Data driven methods for DSO smart grid operation in the context of flexible energy systems

Blomgren, Emma Margareta Viktoria

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Ph.D. Thesis
Doctor of Philosophy in Engineering

DTU Compute
Department of Applied Mathematics and Computer Science

Data driven methods for DSO smart grid operation in the context of flexible energy systems

Emma Margareta Viktoria
Blomgren

Kongens Lyngby 2022

Data driven methods for DSO smart grid operation in the context of flexible energy systems
Submitted August, 2022

Author
Emma Margareta Viktoria Blomgren
E-mail: emvb@dtu.dk

Supervisors
Professor Henrik Madsen
Section for Dynamical Systems
Department of Applied Mathematics and Computer Science
Technichal University of Denmark
E-mail: hmad@dtu.dk

Dr Razgar Ebrahimy
Section for Dynamical Systems
Department of Applied Mathematics and Computer Science
Technichal University of Denmark
E-mail: raze@dtu.dk

Dr Seyyed Ali Pourmousavi Kani
School of Electrical and Electronic Engineering
The University of Adelaide
E-mail: a.pourm@adelaide.edu.au

DTU Compute
Department of Applied Mathematics and Computer Science
Technical University of Denmark

Matematiktorvet
Building 303B
2800 Kongens Lyngby, Denmark
Phone +45 4525 3031
compute@compute.dtu.dk
www.compute.dtu.dk

Summary (English)

The pathways to decarbonize the energy sector is not only changing the ways in which we consume and produce electricity, but the entire power system at large. As generation is moved to the edge of the grid, while, among other things, electrification of transport and heating leads to changes in load behaviour, new situations arise in the power distribution systems. The increasing levels of intermittent distributed generation (DG) and the changing demand might stress the distribution grids, which highlights the need for new methods to operate power distribution systems.

For low voltage (LV) grids in the distribution systems, the observability is typically zero, thus implying no ability to respond to grid issues that might arise from the more volatile and stochastic generation and demand. Meanwhile, grid equipment, such as transformers, are more often loaded above their rated capacities. Rated limits are conventionally static, while the actual loading capability is dynamic due to seasonal (daily, yearly, etc.) variations in the operating environment. Thereby, static ratings leave unused capacity in the grid.

In this thesis, focus is given to developing new methods for online monitoring and forecasting of both grid and grid equipment states during operation. More specifically, a per phase node voltage estimation method is developed for unbalanced radial LV grids, proven to have reasonable accuracy with root mean squared errors ranging from 0.002 – 0.0004 p.u. depending on the node. A transformer thermal model is further developed for the application of dynamic transformer rating, proven to be capable of providing 6-hour forecasts.

While the mentioned models address real-time operation tools to gain information about the operating conditions, the thesis further develops an operational framework to solve emerging grid issues. The Smart Energy-Operating System (SE-OS) offers a platform to request ancillary services provided through flexible resources, and the distribution system operator (DSO) framework is derived from these principles.

The suggested operational framework involves ancillary services provided by an aggregator, energy communities and battery energy storage systems. To support the coordination of flexibility based ancillary services, a price elasticity model for aggre-

gators is developed. The method enables analysis and evaluation of the capability of flexible resources to provide the ancillary services needed by the DSO.

The operational framework further incorporates data-driven online monitoring and forecasting tools, for which the developed methods are intended. This enables a shift in DSO operation strategy from the traditionally passive to adaptive operation.

Summary (Danish)

Vejene til at dekarbonisere energisektoren ændrer ikke kun den måde, vi forbruger og producerer elektricitet på, men hele elsystemet som helhed. I takt med at produktionen flyttes til de ydre deler af elnettet, mens blandt andet elektrificering af transport og opvarmning fører til ændringer i elforbruget, opstår der nye situationer i eldistributionssystemerne. De stigende niveauer af intermitterende distribueret produktion (DG) og den skiftende elforbrug kan stresse distributionsnettene, hvilket understreger behovet for nye metoder til at drive eldistributionssystemer.

For lavspændingsnet (LV) i distributionssystemerne er observerbarheden typisk nul, hvilket betyder, at der ikke er nogen evne til at reagere på netproblemer, der måtte opstå fra den mere voltaile og stokastiske elproduktion og efterspørgsel. I mellemtiden er netudstyr, såsom transformatorer, oftere belastet over deres nominelle kapacitet. Nominelle grænser er konventionelt statiske, mens den faktiske lasteevne er dynamisk på grund af sæsonbestemte (daglige, årlige osv.) variationer i driftsmiljøet. Derved efterlader statiske ratings uudnyttet kapacitet i elnettet.

I dette Ph.D.-afhandling fokuseres der på at udvikle nye metoder til online overvågning og prognose af både net- og netudstyrstilstande under drift. Mere specifikt er der udviklet en pr. fase spændingsestimeringsmetode til ubalancerede radiale LV net, som har vist sig at have rimelig præcision med rodmiddel kvadratfejl af 0.002 – 0.0004 p.u. afhængigt af noden. En termisk transformator model er videreudviklet til anvendelse af dynamisk transformator rating, bevist at være i stand til at levere 6-timers prognoser.

Mens de nævnte modeller omhandler realtidsdriftsværktøjer for at få information om driftsbetingelserne, videreudvikler afhandlingen en operationel ramme til at løse de lokale netproblemer. Smart Energy-Operating System (SE-OS) tilbyder en platform til at anmode om systemydelser leveret gennem fleksible ressourcer, og distributionssystemoperatørens (DSO)-rammen er afledt af disse principper.

Den foreslåede operationelle ramme involverer systemydelser leveret af en aggregator, energifællesskaber og batterienergilagringssystemer. For at understøtte koordineringen af fleksibilitetsbaserede systemydelser udvikles en priselasticitetsmodel for aggre-

gatorer. Metoden muliggør analyse og evaluering af fleksible ressourcer evne til at levere de systemydelser, som DSO'en har brug for.

Den operationelle ramme omfatter yderligere datadrevne online overvågnings- og prognoseværktøjer, som de udviklede metoder er udviklet til. Dette muliggør et skift i DSO-driftsstrategien fra traditionelt passiv til adaptiv drift.

Preface

This Ph.D. thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark in fulfillment of the requirements for acquiring a Ph.D. degree.

The Ph.D. study has been financially supported by the project Flexible Energy Denmark (FED), which is funded by Innovation Fund Denmark under Grant No. 8090-00069B.

The Ph.D. thesis project focuses on developing data-driven models for radial low voltage DSO smart grids within the context of flexible energy systems. The thesis proposes an adaptive operational framework for DSO smart grid operation. It further develops methods related to online monitoring and forecasting for DSOs as well as analytical methods for aggregators to evaluate price responsive demand.

The thesis includes a summary report and four research papers. The Ph.D. project was carried through between August 2019 and August 2022.

Kongens Lyngby, July 15, 2022

Emma Margareta Viktoria
Blomgren

Acknowledgements

Firstly, I would like to acknowledge my supervisors. Thank you, Professor Henrik Madsen, for giving me the opportunity of conducting the PhD within this field and the supervision on the topics. A special thanks goes to Dr. Razgar Ebrahimy, for supervising and guiding me through these three years, always taking the time for a discussion. Also, many thanks to my supervisor Dr. Seyyed Ali Pourmousavi Kani, for interesting, stimulating and motivational discussions.

I would further like to thank my colleagues, Dr. Mohsen Banaei and Dr. Francesco D'Ettorre, for insightful discussions, collaboration and support, not least when my computer broke down before the submission of a paper. Thank you Dr. Giulia De Zotti, for the extra guidance, support and collaboration.

Thank you Professor Olof Samuelsson, for the interesting discussions, mentoring on electrical power systems and the patience to listen to me. Another thanks goes to Claus Schack Urup at Trefor, for the many meetings and technical support on the electrical grids and to Pranay Krishen for discussions about the data.

Finally, I am so grateful for the support from my parents and the joy and kind words from my friends. And for my partner Felix, I am so grateful for your love, support and understanding throughout this period. Tack!

List of Publications

Publications included in this thesis

Paper A
Emma M. V. Blomgren, Razgar Ebrahimy, Seyyed Ali Pourmousavi Kani, Jan Kloppenborg Møller, Francesco D'Ettorre, Mohsen Banaei, Henrik Madsen,
"Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty",
Proceedings of 2021 IEEE PES Innovative Smart Grid Technologies Conference Europe (2021)

Paper B
Emma M.V. Blomgren, Francesco D'Ettorre, Olof Samuelsson, Mohsen Banaei, Razgar Ebrahimy, Magne E. Rasmussen, Nicolaj H. Nielsen, Anton R. Larsen, Henrik Madsen,
"Grey-box modeling for dynamic rating of oil-immersed transformers in power distribution networks"
(submitted to) *Electric Power Systems Research*

Paper C
Emma M.V. Blomgren, Mohsen Banaei, Razgar Ebrahimy, Olof Samuelsson, Francesco D'Ettorre, Henrik Madsen,
"Intensive data-driven model for real-time observability in low voltage radial DSO grids"
(submitted to) *Applied Energy*

Paper D
Tilman Weckesser, Dominik Franjo Dominković , Emma M.V. Blomgren, Amos Schledorn, Henrik Madsen,
"Renewable Energy Communities: Optimal sizing and distribution grid impact of photo-voltaics and battery storage"
Applied Energy (2021)

Other publications not included in this thesis

Paper 1 Peter Nystrup, Henrik Madsen, Emma M.V. Blomgren, Giulia de Zotti, "Clustering commercial and industrial load patterns for long-term energy planning" *Smart Energy* (2021)

Paper 2
Mohsen Banaei, Francesco D'Ettorre, Razgar Ebrahimy, Emma M. V. Blomgren, Henrik Madsen,
"Mutual Impacts of Procuring Energy Flexibility and Equipment Degradation at the Residential Consumers Level"
Proceedings of 2021 IEEE PES Innovative Smart Grid Technologies Europe (2021)

Paper 3
F. D'Ettorre, M. Banaei, R. Ebrahimy, S. Ali Pourmousavi, E.M.V. Blomgren, J. Kowalski, Z. Bohdanowicz, B. Łopaciuk-Gonczaryk, C. Biele, H. Madsen,
"Exploiting demand-side flexibility: State-of-the-art, open issues and social perspective"
Renewable and Sustainable Energy Reviews (2022)

Paper 4
Mohsen Banaei, Francesco D'Ettorre, Razgar Ebrahimy, S. Ali Pourmousavi, Emma M.V. Blomgren, Henrik Madsen,
"A Stochastic Methodology to Exploit Maximum Flexibility of Swimming Pool Heating Systems"
(Submitted to) *International Journal of Electrical Power and Energy Systems*

Contents

List of Figures

List of Tables

Part I

Summary Report

CHAPTER 1
Introduction

1.1 Context and Motivation

While the severity of climate change is reported in the latest report from the Intergovernmental Panel on Climate Change (IPCC), adaptive measures are advocated to mitigate its impact. Among the adaptive measures mentioned in the report are smart grid-technologies, climate responding energy markets and improved capacity to respond to shortfalls in energy supply [1]. Further, actions taken within the European Union (EU) include an aim at climate neutrality in 2050 as well as a goal of 32% share of renewable energy in the energy mix by 2030 [2, 3]. To support the indispensable decarbonization of the energy sector, the power system is undergoing a paradigm shift, meaning that the ways in which we consume, produce and operate in the power system have to change.

