because the evaluated network is put together mainly by the residential areas, rather than commercial. The needed weight value is comprised the sum of the total energy demand in a node related to the minimum and maximum energy consumption in a node.
Generation	A single distributed generation source is assumed to generate 1000 kWh of electrical energy, although the submarine cable is assumed to have a smaller value because of the losses in transmission.
Terrain	To differentiate the nodes and the arcs, additionally to electrical aspects, environmental indices are used to evaluate a network.
Probability of fault & maintenance costs	As mentioned above not only electrical indices are used, also economical characteristics of a network are needed to be assessed. For the total maintenance costs [24]–[26] , the repair costs and a probability of fault is needed to be assessed for the transmission line and the substation. For this fault value is assumed based on the terrain influence on the probability of fault. Maintenance costs = Cost of repair * Probability of fault value [24]–[26]

This group of constraints define the main properties of conventional generators, wind plants and batteries. Upper limits on active and reactive power from conventional generators and wind plants are defined in constraints 6, 7, 8 and 9. While constraints 10, 11, 12 and 13 control the battery operations in terms of capacity, State of Charge (SOC), rating in and rating out respectively.

C. Grid Restructuring

$$P_{i,j,t} \leq (P^A - P^B + P^C) \qquad \forall_{i,j,t} | X_{i,j} = 0; A_{i,j} = 0 \tag{14}$$

$$\begin{aligned} P_{i,j,t} \leq & \sum_c (1 - R_{i,j,c}) * (P^A - P^B + P^C) \\ & + \sum_c R_{i,j,c} * (P^D - P^E + P^F) \qquad \forall_{i,j,t} | X_{i,j} = 1 \end{aligned} \tag{15}$$

$$P_{i,j,t} \leq \sum_c Y_{i,j,c} * (P^D - P^E + P^F) \qquad \forall_{i,j,t} | A_{i,j} = 1 \tag{16}$$

$$P^A = K_{i,j} * \Big(\frac{V_{i,t}}{T^{tr}_{i,j}}\Big)^2 \qquad \forall_{i,j,t} \tag{17}$$

$$P^B = \Big(\frac{V_{i,t}}{T^{tr}_{i,j}}\Big) * V_{j,t} * K_{i,j} * \cos(\delta_{i,t} - \delta_{j,t}) \qquad \forall_{i,j,t} \tag{18}$$

$$P^C = S_{i,j} * \sin(\delta_{i,t} - \delta_{j,t}) \qquad \forall_{i,j,t} \tag{19}$$

The traditional OPF equations are defined in this group of constraints in a way that incorporates the possibility to reconfigure the network. Constraint 14 defines the active power as the sum of three terms P^A, P^B and P^C that contains the power flow equations as described in constraints 17, 18 and 19. Constraint 15 defines how reconfiguration can happen: if an existing cable is not replaced with a new one of type c, then the binary variable $R_{i,j,c}$ will be equal to 0, therefore the second term of constraint 15 will be equal to zero and the active power will be defined as in 14. On the other hand, if an existing cable is replaced with a new one of type c, then the binary variable $R_{i,j,c}$ will be equal to 1, therefore the first term of constraint 15 will be equal to zero and the active power equation will be equal to the second term of constraint 15. The terms P^D, P^E and P^F are formulas equal to P^A, P^B and P^C respectively, where the parameters of existing cables $K_{i,j}$ and $S_{i,j}$ are replaced by the correspondent parameters of new available new cables K_c and S_c. The model has therefore the ability to choose if it is necessary to dismantle and replace an existing cable by choosing a new one among a list of cables with different properties and costs.

Constraint 16 defines how the installation of new cables where no existing connections are available can happen. In