Increasing levels of renewable distributed generation (DG) along with technology that allows to shift the consumption to times with intermittent renewable generation, i.e. flexible consumers and energy management systems, constitute some of the promising solutions towards a carbon neutral power system [4, 5, 6]. As the generation is moved to the lower topological levels of the power system simultaneously as demand behaviour is changing through new types of loads such as electric vehicles, it might stress the low voltage (LV) networks, which were not originally designed for these power flow situations. Voltage violations, congestion and reverse power flow in the LV systems, might be some of the concerns for a distribution system operator (DSO) [7, 8, 9].

Meanwhile, the real-time observability in LV networks are typically nonexistent or very low, making it impossible to mitigate any grid limit violations during operation. Furthermore, the ratings of grid equipment are static although the operating conditions are varying, leaving unused capacity in the grid. To enable a shift from the traditionally passive operation approach [10] to a more active grid management enabling full utilization of the grid and renewable DGs, new operation models and methods in the distribution systems are required.

To implement such active operation schemes, the DSOs firstly need increased observability of grid equipment and grid states. With such information the DSO could

evaluate the physical limits of the grid in relation to real time demand and generation. To solve any detected local grid issues, it is also important that the new operational framework allows for participation to request local ancillary services [11]. For instance, the Smart Energy - Operation System (SE-OS) provides such a platform, where ancillary services are provided through activating flexible consumers [12].

To successfully apply methods for increased observability in DSO grids, the methods need to be data-driven to reflect the state of the grid in real-time. Meanwhile, physics-informed methods could provide the DSO with a certain degree of explainability to derive causes for local grid issues, while enabling analysis of various scenarios that could occur during operation. The methods further need to be able to give probabilistic monitoring and forecasting outputs to support risk averse operation strategies for the DSO. Moreover, appropriate forecast horizons with reasonable accuracy are crucial for the DSOs ability to react to certain grid situations in time and solve local grid issues.

Furthermore, digitization of LV grids is key in enabling novel and active operation strategies [13]. While smart meters are being installed at scale at the consumer premises, their usage for operational purposes is limited due the cost of communication infrastructure [14]. It is thus of importance that the required installation of sensors for observability solutions are cheap and efficient in terms of data flow and infrastructure. This is essential to keep in mind when developing the new operational tools in order to achieve practical and scalable solutions bridging the gap towards application at scale.

As the DSOs become aware of local grid issues in real-time through online monitoring and forecasting tools, they could also solve any local grid issues, for instance by requesting local ancillary services. Demand response (DR) constitutes a promising alternative to provide ancillary services in distribution grids [15, 16] and the aforementioned SE-OS a platform to exchange these services. It is thus of importance to define new DSO operational frameworks, within this context to identify how to coordinate with actors within the platform, such as aggregators, providing the ancillary services. Furthermore, the aggregators need probabilistic analytical methods to ensure that requested ancillary services can be delivered from their flexible resources to the DSOs. By solving local grid issues through DR ancillary services within the SE-OS the DSO operation would not only be active, but also adaptive.

1.2 Thesis Objectives

The objective of this thesis is to develop an operational framework for adaptive DSO smart grid operation consisting of data-driven approaches.

Objectives related to DSO smart grid operation are:

- To develop a data-driven model, predicting the transformer temperature for online dynamic transformer rating in LV grids.
- To develop a data-driven model for real time online voltage estimations in LV grids.
- To develop practical methods, in terms of computational burden and required setup, based on real-world installations to increase observability in the grid.
- To develop probabilistic methods for dynamic operation that can be used for risk averse operation strategies.

To further enable adaptive DSO operation in the context of utilizing flexible resources in the SE-OS, the objectives are:

- To develop a probabilistic model to estimate the quantity and uncertainty of price responsive demand response.
- To develop a framework for adaptive DSO operation of LV grids, including the aggregator-DSO communication concepts for requesting flexibility as well as a variation of flexible resources.

1.3 Thesis Structure/Outline of Thesis

This thesis consists of two parts. Part I presents a summary of the research conducted in the thesis. Chapter 2 provides relevant background information and introduction to concepts that lays the ground for the following chapters. Chapter 3 presents the experimental setup and data collection from two living lab installations. Chapter 4 presents methods and discussion of the results related to aggregators, whereas chapter 5 follows the same structure focusing on DSO operation instead. In chapter 6 conclusions are outlined and suggestions for future work provided. Finally, Part II presents preprints of papers included in the thesis.

CHAPTER 2
Background

This chapter presents relevant theory and information, laying the ground for the following chapters of this thesis. It presents the concept of the smart-energy operation-system, for which the methods in paper A, B and C are developed. Furthermore, relevant theory and state of the art for local DSO grid issues such as congestion and rating of transformers as well as voltage drop and rise is presented.

2.1 Smart-Energy Operation-System

The work in this thesis is derived from the concept of Smart-Energy Operating-System (SE-OS) [12, 17]. The SE-OS framework is a response to the Danish goals to eliminate fossil fuel plants and to achieve a 50 % coverage of the total electricity consumption through wind power by 2050. To reach these goals conventional operation methods and provision of ancillary services are insufficient as the high share of intermittent electricity production requires demand matching to the production instead of the traditional vice versa approach. Thus, the SE-OS compromises a framework to enable such energy flexibility solutions at different topological levels in the power system to meet the requirements of a transition towards a carbon neutral energy system.

Traditional ancillary services (AS) markets and methods have served the conventional centralized system well. However, to fully utilize flexible resources and to efficiently operate decentralized generation (DG), the methods involving centralized large-scale optimization will be slow due to the increasing complexity of the system. Realizing that the complex computations are required at each time step, further limits the applicability of conventional methods to the future power system. Another important issue is that for the future weather-driven low-carbon power system the flexibility has to be provided at end-users (residential buildings, industry, etc) at the low voltage (LV) levels in the power system. However, conventional market bidding and clearing will not work on this low topological level, with the increasing number of single AS providers such as DGs and flexible resources. Instead the SE-OS builds on unidirectional communication, forecasting and control-based approaches. Here aggregators play a crucial role in managing and optimizing the flexibility by distributed energy resources (DERs).

The framework applies both direct an indirect control mechanisms to activate electrical DERs. By using direct control, an aggregator would transmit signals to turn on and off DERs. If instead adapting indirect control, signals are broadcasted to DERs, who themselves can run internal optimization through for example an energy management system (EMS) and decide whether they would like to shift their consumption/production or not. In SE-OS mostly price-based signals are adopted in the indirect control mechanisms, but CO2-signals can and has also been implemented. The designed indirect control signal, or rather penalty signal, function is illustrated in Figure 2.2 for the configuration of the low-levels in the SE-OS. Here the prices generated by the function can be designed with different objectives for ancillary services, such as voltage or frequency control, but multilevel control objective are also possible.

The SE-OS framework, seen in Figure 2.1, consists of different stochastic optimization layers, thereby avoiding large-scale optimization. These optimization layers naturally, has different spatial-temporal resolutions for the aggregation of DERs [17]. The top layer represent day-ahead and intraday markets for larger regions and a lower temporal solution along with the conventional market and system optimizations. The lower layers in Figure 2.1 has an increasing spatial-temporal resolution in the aggregation and optimization of flexible DERs [18]. Thereby, the framework addresses the dynamic and stochastic behaviour of DERs, and control-based technologies can be embedded in the SE-OS [19, 20, 21]. In the layers the aggregators also use the flexibility function in Figure 2.2 to predict the response from DERs using time series of price signals [22, 23].

The methods in the SE-OS use various data sources such as meteorological, electricity market and load demand behaviour data to build algorithms to operate in an open-loop scheme. Forecasts and control schemes are thus updated for each time step during operation. The framework requires real-time online forecasting and operation algorithms, for which new technological setup and installation of devices are needed, both at DER and power system levels.

The SE-OS framework has been used to implement solutions enabling flexibility for wastewater treatment plants, supermarket cooling systems, heat pumps, buildings and district heating [24, 25, 26]. It has further demonstrated that it is possible to unlock flexibility in already existing resources through real-time data-driven methods utilizing available data through meteorological forecasts and predictive algorithms. The implementations further highlight the importance of considering stochastic and dynamic behaviour when dealing with energy flexibility as it deviates from traditional consumption-production behaviour. This is an important consideration for aggregators that wish to enter the market. For further details on the SE-OS the reader is referred to chapter 6 in [12].