TABLE II: Reliability indices for RNR

Arc	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Node x	1	2	3	3	6	6	12	7	9	7	9	10	13	4	4
Node y	2	4	4	5	11	12	13	9	14	8	10	11	14	7	9
Transmission length, km	2	3	2	5	4	8	3	3.00	5.00	6.00	3.00	3.00	3.00	2.00	2.00
Transmission voltage,kV	10	110	35	110	110	110	110	10.00	10.00	110.00	110.00	110.00	110.00	110.00	35.00
Line type[1,2,3,4,5]	1	3	4	5	5	3	5	1	1	5	3	5	5	5	2
Life expectancy value	0.1	0.1	1	1	1	0.1	1	0.10	0.10	1.00	0.10	1.00	1.00	1.00	1.00
Conductor type [1,2,3,4,5]	1	1	3	3+5	3+5	1	3+5	1	1	3+5	1	3+5	3+5	3+5	2
Line weight value, kg/km	0.36	0.81	0.33	0.00	0.00	0.36	0.00	0.36	0.36	0.00	0.81	0.00	0.00	0.00	0.43
Conductor weight value,kg/km	0.06	0.10	0.06	0.00	0.00	0.10	0.00	0.06	0.06	0.00	0.10	0.00	0.00	0.00	0.29
Insulation weight value,kg/km	0.38	1.00	0.40	-	-	1.00	-	0.38	0.38	-	1.00	-	-	-	0.00
OHL pole value, pcs	-	-	0.00	0.60	0.60	-	1.00	-	-	0.60	-	1.00	1.00	1.00	-
Climate change value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Fossil depletion value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Freshwater ecotoxicity value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Freshwater eutrophication value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Human toxicity value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Marine eutrophication value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Metal depletion value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Ozone depletion value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Particulate matter formation value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Photo-chemical oxidatnd formation value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Terrestrial acidification value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Terrestrial ecotoxicity value	0.01	1.00	0.09	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.12
Terrain	4	4	5	5	4	8	9	4	5	3	7	9	5	4	5
Terrain value	0.38	0.38	0.50	0.50	0.38	0.88	1.00	0.38	0.50	0.25	0.75	1.00	0.50	0.38	0.50
Fault value	0.75	0.75	0.88	0.88	0.75	0.25	1.00	0.75	0.88	0.63	0.50	1.00	0.88	0.38	0.88
AIT value	0.51	0.31	0.73	0.41	0.74	0.36	0.51	0.79	0.59	0.00	0.54	1.00	0.82	0.85	0.69
AIF Value	0.51	0.55	0.69	0.79	0.63	0.00	1.00	0.71	0.97	0.57	0.39	1.00	0.74	0.17	0.83
AID value	0.22	0.13	0.17	0.05	0.21	1.00	0.00	0.18	0.03	0.03	0.33	0.10	0.17	0.80	0.10
Availability value	0.78	0.87	0.83	0.95	0.79	0.00	1.00	0.82	0.97	0.97	0.67	0.90	0.83	0.20	0.90
Investment value	0.78	0.87	0.83	0.95	0.79	0.00	1.00	0.82	0.97	0.97	0.67	0.90	0.83	0.20	0.90
Repair costs value	0.25	0.28	0.93	0.98	0.91	0.00	1.00	0.26	0.31	0.99	0.22	0.96	0.93	0.66	0.29
Maintenance costs value	0.46	0.47	0.82	0.84	0.67	0.00	1.00	0.46	0.59	0.55	0.23	0.98	0.82	0.20	0.58

particular, if a new cable of type c is going to be installed between two nodes, the binary variable $Y_{i,j,c}$ will be equal to 1 and the active power will be equal to the terms P^D, P^E and P^F that have been explained above. On the other hand, if no cables are going to be installed, the binary variable $Y_{i,j,c}$ will be equal to 0 and no power flow will be allowed between the two nodes.

Similarly, for reactive power the same thoughts above can be applied as shown in constraints 20, 21, 22, 23, 24, 25. In this case Q^D, Q^E and Q^F are formulas equal to Q^A, Q^B and Q^C respectively, where the parameters of existing cables $K_{i,j}$, $S_{i,j}$ and $S^{sh}_{i,j}$ are replaced by the correspondent parameters of available new cables K_c, S_c and S^{sh}_c.