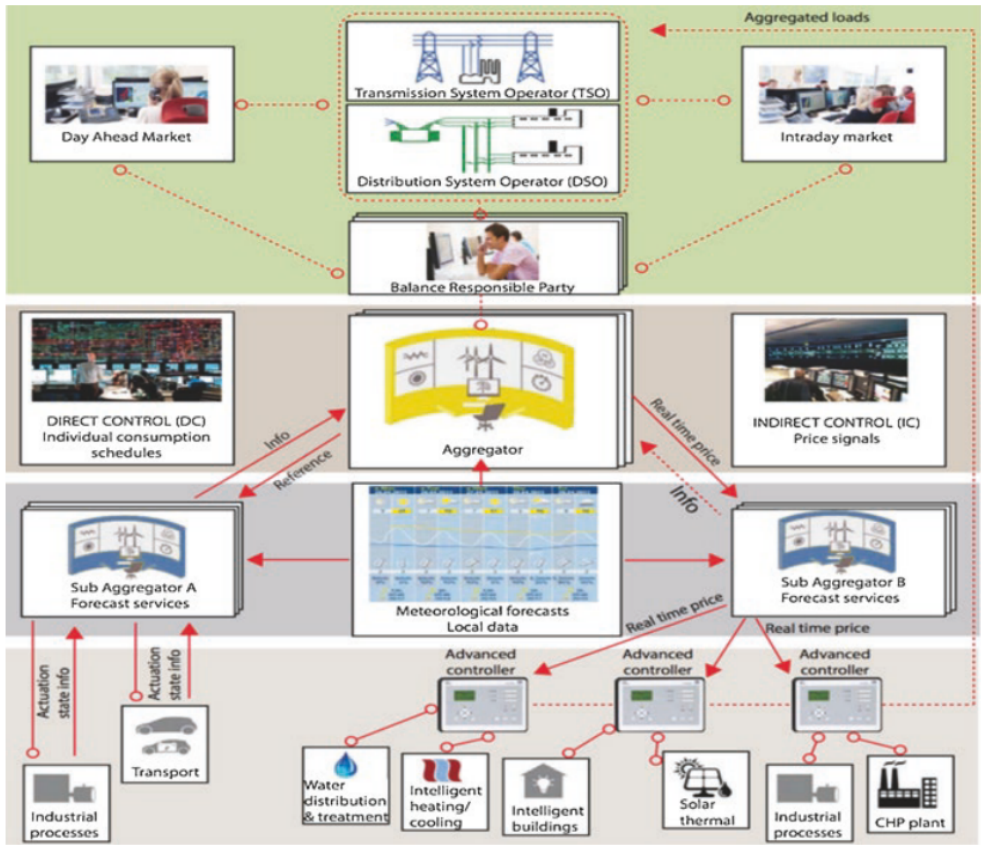


Figure 2.1: Workflow for Smart-Energy Operation-System (Figure from [12]).

2.2 Distribution system operation and local grid issues

Distributed generation (DG), such as wind and photo-voltaic (PV) panels, are typically placed in the power distribution grids. This does not only mean increasing levels of stochastic production in the distribution grids [27], but potentially also bidirectional power flows. Increasing customer load, for example through increasing levels of electric vehicles (EVs), further amplifies the already stochastic nature of customer load behaviour. Distribution power systems were originally designed for unidirectional power flows from higher to lower voltage levels. The operation has also been passive and traditionally, the grid development strategy has been to 'fit-and-forget'

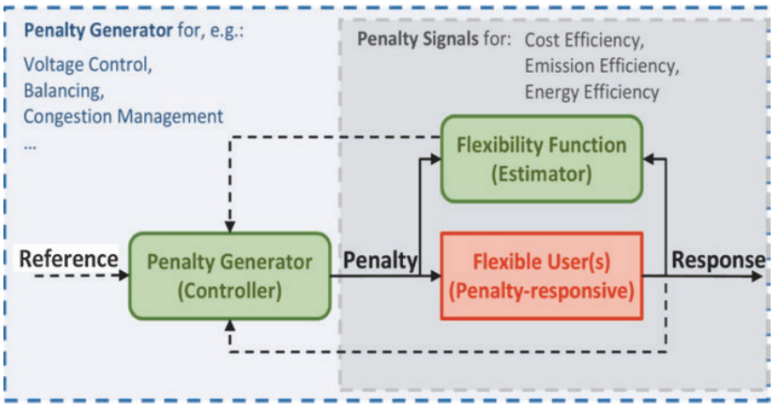


Figure 2.2: Penalty signal and flexibility function in SE-OS with a power systems focus (Figure from [12]).

[10]. The developments in generation and consumption, now compromises various concerns for DSOs and calls for new methods for distribution grid operation [11]. For instance, distributed generation introduces harmonics affecting the power quality and can also cause bidirectional power flows, whereas voltage drops and rises might be a consequence from more volatile generation and consumption [10, 7, 8, 9]. Increasing fault levels and congestion in distribution grids are also potential concerns and consequences if adaptive measures are not taken by the DSOs.

2.2.1 Low voltage distribution grids

The research in this thesis concerns low voltage distribution grids, also known as secondary distribution systems. The low voltage system distributes the electricity from transformers connected to the medium voltage grids to cable cabinets, to which the customers are connected, or alternatively directly to customers metering devices. The transformers are in general small and the neutral-to-line voltage is normally 230 V (i.e. 0.4 kV line-to-line) in Europe. The distribution networks are typically radial grids (subsection 2.3()), meaning there is only one path for the power flow [28], as opposed to meshed grids (2.3(a)) commonly seen at the higher topological levels in the power system. Furthermore, the low voltage distribution systems are characterized by a low X/R ratio (reactance over resistance), and unbalance between the phases since the various customers loads are directly connected to the different phases.

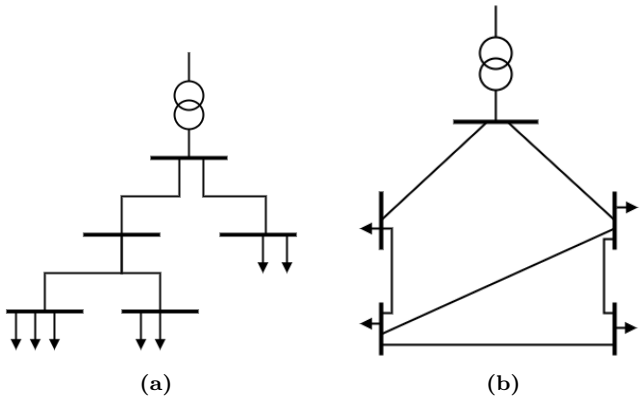


Figure 2.3: Example drawings of radial (a) and meshed (b) power grids.

2.2.2 Transformer congestion and temperature dynamics

Congestion occurs when the loading exceed the limits of grid equipment such as cables or transformers. Power transformers are one of the most costly grid equipment in the power system [29], and it it thus of great concern that they are safely operated to avoid equipment failure or reduction in life expectancy.

It is a standard practice to operate power transformers based on their nameplate ratings [30]. For each transformer, the nameplate provides the working conditions within which the transformer can be safely operated, without compromising its ageing and life expectation. These ratings have traditionally been set as static limits that have been developed in controlled environments with conservative margins. For instance the ambient temperature at which the transformer is rated is normally higher than the actual ambient temperature, especially in the Nordic countries [31]. During transformer operation, the conditions might be more advantageous than the static rating conditions, leaving unused capacity to be extracted. If instead adapting the rating to the actual working conditions this extra capacity can be unlocked, and this concept is known as dynamic transformer rating (DTR) [32, 33], which is similar to the concept of dynamic line ratings (DLR) [34].

The limiting factor for loading a transformer beyond the nameplate rating is the temperature, both in terms of operation and thermal ageing. Loading a transformer beyond the nameplate rating increases the leakage flux both to the core and outside, which heats metallic parts of the transformer. This can cause higher temperatures of windings, oil and core, while stressing components such as tap-changers and cables, potentially beyond their designed margins [35, 36]. Additionally, the composition of the insulation oil might change and gas content increase [35]. It is thus of importance

to have real-time monitoring of the transformer temperatures to ensure safe operation [32].

As explained in Annex F in standard IEC 60076-7:2018 [35], the thermal conditions within the transformer is driven by heat generated from load (i.e. ohmic winding, core and stray losses), and no-load losses. The thermal system also includes thermal capacitance of oil and winding and is cooled down by the transformer oil and the ambient temperature between which there are thermal resistances. Moreover, IEEE standard Std. C57.110-1998 [36] defines the load loss as the sum of load loss related to I^2R and stray loss, including both winding stray loss or eddy-current loss (P_{EC} in (2.1)), and stray loss in other components (P_{OSL} in (2.1)). It should however, be noted that all of these losses are driven by current, meaning that they increase as the current increases. Eddy-current losses, P_{EC}, are specifically driven by increased harmonic currents. The standard IEC 60076-7:2018 further states that outdoor ambient conditions such as solar radiation, wind and rain might affect the loading capacity of small transformers, but it is not included in the calculation methods in the standard [35].

$$P_{LL} = P + P_{EC} + P_{OSL} \tag{2.1}$$

According to the standard IEC 60076-7:2018 four temperatures under three different conditions are crucial to the transformer operation and lifetime and should be considered for loading over rating. The conditions consider normal cyclic loading, long-time emergency loading and short-time emergency loading and the maximum temperature are seen in Table 2.1 respectively, however not for short-time emergency loading for which the temperature might be hard to control. Further, gas bubbles might be developed in the oil if the hot-spot temperature exceeds 140 $°C$ [35]. The transformer temperature can be monitored using fiber optics or thermal models. The methods developed in this thesis primarily aims to support the grid operator under normal cyclic conditions. However, insights for long-time emergency loading could be of interest (note that methods would need to be tested and validated under such conditions).

Temperature in $°C$:	**Normal cyclic loading**	**Long-time emergency loading**
Winding hot-spot	120	140
Other metallic hot-spot	140	160
Inner core hot-spot	130	140
Top-oil	105	115

Table 2.1: Maximum temperatures for small transformers from IEC standard 60076-7:2018 [35].

2.2.3 Voltage estimation/state estimation

Keeping the voltage stable in a power grid is of major importance to avoid system or equipment failure. As stated in Section 2.2.1, DSO LV grids are most commonly radial grids. Assuming we have a simplified radial as in Figure 2.4 (with Thevenin equivalent as in Figure 2.5) the relation between sending-end voltage (V_S) and receiving-end voltage (V_R) is described by:

$$V_S = V_R + \Delta V \tag{2.2}$$

where the voltage drop along the radial, ΔV, is described by the following equation from IEEE Standard 141-1993:

$$\begin{aligned} \Delta V &= V_S + IR\cos\theta + IX\sin\theta - \sqrt{V_S^2 - (IR\cos\theta + IX\sin\theta)^2} \\ &\approx IR\cos\theta + IX\sin\theta \end{aligned} \tag{2.3}$$

where I is the current as in Figure 2.5, R and X are the resistance and reactance of Z_L in Figure 2.5 (i.e. $Z_L = R + jX$), $\cos\theta$ is the power factor of the load and $\sin\theta$ is the reactive factor of the load [37]. The second line of Eq. (2.3) represents an approximation of the voltage drop which is commonly used for engineering purposes [38]. If the X/R ratio is low, as it typically is in low voltage grids, it might be possible to simplify the equation further by neglecting the second term of the approximation of Eq. (2.3) (i.e. $IX\sin\theta$).

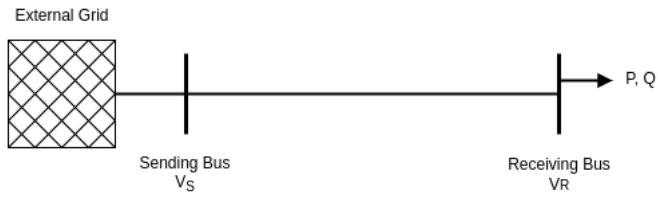


Figure 2.4: Assumed/simplified radial for voltage drop equations. V_S and V_R represent sending-end and receiving-end voltages, respectively.

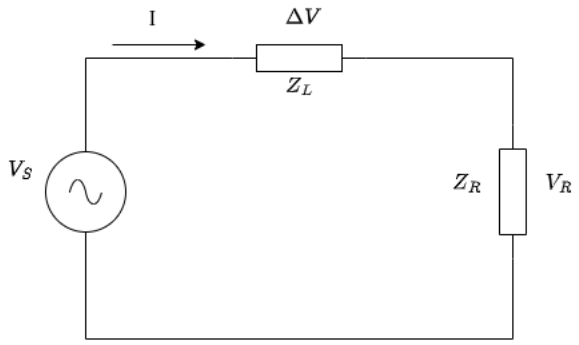


Figure 2.5: Thevenin equivalent of Figure 2.4, where Z_L is the impedance of the line (such that $Z_L = R + jX$, in Eq. (2.3)) and Z_R represents the load in Figure 2.5.

Looking at Eq. (2.3) with Figure 2.4 and Figure 2.5, it is realized that a higher load will lead to a larger voltage drop and lower receiving-end voltage. A lower or negative load (for example through PV generation) will on the other hand lead to a smaller voltage drop or even voltage rise and therefore a higher receiving-end voltage.

The concerns for a DSO is understood by studying Figure 2.6, describing the relation between receiving-end over sending-end voltage versus active power at receiving-end over short circuit capacity. If the receiving-end voltage is too high or too low it will cause the power to drop drastically (system failure). With the recent developments in demand behaviour and increasing levels of EVs, such drastic voltage drops could for instance be caused by:

- many EVs being plugged in and charged at the same time.
- other and many loads occurring simultaneously, such as heat pumps.
- many DERs providing frequency down-regulation at the same time.

The listed events all result in a large I in Eq. (2.3), meaning a larger voltage drop along the radials. On the other hand drastic voltage rise could occur through:

- significant PV or other DG generation (i.e. negative I in Eq. (2.3)).
- many DERs providing frequency up-regulation (meaning very low I in Eq. (2.3)).

It should be noted that these voltage drops and rises will occur if the power grid is not dimensioned (meaning Z_L in Figure 2.5 adjusted) for these changes in consumption and DG in LV grids or if no adaptive measures are taken, for example by utilizing flexibility (or adapting flexibility algorithms). Utilizing flexibility to mitigate voltage

issues in distribution power systems will potentially have a more efficient impact than reactive power control due to the low X/R ratio. Due to this behaviour it is important that voltage adjusting flexibility is locally prioritized above frequency regulating flexibility. To avoid theses hazardous situations, the TSOs normally set voltage limits, and in Denmark the limits are $230V \pm 10\%$ [39].

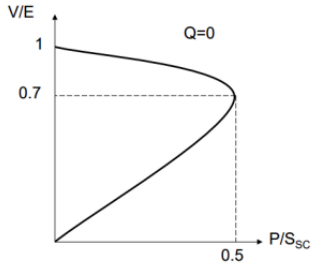


Figure 2.6: The ratio between receiving-end over sending-end voltage versus active power at receiving-end over short circuit capacity, presented as Figure 3 in [40].

For the described reasons it is becoming increasingly important for DSOs to keep track of the voltage. In transmission power systems, state estimation has been used by TSOs for decades to estimate voltages and other states. However, due to the commonly low X/R ratio in distribution systems, the DC state estimation techniques and simplifications are not applicable [41, 14]. Unbalance between the phases in distribution system further limits the applicability of conventional state estimation methods [14] and DSOs need other real-time 3-phase state estimation methods convert to an active grid management [28, 42]. Other important considerations for DSO state estimation are cyber-security, GDPR (if using data from smart meters), data flow architecture/infrastructure and bandwidth, more frequent network configurations (replacing or installing cables, connecting new loads/customers) and increasing number of states [14, 42].

2.3 Towards adaptive/dynamic DSO smart grid operation - identified gaps

The SE-OS framework offers potential benefits for DSOs, such as local ancillary services solutions in distribution grids, to solve potential local grid issues discussed in the previous section (Section 2.2). Through the multi objective price signal generator in Figure 2.2 the objective can be set to provide various ancillary services at different topological levels in the grid. While frequency control management and signals can be

optimized and generated at a higher aggregator layer, voltage, congestion and other local distribution system issues can be solved by a local aggregator at a lower layer [12]. Thereby, local ancillary services can be provided for the DSO at the distribution system layer in the SE-OS.

2.3.1 Aggregator layer

However, to utilize this potential measures need to be taken from both aggregator and DSO sides. Following the description of voltage deviations and transformer congestion in Section 2.2, it is understood that there is a spatial factor to the need for flexibility in distribution grids as opposed to flexibility provided for, e.g., frequency control. This is addressed by the SE-OS framework by defining different spatial-temporal resolutions in the different layers [12]. For aggregators to enter this framework with customer pools that can deliver the required flexibility, the uncertainty of demand response (DR) needs to be addressed in both spatial and temporal scales [17, 18]. In [43], the authors found that the flexibility potential varies for different sectors in Northern Europe. Therefore, the aggregators need probabilistic price elasticity models reflecting the various customers types. This is firstly needed to identify a customer pool (in quantity and types of consumers) in distribution grids that can fill the DSOs requirements for flexibility at different times. Secondly, it is required to generate an appropriate and effective price signal to the customer pool. That the price elasticity models are probabilistic are important due to the stochastic behaviour of DR [17].

Through the extensive communication with DSOs that has been ongoing throughout the work of this thesis, it has been understood that the uncertainty of DR is a barrier or concern for DSOs to incorporate flexibility in their operation strategy. This further highlights the need for probabilistic price elasticity models for aggregators, such that the uncertainty can be communicated to DSOs. If the aggregator's communicated uncertainty/certainty matches the expectations and requirements from the DSO's side, then flexibility can be traded with to solve local grid issues.

Several papers can be found on short-term price elasticity for electricity consumption (e.g [44] and [45]). However, many of the presented models in these studies are deterministic and not applicable in the SE-OS framework. There are, however, probabilistic models presented in the literature. For instance, [46] presents a DR model using conditional probability density and [47] presents a quantile linear-regression model. Nevertheless, the applicability in the context of the SE-OS framework is limited, as the former is based on two-way communication, whereas unidirectional is adapted in the framework, and the latter considers single residential customers ,which limits the aggregators ability to explore the aggregated customer pool behaviour with various types of customers. Furthermore, it is important that an appropriate scenario

is assumed for the model development. For instance, [48] found that automation strongly increases the price responsiveness. To fit the model development to the SE-OS framework, a future scenario was adopted in which customers are equipped with Energy Management Systems (EMS). For the scenario adopted, limited research and data was available. For instance, in [47] the authors assume customers having home energy managements systems, but limitations in the context of this work has already been stated. For a more detailed literature review on price elasticity models the reader is referred to [49] (also attached as paper A in this thesis).

2.3.2 DSO layer

To be able to adapt to new grid conditions and take advantage of ancillary services provided through the SE-OS framework, DSOs need new methods to support their grid operation. As previously mentioned in Section 2.2, DSO grids have traditionally been characterized by passive operation [10]. To adopt an active strategy, increased observability is required as a first step. Without information on the current state of the grid it will be difficult for the DSO to know when and how much flexibility to request. This has increased importance as the levels of DGs, heat pumps and EVs are increasing in the distribution grids and an increased observability is required to be able to ensure safe grid operation and power quality [10, 50].

The DSO methods for increased observability needs to be data-driven and risk informative as well as being able to deliver model estimation and forecasting in real-time. This is especially understood from the SE-OS framework, as control signals and forecasts are updated for each time step and data is collected from devices to ensure synergies between the markets mechanisms and physical system. That the methods are able to deliver forecasts and risk information, such as prediction intervals, are important for the DSO to be able to take actions ahead in time though the SE-OS to avoid any predicted grid violations.

In this thesis methods for increasing the observability of both transformer congestion (temperature) and system voltage states are developed in paper B and C. In the literature, research papers can be found aiming to increase the observability, but has limited applicability in the SE-OS framework. For instance in [32, 51], thermal models for transformers are developed for monitoring or design purposes, but are not proven to have forecasting capabilities, required in the SE-OS framework. Instead e.g. [52, 53] propose transformer thermal models for predicting the temperature, but the models either require input data that cannot be practically measured or does not consider the impact from environmental surroundings, which small transformers are affected by.

Furthermore, there is a need for adaptive and efficient feeder modeling for state estimation, as highlighted by the authors in [54]. Concerning state estimation to

increase the observability for DSOs, the authors in [42] requests more real-world implementations and more accurate models as 1% observation error can lead to a 10 % state estimation error. Alternatively, models enabling uncertainty management could also be useful.

Grey-box modeling offers a method to derive models that are both physics-informed and data-driven [55, 56]. The models are useful for providing estimates and predictions as well as associated uncertainty of dynamic and stochastic system behaviour [24, 25]. These types of models are a mixture between deterministic modeling approaches (white-box models), purely relying on theoretical knowledge, and black-box modeling approaches, exclusively using statistics and data to derive the model structures (Figure 2.7). In the context of SE-OS, grey-box models have proven to be useful and are thus explored in paper B and C.

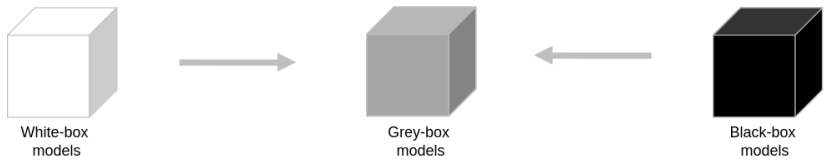


Figure 2.7: Concept of grey-box modeling.