$$Q_{i,j,t} \leq -Q^A - Q^B + Q^C \qquad \forall_{i,j,t} | X_{i,j} = 0; A_{i,j} = 0 \tag{20}$$

$$\begin{aligned} Q_{i,j,t} \leq & \sum_c (1 - R_{i,j,c}) * (-Q^A - Q^B + Q^C) \\ & + \sum_c R_{i,j,c} * (-Q^D - Q^E + Q^F) \qquad \forall_{i,j,t} | X_{i,j} = 1 \end{aligned} \tag{21}$$

$$Q_{i,j,t} \leq \sum_c Y_{i,j,c} * (-Q^D - Q^E + Q^F) \qquad \forall_{i,j,t} | A_{i,j} = 1 \tag{22}$$

$$Q^A = \Big(S_{i,j} + \frac{S^{sh}_{i,j}}{2}\Big) * \Big(\frac{V_{i,t}}{T^{tr}_{i,j}}\Big)^2 \qquad \forall_{i,j,t} \tag{23}$$

$$Q^B = \frac{V_{i,t}}{T^{tr}_{i,j}} * V_{j,t} * K_{i,j} * \sin(\delta_{i,t} - \delta_{j,t}) \qquad \forall_{i,j,t} \tag{24}$$

$$Q^C = S_{i,j} * cos(\delta_{i,t} - \delta_{j,t}) \qquad \forall_{i,j,t} \tag{25}$$

It is straightforward that the above formulation allows also the possibility to simply dismantle existing cables without replacing them. In this case it is enough to provide a list of cables that contains also a type c with K_c, S_c and S^{sh}_c equal to zero. If chosen, this will simply correspond to absence of connection.

Reconfiguration and new potential connections can happen only in those arcs that the operator is willing to check. Not all the arcs of the grid will be subjected to such decision, therefore binary parameters $X_{i,j}$ and $A_{i,j}$ are used to select which arcs to reconfigure and which new connections to evaluate respectively.

D. Grid General Management

$$\begin{aligned} & \sum_g P_{g,i,t} + \sum_w P_{w,i,t} - \sum_j P_{i,j,t} + \sum_j P_{j,i,t} \\ & - \sum_s P^{out}_{b,i,t} - \sum_s P^{in}_{b,i,t} = P^L_{i,t} \qquad \forall_{i,t} \end{aligned} \tag{26}$$

$$\begin{aligned} & \sum_g Q_{g,i,t} + \sum_w Q_{w,i,t} - \sum_j Q_{i,j,t} + \sum_j Q_{j,i,t} \\ & - \sum_s Q^{out}_{b,i,t} - \sum_s Q^{in}_{b,i,t} = Q^L_{i,t} \qquad \forall_{i,t} \end{aligned} \tag{27}$$

$$Z_{i,j,t} = \frac{1}{\sqrt{3}} * V_{i,t} * \sqrt{(P_{i,j,t})^2 + (Q_{i,j,t})^2} \qquad \forall_{i,j,t} \tag{28}$$

$$P_{i,j,t} \leq BigM * dir_{i,j,t} \qquad \forall_{i,j,t} \tag{29}$$

$$P_{j,i,t} \leq BigM * (1 - dir_{i,j,t}) \qquad \forall_{i,j,t} \tag{30}$$

$$Q_{i,j,t} \leq BigM * dir_{i,j,t} \qquad \forall_{i,j,t} \tag{31}$$

$$Q_{j,i,t} \leq BigM * (1 - dir_{i,j,t}) \qquad \forall_{i,j,t} \tag{32}$$

$$\sum_{c} Y_{i,j,c} \leq 1 \qquad \forall_{i,j} \tag{33}$$

$$\sum_{c} R_{i,j,c} \leq 1 \qquad \forall_{i,j} \tag{34}$$

$$\underline{V} \leq V_{i,t} \leq \overline{V} + V^{SVC} * D_i \qquad \forall_{i,t} \tag{35}$$

$$\underline{\delta} \leq \delta_{i,t} \leq \overline{\delta} \qquad \forall_{i,t} \tag{36}$$

$$\underline{Z} \leq Z_{i,t} \leq \overline{Z} \qquad \forall_{i,j,t} \tag{37}$$

This set of constraints describe the main properties to take into account for the grid management. In particular flow balance for active and reactive power is defined in 26 and 27 respectively; the current is defined in 28; the flow direction is described through constraints 29, 30, 31 and 32; constraints33 and 34 limit the choice of new cables to 1; finally constraints 35, 36 and 37 define limits on the voltage, phase angle and current. Regarding constraint 36, the voltage upper limit is linked to the decision of installing a SVC device. In particular, when a SVC device is installed on a node i, the binary variable D_i is equal to 1 and the voltage upper limit increases of a value V^{SVC}. This can make a difference in the decision of dismantling a cable or installing a SVC device.