Furthermore, the DSOs need to adapt the observability models in their operation strategy. Therefore, this thesis focuses on developing an adaptive DSO smart grid operation strategy in the context of SE-OS. This would not only offer the utilization of DER to solve local grid issues, but also to use data-driven approaches to fully utilize grid equipment. The recent digital development of better communication infrastructure and cheaper devices offers new ways to operate the grid to its full potential, delaying the replacement of grid equipment [29] and potentially increasing the electricity distribution at certain times. The latter should be of critical interest for DSOs as they might be able to safely transmit more electricity without the long procedure of replacing equipment. For this purpose the DSO operation framework is developed and addressed in paper paper B and C.

2.4 Chapter Summary

As seen in the literature many electricity price elasticity models are deterministic and adapts scenarios with little applicability in the SE-OS framework. For aggregators to successfully explore price elasticity in DR for various customer types to ensure

a customer pool capable of supporting the DSOs need for ancillary services as well as supporting the communication between the two actors different approaches are required. To fill this gap an approach is adapted assuming a future scenario, in which customers have EMS. Furthermore, the approach explores variations and uncertainty of DR from various customer types allowing the aggregator to adapt its customer pool and business strategy.

Moreover, several works found in literature aiming to support DSO grid operation are either simulation studies, does not report enough information about uncertainties, relies on measurements that are unpractical to measure at scale and cannot be utilized for forecasting purposes. To design useful methods for adaptive DSO smart grid operation in the context of SE-OS, special focus is given to:

- Physics-informed and data-driven models (grey-box models), to enable real-time monitoring and forecasting of the system.
- Models capable of providing forecasts and information of associated uncertainty.
- Using available data as well as data inputs that are practical to measure with a setup that is scalable in DSO grids.

CHAPTER 3
Experimental setup

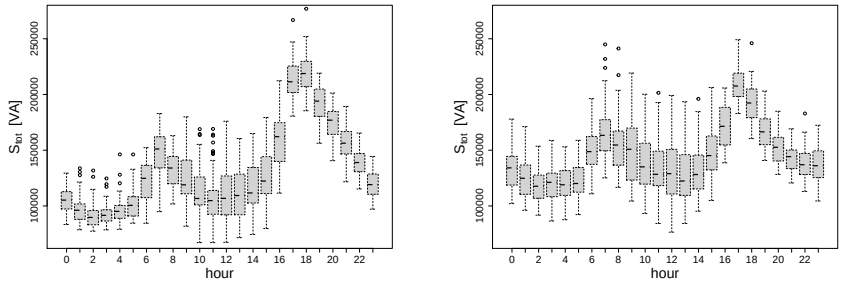
This chapter presents the experimental setup and data that is used to develop monitoring and forecasting tools for DSOs.

3.1 Experimental setup - living lab description

The research in this thesis has been closely related to the Flexible Energy Denmark (FED) project [57]. FED is a Danish project focused on enabling flexibility in the power system through digitization. Through FED two living labs (LLs) from a DSO in Jutland, Denmark, were established [58]. LLs here refers to real-world grids that are used for testing new technologies. The two LLs compromise two LV radial grids at 0.4 kV, with 170 residential customers (this LL is hereafter referred to as LL1 or grid 1) and 140 residential customers as well as a small industry (this LL is hereafter referred to as LL2 or grid 2), respectively. In each LL there are 5 - 10 customers with electric vehicles (EVs) and 15 - 20 customers with photovoltaic (PV) panels (rated sizes of 3 - 6 kW). The heating of the households occurs through a mixture of heat pumps and district heating.

Both LLs are connected to the external grid with 3-phase 10/0.4 kV oil cooled transformers, where the transformer in LL1 (hereafter transformer 1) is rated at 400 kVA and the one in LL2 (hereafter transformer 2) is rated at 200 kVA. Both transformers are installed in ventilated metal housings, placed outdoors. The apparent power profile for each transformer is visualized through the boxplots in Figure 3.1. (Note that the apparent power is chosen as it is rather the apparent power than the active power that is relevant for the loading or rating of the transformer.) In the figure it can be seen that the peak loads of both transformers are in the range of 200 to 250 kVA. Transformer 2 seems to have a relatively larger base load, compared to transformer 1, which could be because of the industrial load in LL2.

In the LLs 22 electronic measuring devices have been installed. In Figure 3.3 and in Figure 3.4, it can be seen how the devices are distributed in the two LLs. Since only 22 devices were available, it was decided to install them on three out of the five feeders in each grid, to gain enough observability at these feeders for the model development in paper B and C. At the transformers all feeders as well as temperatures



(a) Transformer 1 **(b)** Transformer 2

Figure 3.1: Boxplots of apparent power for both transformers using data from November 2021.

are measured and therefore two devices were required. The LLs are real-world grid and thus, they can not always be forced into ideal situations often assumed when conducting simulation studies. For instance, in this setup we aimed to have one device at one middle node along the radial and some devices at the end-nodes of the feeders to measure both voltage and current of ingoing and outgoing cables. Due to space limitations in the cable cabinets this was not always possible. Table 3.1 indicates which measurements are delivered by each of the devices, either voltage and current measurements or alternatively only voltage. If current measurements are available, measurements of active power, power factor and harmonic distortion are also available. All data is delivered per phase and once per second (neutral phase data is only available for some of the devices). Furthermore, all impedances of the cables as well as cable lengths are known.

3.2 Modeling approach based on experimental setup

As previously stated, the setup of the two LLs were designed throughpaper B andC in this thesis. However, the real-world reality in LLs put constraints on the setup design, that simply had to be accepted. Therefore, the developed methods reflect a real-world scenario, rather than an ideal case that are often seen in simulation studies, but sometimes not in the real-world. The approach for the model development processes was to use the available data that could practically be measured, and it was inevitably achieved through the LL setup design.

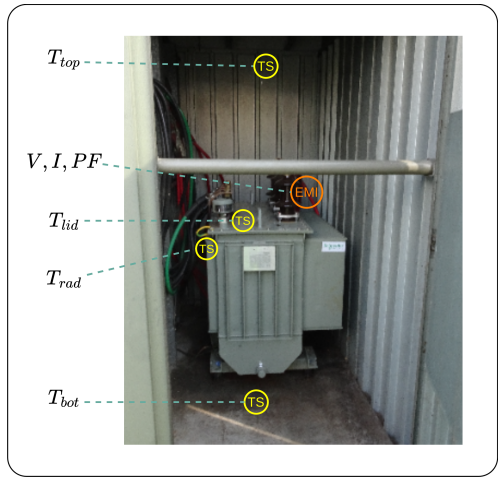


Figure 3.2: Experimental setup for the transformers, with temperature sensor (TS) and electronic measurement instruments (EMI). Figure is from paper B.

Device	Voltage	Current	Temperature
Grid 1:			
$T1_1$,$T1_2$	X	X	X
BR_M	X	X	-
BR_E	X	X	-
G_M	X	X	-
G_{E1}	X	X	-
G_{E2}	X	-	-
G_{E3}	X	-	-
B_M	X	X	-
B_{E1}	X	X	-
B_{E2}	X	X	-
B_{E3}	X	-	-
Grid 2:			
$T2_1$,$T2_2$	X	X	X
BL_M	X	X	-
BL_{E1}	X	-	-
BL_{E2}	X	-	-
Y_M	X	X	-
Y_E	X	X	-
R_M	X	X	-
R_{E1}	X	-	-
R_{E2}	X	X	-

Table 3.1: Devices installed in the LLs (in Figure 3.3 and Figure 3.4), X indicates that voltage or current measurement is available. If current measurements are available, data for harmonic current content, active power and power factor are also available.

3.3 Data Cleaning Pipelines

To develop the methods in this thesis various data sources have been used.

Meter data – electrical meter data from the DSO was available in parquet file format that was sorted through spark SQL. The data included both active and reactive power measurements in a 1 hour time resolution.

Grid topology data – was available in several excel files and were sorted in python, to build a grid model using pandapower [59].

Measuring device data – the data from the installed measuring devices were available in json format and were fetched from a third party SFTP server. The devices deliver per phase data once per second.

Meteorological data – was fetched in json format from the open data from the Danish Meterological Institute (DMI) [60], using their API [61]. The time resolution was 10 minutes.

To fetch and clean this data a set of data cleaning pipelines were written in python. One pipeline was written to fetch the measuring device data from an SFTP server. The pipeline fetched and extracted relevant measurements for defined devices and dates automatically, through small data sets that were later combined. The pipeline also filtered the data sets to different time resolutions and combined them with relevant meteorological data for model development (note that this was only possible for 10 minute time resolutions).

The data from the measuring devices and meter data was further analyzed to ensure a correct installation. For this purpose another pipeline was written. The pipeline first constructed a graph (the data structure) representation of the LV network using Networkx graphs in python [62] and the grid topology excel files. The meter data could then be sorted and combined for different feeders, or even section of feeders, by cutting edges between relevant nodes in the graphs. It proved to be an efficient way of sorting customer electrical load data for various spatial scales in the grid. Through analyzing the data from the meters and the measuring devices for different feeders and section of feeders, a correct installation could be ensured. Through the pipeline and data analysis, incorrect installations in the setup were identified and fixed. Furthermore, the analysis showed that the reactive power data recordings from the meters contained so many erroneous measurements that the data was unusable. Furthermore, time shifts between data sets were also identified.

A simplified illustration of the data bases and pipelines is depicted in Figure 3.5.

3.4 Chapter Summary

In this chapter it has been explained how the experimental setup was designed through the thesis project and collaboration with the DSO and FED. The real-world conditions put constraints on the preferred installation of the 22 electrical measuring devices due to space limitations in the cable cabinets. Nevertheless, as the modeling approach is to use the data available to construct models that are based on a setup that is practical and achievable for the DSO, this was accepted.

Available data sets were also outlined and a couple of data cleaning pipelines described. The pipelines fetched data using APIs, cleaned and filtered the data to relevant time resolutions. The main contributions can be summarized as follows:

- Designing a real-world setup with 22 electrical measuring devices for two LV grids.
- The design includes a specific installation setup for temperature measurements at the transformers.
- Writing data cleaning pipelines fetching data from different sources of data to output relevant data sets for model development and analysis.
- Identifying incorrect installations through data analysis of the cleaned data sets.