V. COMPUTATIONAL EXPERIMENTS

Computational experiments have been performed on IEEE 14 bus system represented in fig. 1, using the data-set contained in table II. It is assumed that the arcs $4-9, 4-7, 5-6, 11-10, 13-14$ are non-existing and potential connections should be evaluated. Therefore the system is now split into three microgrids as highlighted in fig. 1. Microgrids 1 and 2 are equipped with conventional and renewable sources respectively, while microgrid 3 is without any resource and can be considered as an emerging district that has been created and that needs to be connected to a neighbourhood area. The microgrid 1 considers restructuring of the existing network to accommodate the emerging district. It is straightforward that restructuring is not considered for the emerging district, because it is assumed that a new microgrid will have new and up to date equipment. Hence the trade-off between the maintenance cost of existing network and the replacement costs to accommodate new emerging demand is analysed. Moreover reconfiguration is allowed on arc $1-2, 2-4, 4-3, 2-3, 2-5, 1-5, 4-5$ in order to verify how the establishment of new connections are affecting the reliability of the system.

As a result, new cables installations are created on arcs $4-7, 4-9$ and a cable $1-2$ is replaced with a new cable provided with higher sustenance. Note that microgrid 2 remains isolated because it already has enough power from the renewable plant.

VI. CONCLUSION

A methodology to analyse how connecting emerging districts to existing microgrids can affect the reliability of the whole system has been presented. The technical aspects of AC-OPF have been thoroughly taken into account and the reliability oriented Network Restructuring RNR framework has been developed and implemented. The results showed that reliability aspects are crucial when evaluating new investments in grid expansion: new connections should always be coupled with a more holistic evaluation of the conditions of the existing networks as they may require further investments in upgrades to fulfill the new requirements. When the system operator considers investments for power network expansion, it should also consider restructuring of the existing network at the same time. The presented model RNR is able to address both decisions holistically and therefore more investigation is required in this area.

NOMENCLATURE

Indexes

t time step
i, j nodes of the grid
c available new cables
g conventional generators
w wind plants
b batteries

Parameters

C^{op} Operational costs
C^{inv} Investment costs
C_g Operational cost of conventional generator g
$\overline{P}_{g,t}$ Upper limit on active power from conventional generator g at time t
$\overline{Q}_{g,t}$ Upper limit on reactive power from conventional generator g at time t
$\overline{P}_{w,i,t}$ Upper limit on active power from wind plants w on node i at time t
$\overline{Q}_{w,i,t}$ Upper limit on reactive power from wind plants w on node i at time t
B^{cap}_b Capacity of battery b
B^{eff}_b Efficiency of battery b
B^{rate}_b Rating of battery b
$T^{tr}_{i,j}$ Tap ratio of transformer placed between nodes i and j
$K_{i,j}$ Conductance of existing cables placed between nodes i and j
$S_{i,j}$ susceptance of existing cables placed between nodes i and j
$S^{sh}_{i,j}$ shunt susceptance factor of existing cables placed between nodes i and j
$S_{i,j}$ Susceptance of existing cables placed between nodes i and j
$S^{sh}_{i,j}$ Shunt suceptance of existing cables placed between nodes i and j
$C_{i,j}$ Representative cost of existing cables due to their history of maintenance operations
$A_{i,j}$ Binary parameter defining if a new potential cable can be installed between nodes i and j
$X_{i,j}$ Binary parameter defining if an existing cable between nodes i and j should be checked for possible replacement
K_c Conductance of new cables of type c
S_c Susceptance of new cables of type c
S^{sh}_c Shunt susceptance of new cables of type c
CRF_c Capital recovery factor of new cables of type c
C_c Investment cost of new cables of type c
V^{SVC} Possible incremental voltage due to installation of a SVC device
$\underline{V}, \overline{V}$ Minimum and maximum limits for voltage
$\underline{\delta}, \overline{\delta}$ Minimum and maximum limits for phase angle
$\underline{Z}, \overline{Z}$ Minimum and maximum limits for current
C^{SVC} Investment cost of an SVC device
$P^L_{i,t}$ Active load in node i at time t
$Q^L_{i,t}$ Reactive load in node i at time t

Variables

$P_{g,i,t}$ Active power from conventional generator g in node i at time t
$Q_{g,i,t}$ Reactive power from conventional generator g in node i at time t
$W_{g,i,t}$ Binary variable equal to 1 if the conventional generator g in node i is working at time t
$P_{w,i,t}$ Active power from wind plants w in node i at time t
$Q_{w,i,t}$ Reactive power from wind plants w in node i at time t
$B^{SOC}_{b,i,t}$ State of charge of battery b in node i at time t
$P^{out}_{b,i,t}$ Active power from battery b in node i at time t
$P^{in}_{b,i,t}$ Active power into battery b in node i at time t
$Q^{out}_{b,i,t}$ Reactive power from battery b in node i at time t
$Q^{in}_{b,i,t}$ Reactive power into battery b in node i at time t

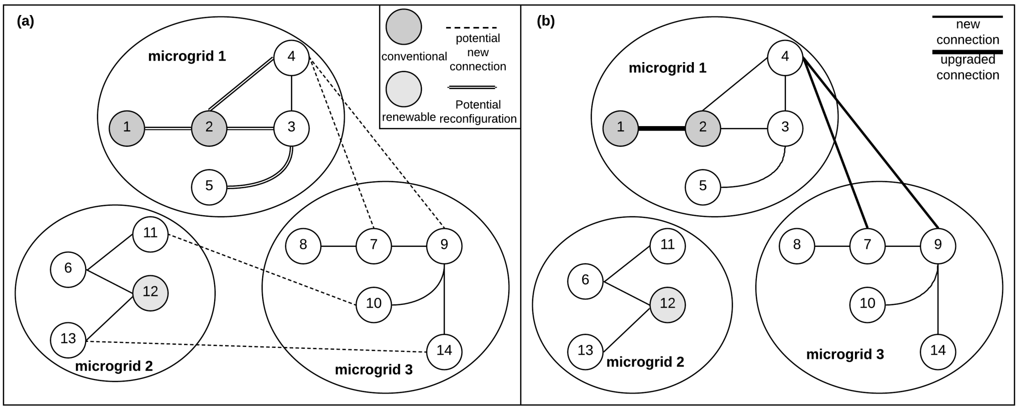


Fig. 1: The microgrid structures: (a) expansion (b) expansion with reconfiguration

- $V_{i,t}$ Voltage value in node i at time t
- $\delta_{i,t}$ Phase angle value in node i at time t
- $Z_{i,j,t}$ Current value between nodes i and j at time t
- $dir_{i,j,t}$ Binary variable equal to 1 if the power flow is from node i to node j, 0 otherwise
- $Y_{i,j,c}$ Binary variable equal to 1 if a potential new cable of type c is installed between nodes i and j, 0 otherwise
- $R_{i,j,c}$ Binary variable equal to 1 if an existing cable between nodes i and j is replaced by a new cable of type c, 0 otherwise
- D_i Binary variable equal to 1 if an SVC device is installed on node i, 0 otherwise

REFERENCES

[1] K. Lehmann, A. Grastien, and P. Van Hentenryck, "Ac-feasibility on tree networks is np-hard," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 798–801, 2016.

[2] D. S. Hochbaum, *Approximation algorithms for NP-hard problems*. PWS Publishing Co., 1996.

[3] A. J. Conejo and L. Baringo, "Optimal power flow," in *Power System Operations*. Springer, 2018, pp. 165–196.

[4] X. Zhang and A. J. Conejo, "Robust transmission expansion planning representing long-and short-term uncertainty," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1329–1338, 2018.

[5] N. Amjady, A. Attarha, S. Dehghan, and A. J. Conejo, "Adaptive robust expansion planning for a distribution network with ders," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1698–1715, 2018.

[6] A. Attarha, N. Amjady, and A. J. Conejo, "Adaptive robust ac optimal power flow considering load and wind power uncertainties," *International Journal of Electrical Power & Energy Systems*, vol. 96, pp. 132–142, 2018.

[7] X. Zhang and A. J. Conejo, "Candidate line selection for transmission expansion planning considering long-and short-term uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 100, pp. 320–330, 2018.

[8] A. J. Conejo and R. Sioshansi, "Rethinking restructured electricity market design: Lessons learned and future needs," *International Journal of Electrical Power & Energy Systems*, vol. 98, pp. 520–530, 2018.