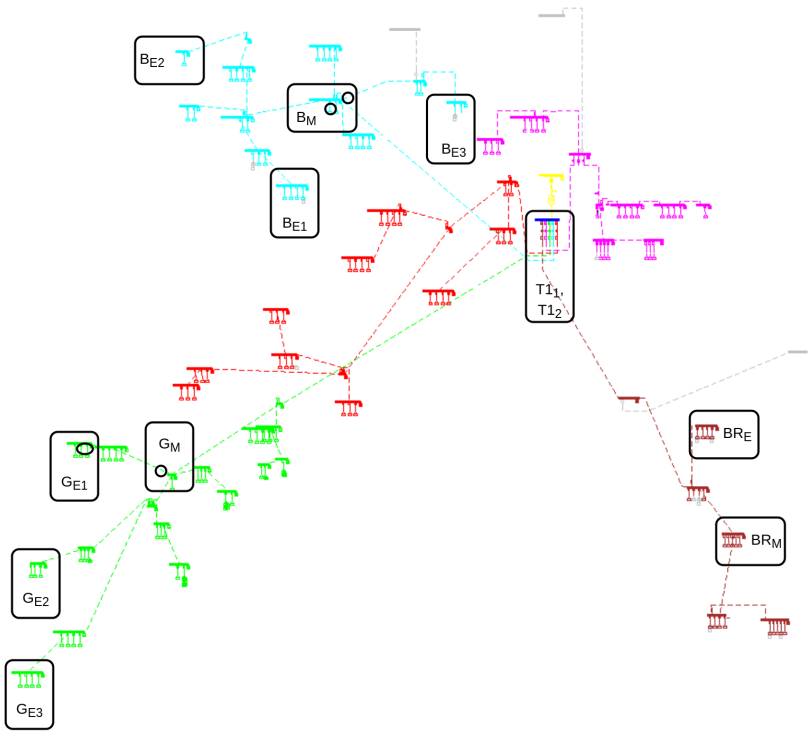


Figure 3.3: Grid topology and installation of devices for grid 1/LL1. The name of the devices is constructed from the first letter(s) of the color of the feeder, while subscript M indicates that it is a middle node, and subscript E indicates an end-node. Figure is from paper C .

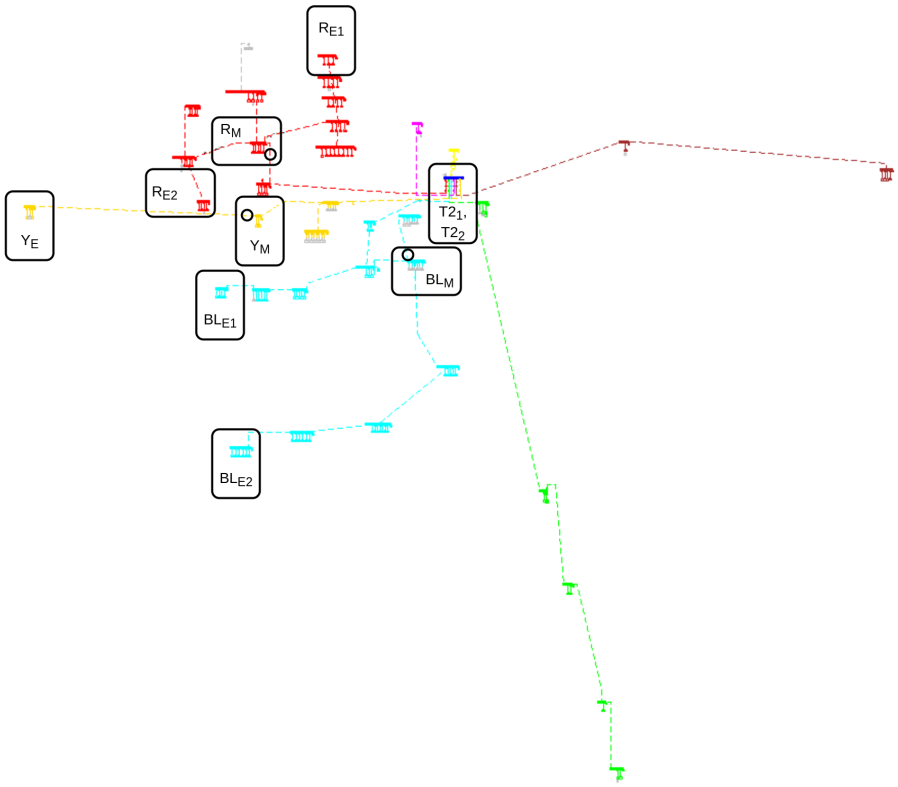


Figure 3.4: Grid topology and installation of devices for grid 2/LL2. The name of the devices is constructed from the first letter(s) of the color of the feeder, while subscript M indicates that it is a middle node, and subscript E indicates an end-node.

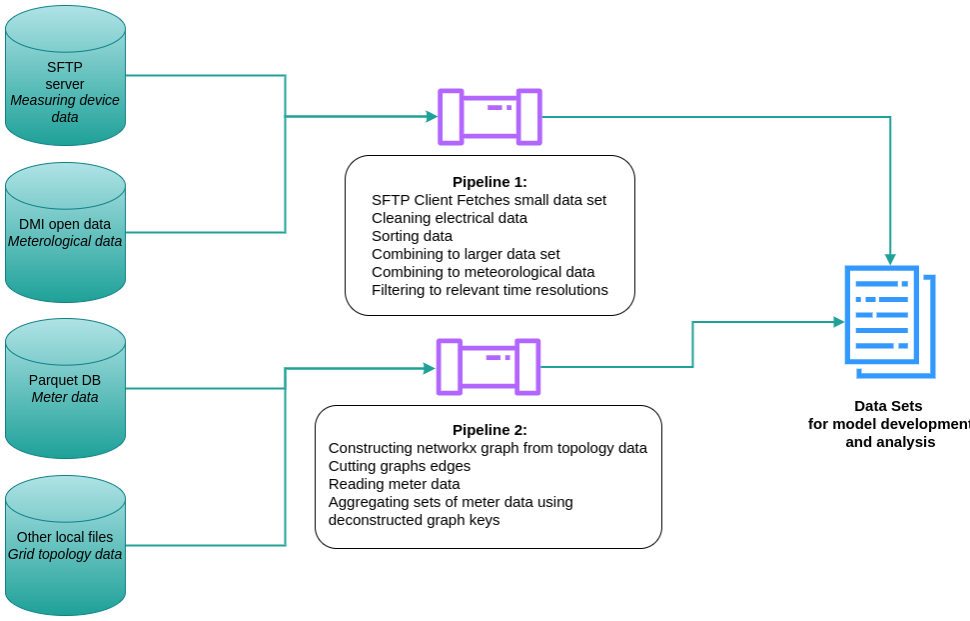


Figure 3.5: A simplified illustration of data bases, data flows and pipelines to construct relevant data sets for model development.

CHAPTER 4
Aggregators in SE-OS

The aggregators are key actors in the SE-OS framework. In the framework their task is to aggregate the DR at various spatial-temporal resolutions to provide ancillary services at different layers in the framework (Figure 2.1). This chapter presents and motivates methods for aggregators to build useful customer pools by developing a model to describe the price responsiveness across different sectors using a probabilistic approach. The chapter further discusses the results and contributions of paper A in the context of SE-OS to provide ancillary services for DSOs.

4.1 Considerations towards useful price response models for aggregators

Following the presented background in Chapter 2, there are several important considerations when modeling the price responsiveness of DR.

As the SE-OS involves spatial-temporal layers covering the entire power system, there will be many and diverse types of potential customers for the aggregator. Meanwhile, the flexibility potential has proven to vary across different sectors [43]. Therefore, the aggregator needs information on the flexibility potential for various types of customers to explore the flexibility of customers pools in various DSO grids. Since the aggregator is supposedly interested in the aggregated flexibility potential, the aggregators need information about the DR from aggregated consumers, as opposed to single customer behaviour with little relevance in this context [16]. Thus in paper A, focus is given to modeling aggregated response from various types of customers.

In the SE-OS indirect control through unidirectional communication is commonly applied, often using price signals as control mechanism. It is thus useful for the aggregator to investigate the price elasticity. Values of estimated price elasticity varies widely in the literature. For instance, in [63] 538 observations of price elasticity were found and the authors in [64] found that the elasticity varied between -0.2 and -0.8 for data sets representing residential consumers in the US. Meanwhile, the estimated price elasticity seems to depend on assumed scenarios and other author assumptions. Gillan, assuming explicit DR in [44], found different values for price elasticity for

different changes in price and furthermore an increased price elasticity for automated DR, for which a reduced effort cost was also highlighted. Increased price responsiveness for automated customers was also seen in [48].

Hence, to apply a relevant scenario for the price elasticity model in the context of SE-OS, paper A assumes a future scenario, in which consumers are equipped with Energy Management Systems (EMS) [65]. Furthermore, the various values of price elasticity found in literature indicates that deterministic point estimations might not be sufficient. Instead we develop a probabilistic model allowing for non-linear price responsive behaviour. Due to the flexibility of such a model, it is also assumed that the model would be applicable to slightly varying scenario assumptions and DR behaviour.

Due to the unidirectional communication flow giving the customer the choice to choose if they would like to adapt their consumption, the response from the flexible resources is not as certain as it is in two-way communication approaches, in which the DER commits to providing flexibility. Therefore, probabilistic models are of major importance for aggregators to describe the uncertainty of DR in this context. Such models are also required by the aggregator to manage the uncertainty and apply risk informed operation strategies. For instance, the aggregator could apply conditional value at risk (CVaR) or scenario-based stochastic programming [66, 67]. In this case, deterministic point estimations are insufficient as input to these methods and thus paper A applies a probabilistic price elasticity model to support these types of operation strategies.

The objective of the research in relation to the aggregator can be summarized as follows:

- Develop a model to describe price responsive behaviour of flexible consumers.
- Adapt a relevant future scenario for the consumer behaviour applicable to the SE-OS.
- Provide a probabilistic output from the model, than can be used to 1) estimate a customer pool capable of providing ancillary services to DSO 2) communicate the uncertainty of the response to the DSO.