[9] C. L. Borges and D. M. Falcao, "Optimal distributed generation allocation for reliability, losses, and voltage improvement," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 6, pp. 413–420, 2006.

[10] J. Pylvänäinen, P. Verho, J. Järvinen, S. Kunttu, and J. Sarsama, "Advanced failure rate and distribution network reliability modelling as part of network planning software," in *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*. IET, 2005, pp. 1–5.

[11] E. HILLBERG, "System integrity protection schemes–increasing operational security and system capacity e. hillberg, f. trengereid 2, ø. breidablik 2, k. uhlen, g. kjølle, s. løvlund 2, jo gjerde 2."

[12] G. Kjølle, O. Gjerde, and M. Hofmann, "Monitoring vulnerability in power systems," *Paper accepted for PMAPS*, 2012.

[13] R. Karki, R. Billinton, and A. K. Verma, *Reliability Modeling and Analysis of Smart Power Systems*. Springer, 2014.

[14] M. Čepin, *Assessment of power system reliability: methods and applications*. Springer Science & Business Media, 2011.

[15] A. Volkanovski, M. Čepin, and B. Mavko, "Application of the fault tree analysis for assessment of power system reliability," *Reliability Engineering & System Safety*, vol. 94, no. 6, pp. 1116–1127, 2009.

[16] R. N. Allan *et al.*, *Reliability evaluation of power systems*. Springer Science & Business Media, 2013.

[17] R. Billinton and R. N. Allan, *Reliability assessment of large electric power systems*. Springer Science & Business Media, 2012.

[18] Nexans, "6-36kv medium voltage underground power cables xlpe insulated cables," http://www.nexans.co.uk/, 2017, [Online; accessed 01-Oct-2017].

[19] Elandcables, "High voltage / 2xs(f)2y, a2xs(f)2y 110kv power cable," https://www.elandcables.com/, 2017, [Online; accessed 01-Oct-2017].

[20] B. B. Zad, J. Lobry, and F. Vallée, "Impacts of the model uncertainty on the voltage regulation problem of medium-voltage distribution systems," *IET Generation, Transmission & Distribution*, 2018.

[21] G. Nagel, "System dependant idmt settings: direct buried underground cable," 2014.

[22] V. J. H. Jiménez, E. D. Castronuovo, and I. S. Rodríguez-Morcillo, "Optimal statistical calculation of underground cable bundles positions for time-varying currents," *International Journal of Electrical Power & Energy Systems*, vol. 95, pp. 26–35, 2018.

[23] K. Kopsidas and S. Liu, "Power network reliability framework for integrating cable design and ageing," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1521–1532, 2018.

[24] R. Pletka, J. Khangura, A. Rawlins, E. Waldren, and D. Wilson, "Capital costs for transmission and substations," *Western Electricity Coordinating Council*, p. 35, 2014.

[25] T. Mason, T. Curry, and D. Wilson, "Capital costs for transmission and substations: recommendations for wecc transmission expansion planning," *Black & Veatch*, 2012.

[26] R. Benato and D. Napolitano, "Overall cost comparison between cable and overhead lines including the costs for repair after random failures," *IEEE transactions on Power Delivery*, vol. 27, no. 3, pp. 1213–1222, 2012.

[27] M. A. Duran and I. E. Grossmann, "An outer-approximation algorithm for a class of mixed-integer nonlinear programs," *Mathematical programming*, vol. 36, no. 3, pp. 307–339, 1986.

Curriculum Vitae

1. Personal data

Name	Sambeet Mishra
Date and place of birth	16^{th} August 1989, India
Nationality	Indian

2. Contact information

Address	NRG 512, Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, School of Engineering, Ehitajate tee 5, 19086 Tallinn, Estonia
Phone	+372 58903996
E-mail	sambeet.mishra@TalTech.ee

3. Education

2014–2018	Tallinn University of Technology, School of Engineering, Department of Electrical Power Engineering and Mechatronics Research group: Energy Economics Doctoral Studies
2016–2017	Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management Professional group: Business economics and optimization Visiting researcher & Affiliated
2011–2013	KIIT University, Department of Electrical Engineering, Power & Energy Systems, Master of Technology (M.Tech) *cum laude*
2006–2010	KIIT University, Department of Electrical Engineering, Electrical Engineering, Bachelor of Technology (B.Tech)