4.2 Data Presentation

Due to scarcity of available experimental data representing DR consumers equipped with EMS, synthesized data was used to develop the price elasticity method in paper A. The synthesized data is from [65, 68, 69] and is based on data from a Danish project collaboration between the Danish TSO, Energinet and Dansk Energi (the Elforbrugspanel project). The data set involves 29 customer categories, with different

price responsive characteristics, to account for various customer types across different sectors. 70 unique customers of each category are then exposed to 1000 price-sets, consisting of hourly electricity prices for one day. The electricity price is constructed from a baseline price of 2.25 DKK/kWh and an hourly variable price component, π_h. The total price is consequently $2.25 + \pi_h$ DKK/kWh and the variable component has a cap such that $\pi_h \in [-0.75, 0.75]$. Moreover, rebound is also accounted for assuming a static rebound effect as described in [68]. A more detailed description and a visual example of the data are found in [49] (also attached as supplementary information to paper A in this thesis). In paper A customers from different customer categories are aggregated to residential, light industry and heavy industry clusters (this is further described in [49].

4.3 Probabilistic price elasticity modeling of DR

This section presents applied methods relevant for the probabilistic price elasticity model in paper A.

4.3.1 B-splines

B-splines offers a flexible method of data fitting, allowing to model non-linear behaviour using a linear model structure. B-splines consists of pieces of polynomials of degree p, that meet on x-coordinates called knots [70]. The knots, or knot vector, are a vector of non-decreasing values along the x-axis, i.e. $\mathbf{v} = [\xi_1, ..., \xi_k]$, where $\xi_i \leq \xi_{i+1}$ and k is the number of knots [71]. If the knots are equidistantly spaced the B-splines are said to be uniform B-splines [70]. Each polynomial spline function is defined such that:

$$B_{j,p}(x) = \begin{cases} f_p(x), & \text{if } \xi_j \leq x < \xi_{j+1}. \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

where $f_p(x)$ is a polynomial of degree p and $B_{j,p}$ is the jth spline function of a B-spline with degree p [71]. The total B-spline formula, $S(x)$ can then be described by:

$$S(x) = \sum_{j=1}^{N} a_j B_j(x) \tag{4.2}$$

where N are the number of individual spline functions. A property of B-splines is that its integer values (of $B_j(x)$) sums to unity [70]. Fitting the B-splines to a data set is thus about optimizing the coefficients, a_j, for the individual spline functions.

Since the regression had to be done for multiple clusters for several hours, B-splines proved to have the flexibility not requiring to manually calibrate the regression for each data set. Other methods were considered, for instance linear regression, polynomial regression, and I-splines. However, B-splines was preferred to gain the flexibility

and applicability to multiple data sets, while also offering method extensions, such as monotonically non-decreasing P-splines, as outlined in Section 4.3.3.1.

4.3.2 Distribution of DR

In engineering modeling design, errors are often assumed to be Gaussian. The problem when modeling the aggregator customer pools in the context of providing ancillary services for DSOs is that the number of customers in each LV grid is generally too small to consider the assumption applicable. In 1978, Koenker highlighted an over exaggerated belief in the Gaussian law of error when estimating models to data, instead suggesting a minimization problem to generate quantile regression models, at the time with comparable efficiency to least squares optimization for Gaussian linear model [72].

The maximum customer pool sizes to provide local ancillary services, such as voltage control or congestion management, in the grids studied in the thesis are 170 and 140 customers each (see setup description in Chapter 3). The distribution of a hypothetical DR from the customers in the grids, can therefore not be assumed to be Gaussian. Instead quantile regression is applied where knowledge of the a priori distribution is not required. Additionally, it is seen in the data that is used to build the price elasticity model that a Gaussian description of the distribution is insufficient [49].

Instead it is concluded that the magnitude of the flexibility given a certain hour and price signal can be described by some unknown probability density function (pdf) with a cumulative density function (cdf):

$$P(L_h|\pi_h) = \int_{-\infty}^{L_h} f(x;\pi_h)dx \tag{4.3}$$

In paper A quantile regression method is thus applied to allow for and model various cdf's in the data.

4.3.3 Quantile regression

In quantile regression, no assumed a priori distribution is required [73]. The modeling approach rather estimates the model for a quantile, q_τ, i.e. a model for observing a set of observations, e.g. y, with at least τ probability (i.e. $P(y \leq q_\tau) \geq \tau$) [73]. By conducting several such quantile regressions for different values of τ a cdf can be extracted. The parameters of a quantile regression model are optimized by minimizing an asymmetrically weighted absolute error to fit the model to a certain quantile. For more details on quantile regression the reader is referred to [73].

4.3.3.1 Quantile Regression with P-splines

In paper A quantile regression with B-splines is applied such that the conditional quantile given the data for price and response is described by:

$$Q_{\tau,h}(\hat{L}_h|L_{c,h}\pi_{c,h}) = \sum_{j=1}^{n} \hat{a}_{j,h} B_{j,h}(\pi_h; q) \tag{4.4}$$

where $Q_\tau(\hat{L}_h|L_{c,h}, \pi_{c,h})$ is the estimated function for a quantile τ, given the data set C for hour h, and $\hat{a}_{j,h}$and $B_{j,h}$ are the corresponding estimated B-splines as described in Section 4.3.1.

To successfully apply this modeling approach there are some important considerations. Firstly, to avoid overfitting, B-splines with penalties, also known as P-splines, are used. P-splines apply a penalty term adjusting the smoothness of the fitted curve and thereby correcting for overfitting to the data [74].

Secondly, monotonically non-decreasing P-splines are applied. This constraints the method to model a non-positive relation between electricity price and DR. Due to the exclusively negative estimations of price elasticity found in literature (e.g. [63, 64, 44]), this seems to be a valid model modification. Furthermore, the study in paper A assumes customers with EMS, meaning that it is highly expected that a higher price would result in a lower consumption and vice versa. This was found to be a reasonable model modification in paper A. Should, however, the method be applied to another data set, it could occur that the DR does not follow this logic. In such a case another addition or extension to the model should be applied as this behaviour supposedly is not driven by price, but rather by other factors.

Finally, when fitting quantile regression curves to data sets, it might occur that the regression lines cross. Considering the definition of quantiles [73] it would imply a negative probability of the observation between the fitted curves. An example of this is seen in Figure 4.1, where I-splines [75] have been used to fit quantile regressions to the negative price deviations ($\pi_h < 0$) in hour 18 using a data set with 50 customers. In the figure it can be seen how quantiles for τ_2 and τ_3 as well as τ_4 – τ_7 are crossing. The problem of quantile crossings is also highlighted in [76] and [77], where the latter suggests monotonization of the quantile regressions as a solutions. This is another reason for applying monotonically non-decreasing P-splines in the method of paper A.

Another method to avoid quantile crossings is proposed in [78]. In the suggested methodology constraints are applied to the parameters in each quantile regression, such that $\theta_k \leq \theta_{k+1}$, where θ_k are the parameters for the k'th quantile [78]. In paper A this method is utilized by applying the R package quantregGrowth version 0.4.3 [78, 79, 80, 81].

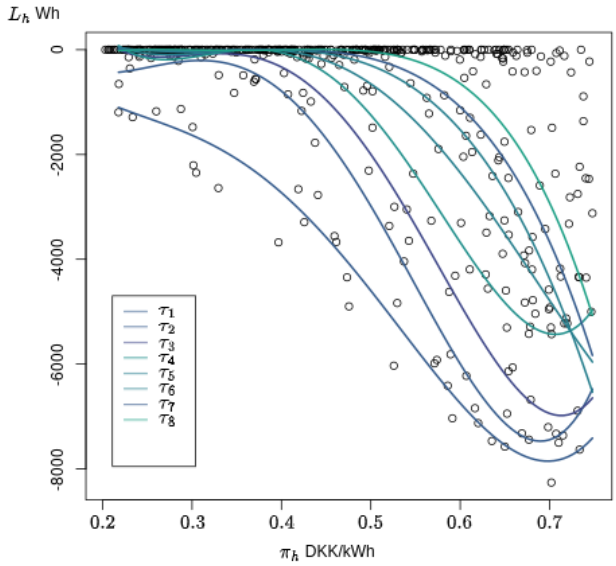


Figure 4.1: Fitted quantile regressions to negative price deviation, π_h in DKK/kWh, at hour 18. The points represent the flexibility magnitude, L_h in Wh, for 50 customers of customer category "House without heating". It can be seen that quantiles for τ_2 and τ_3 as well as $\tau_4 - \tau_7$ are crossing.

4.4 Discussion of Results - insights on DR price elasticity modeling

In paper A the applied method showed a flexible fitting to the data capturing various shapes of cdf's, such as very skewed or bimodal distributions. The variety of distribution functions seen in the fitted models to the data, validates that quantile regression was an appropriate approach, and one assumed a priori distribution function would not have been applicable to the entire data set.

Because of the observed flexibility of describing the cdf's, it is further believed that the method is applicable to other data sets representing other DR scenarios. However, it remains to be tested and validated. Due to the stochastic behaviour and other factors in another data set, it could occur that the response does not follow the assumed logic that an increase in price is followed by a reduction in electricity consumption. One option to apply the method is to loosen the restriction of monotonically non-increasing P-splines. Yet, this solution is not recommended as it firstly,

could cause quantile crossing problems in the quantile regression [76, 77]. Secondly, such a deviation from the logic is probably better described through an addition or extension to the model with relevant inputs for the behaviour, which is the recommended solution. For example, an extension could be to include a variable describing holidays, as these days could affect the price elasticity of DR.

After fitting the quantile regressions to the data sets, cdf's for a certain price deviation and hour could then be extracted. This was done using the fitted quantiles as well as a combination of linear and exponential regression. Thereafter the magnitude of flexibility for customer cluster could be simulated. To demonstrate the usage of the model in paper A, the DR was simulated for different customer clusters using a price deviation of ±0.5 DKK/kWh to achieve ±1 MW of DR for each hour of the day. The simulations were repeated 1000 times and the results are presented in the boxplots in Figure 4.2, Figure 4.3 and Figure 4.4, showing how many customers that were required to achieve a flexibility magnitude of ±1 MW.