Courses:

- 10 hour(s) Professional development from IEEE Smart Grid
- The DTU CEE Summer School 2016 “Uncertainty in Electricity Markets and System Operation”
- PhD Winter school 2017 "Stochastic programming with applications in energy, logistics and finance"
- VU Amsterdam Summer School 2017 "Operations Research: A Mathematical Way to Optmize your World"
- IØ8807 - Advanced Course in Complementarity Models and Equilibrium, Department of Industrial Economics and Technology Management, NTNU

4. Language competence

Odia, Hindi, Sanskrit	native
English	professional
Estonian	A2 level

5. Professional employment

2015– *Today* ...	Tallinn University of Technology, Junior Scientist
2014–2015	Tallinn University of Technology, Engineer

6. Computer skills

- Operating systems: Windows
- Document preparation: LaTeX, Microsoft office, Open Office
- Programming languages: Python, R, Julia, Matlab
- Scientific packages: GAMS, AIMMS

7. Honours and awards

- 2017, Smart specialization scholarship

8. Field of research

- Power & Energy Systems
- Mathematical modelling and optimization
- Energy Economics/Markets
- Energy Informatics

9. Industrial Research Partners

- TrønderEnergi, Power distribution operator company, Trondheim, Norway
- BrooklynMG, LO3 Energy, Blockchain based community powered local microgrid, NewYork, USA.

10. Scientific works *h-index = 7*
Projects: Flex4RES, PUT533, LEP15070, VNF15057, LEP18038, CenSES
Reviewer:

1. International Journal of Energy Research, Publisher: John Wiley & Sons
2. IEEE transactions on Industrial Electronics, Publisher: IEEE

Papers Published:

1. **S. Mishra**, M.Leinakse, and I.Palu "Wind Power Variation Identification using Ramping Behavior Analysis" Elsevier Energy Procedia, pages - 141:565–571, 2017
2. **S. Mishra**, M. Leinakse, I. Palu, and J. Kilter "Ramping Behaviour Analysis of Wind Farms" EEEIC2018, pages - 1(1):1–7, 2018
3. **S. Mishra**, C. Würsig and I.Palu "Multivariate Scenario Generation - an ARIMA and Copula Approach, International Journal of Modeling and Optimization, 2018
4. **S. Mishra**, C.Bordin,J.Fornes,and I.Palu "Reliability Framework for Power Network Assessment" E3S Web of Conferences, 2018

5. **S. Mishra**, C. Bordin, and I. Palu "RNR: Reliability oriented Network Restructuring" International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 2018

Papers under review for journal publication:

1. **S. Mishra**, C. Bordin, A. Tomasgard and I. Palu. 'Coordinated Microgrid Expansion Planning', 2018

Elulookirjeldus

1. Isikuandmed

Nimi	Sambeet Mishra
Sünniaeg ja -koht	16.08.1989, India
Kodakondsus	India

2. Kontaktandmed

Aadress	Tallinna Tehnikaülikool, Inseneriteaduskond: Elektroenergeetika ja mehhatroonika instituut, Ehitajate tee 5, 19086 Tallinn, Estonia
Telefon	+37258903996
E-post	sambeet.mishra@ttu.ee

3. Haridus

2014–...	Tallinna Tehnikaülikool,Inseneriteaduskond: Elektroenergeetika ja mehhatroonika instituut, elektrienergia, doktoriõpe
2016–2017	Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management Professional group: Business economics and optimization, Visiting researcher
2011–2013	KIIT University, Department of Electrical Engineering, Power & Energy Systems, Master of Technology (M.Tech), *cum laude*
2006–2010	KIIT University, Department of Electrical Engineering, Electrical Engineering, Bachelor of Technology (B.Tech)

4. Keelteoskus

Inglise keel	Kõrgtase
Odia, Hindi, Sanskrit keel	emakeel
Eesti keel	A2 level

5. Teenistuskäik

2014– ...	Tallinna Tehnikaülikool, Elektroenergeetika ja mehhatroonika instituut, Nooremteadur

6. Arvutioskus

- Operatsioonisüsteemid: Windows, Linux
- Kontoritarkvara: MS Office, LaTeX
- Programmeerimiskeeled: Python, R, Julia
- Teadustarkvara paketid: GAMS, AIMMS, MATLAB