Although, paper A focuses on up- and down-regulation for frequency regulation

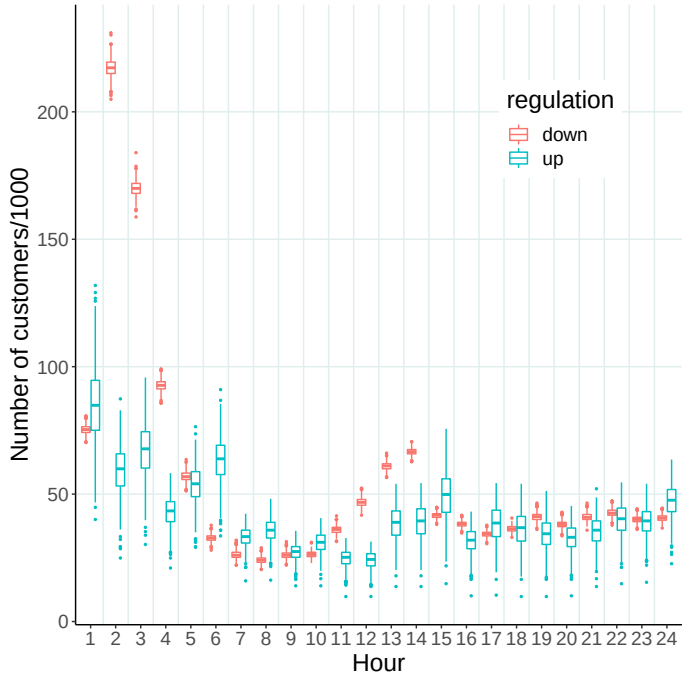


Figure 4.2: Number of activated electricity customers required to achieve ±1 MW in DR from cluster 1 (residential cluster). The figure was presented in paper A.

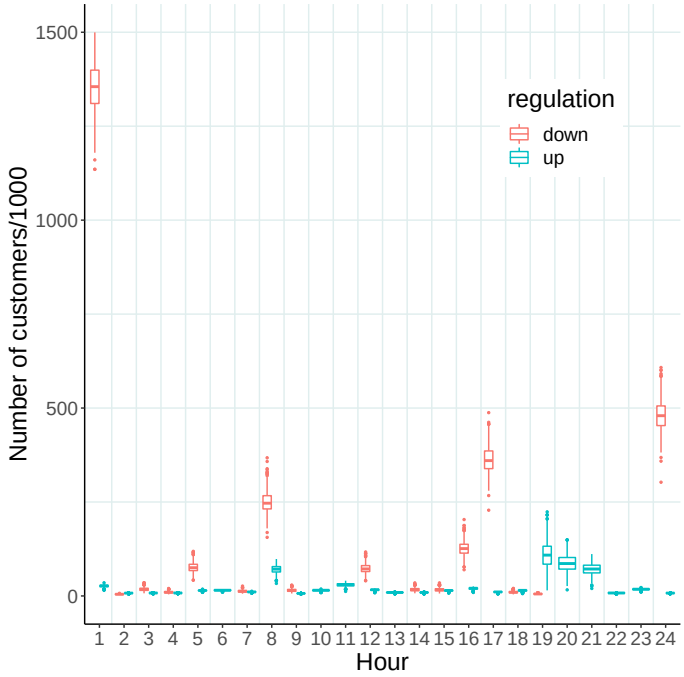


Figure 4.3: Number of activated electricity customers required to achieve ±1 MW in DR from cluster 2 (light industry cluster). The figure was presented in paper A.

when estimating the DR for ±1 MW the results can easily be translated to other ancillary services. For instance, the up-regulation case (i.e. −1 MW DR) would also mean a increasing voltage level if providing voltage control (see Section 2.2.3) or a reduced congestion. In contrast, the down regulation case (i.e. 1 MW DR) would result in down regulation of the voltage level. It should be noted that frequency regulation is more relevant at the upper layers of the SE-OS framework, while the local ancillary services for DSO are more relevant at lower layers.

Given the results, a few examples of insights for aggregator customer pools can be drawn. For instance, a DSO operating a LV grid where PV panels are present, might request voltage down-regulation as an ancillary service. For this purpose the residential customer cluster is a promising candidate to generate the required response. Even though, more customers are required during the day to provide 1 MW of DR, the uncertainty of the amount of required customers is comparably low. Furthermore, less than 75 customers are required, keeping in mind that this would mean approximately less than 50 % of the customers in the studied LV grids. If on the other hand "up-regulation" is required to provide congestion management and there are light in-

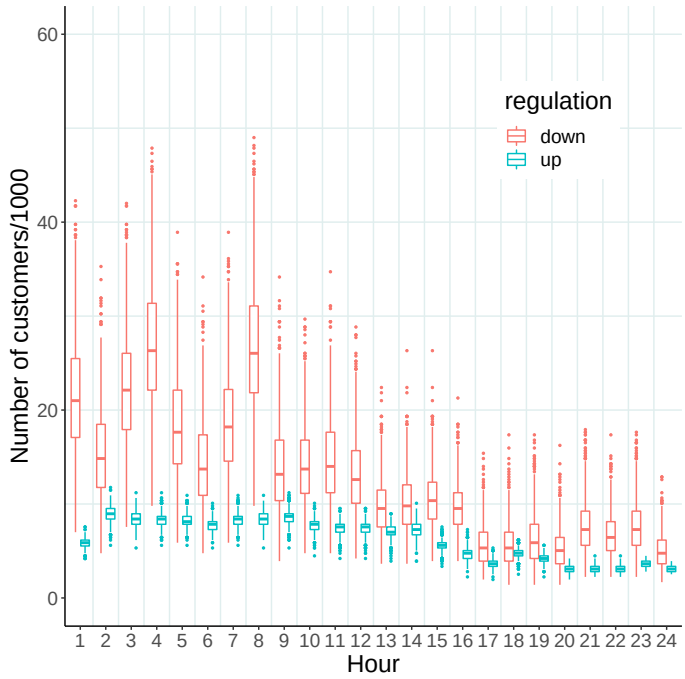


Figure 4.4: Number of activated electricity customers required to achieve ±1 MW in DR from cluster 3 (heavy industry cluster). The figure was presented in paper A.

dustries connected to the LV grid of interest such as in grid 2, it might be smart from the aggregator to recruit them to the customer pool as very few light industries are required to provide a significant response (Figure 4.3).

4.5 Chapter Summary

In summary, the model in paper A has successfully applied a method to provide a probabilistic output of the price responsiveness of DR across different sectors. The model output is useful for the aggregator to plan its operation strategy and composition of customer pools. Furthermore, the clusters can be constructed to represent customers in various DSO grids. However, for aggregator online real-time operation, a time series modeling approach would be more suitable. The output can also be used to support the communication to the DSOs informing about the probability on response form the DR. Such information can be extracted from the generated cdf's given a customer cluster, at different points in time. The contributions related to

this chapters are the following:

- Applying quantile regression and P-splines to model the price responsiveness of various flexible consumers.
- Providing probabilistic model output to construct cumulative density functions.
- Probabilistic analysis of customer pool DR in relation to required ancillary services by the DSO.

CHAPTER 5
DSOs in SE-OS

This chapter presents methods aiming to enable adaptive DSO smart grid operation in the context of the SE-OS. The methods are designed to increase the observability for DSOs in real-time, and can be applied to monitor and evaluate the state of low voltage radial grids. The developed methods have the capability to provide real-time estimations and probabilities forecasts. Through this knowledge the DSO will have tools to make informed decisions in the context of SE-OS and request ancillary services. The chapter further discusses the results and insights of paper B, C and D in the context of adaptive operation in DSO smart grids in the SE-OS framework.

5.1 Concept of adaptive operation in DSO smart grids

For the DSOs to participate in the SE-OS framework and gain the benefits from ancillary services provided by aggregators, there is a need to understand the SE-OS framework from the DSO's perspective. For this purpose the concept of adaptive operation in DSO smart grids is developed. It involves digitization of low voltage (LV) distribution grids to enable online real-time monitoring as a necessary means to adapt to a new situation with increasing levels of DGs and changing behaviour in demand, as described in Section 2.2.

Figure 5.1 presents the conceptual framework for adaptive DSO operation of digitized LV grids. In this framework data is collected from installed measuring devices at the transformers and selected nodes in LV distribution grids. Note that the final installation of devices will differ from the experimental setup (Figure 3.3 and Figure 3.4), as the latter might involve a redundant number of devices for online monitoring. At the transformers the devices collect temperature data as well as electrical measurements from the transformer and all outgoing feeders. Data is also collected at selected nodes in the LV grid, to measure voltage or alternatively also current, power factor and harmonic distortion (see Chapter 3 for a more detailed description). Similarly, to the experimental setup the data may be collected by a third party, and thus it is fetched from third party data bases. In the operational framework, the DSO online forecasting and monitoring tools pull the data from the third party data bases along with

other required data, such as weather data and electric load forecasts, using REST API's. If not cleaned by the third party, the data will be cleaned through pipelines in the online tools (for instance as in Section 3.3).

Through the data-driven algorithms in the online tool, real-time grid and equipment state estimations and forecasts as well as associated uncertainties are provided. Through the output from the online tools, the state of the grid can be evaluated in the DSO control room, as the DSO gains information on if grid states are, or are predicted to be, violated. For instance, the congestion of the transformer is monitored using a thermal dynamic rating model. If local grid issues are predicted, the DSO can then take action by requesting flexibility ancillary services from an appropriate aggregator in the SE-OS framework. The aggregator then broadcasts signals to flexible resources to change their load. Here, voltage adjusting flexibility should be locally prioritized above frequency regulating flexibility at higher layers in the SE-OS, since active power regulation for voltage control might be more effective than reactive power regulation due the low X/R ratio in the LV grids. Furthermore, the aggregator could use the method described in Chapter 4 to calculate how many customers to activate and which price to set through the probabilistic model, and could further inform the DSO about the uncertainty of the DR. An alternative approach, if the DR is not predicted to solve the local grid issues, is to to use a DSO operated Battery Energy Storage System (BESS). The DR will be noticeable in the measuring devices placed in the grid, and load forecasts will be updated. This procedure is iterated for each time step and the DSO will receive updated forecasts and estimations for every time step. Thus, the DSO can monitor if the flexibility is predicted to solve the local grid issues, or if further changes in DR is required. Here, probabilistic outputs from the tools are essential in the DSO's capability of taking informed decisions and accounting for uncertainty/risk when taking action in the SE-OS framework.

To develop useful methods applicable in the described framework, the approach has been to use available data to achieve practically possible methods. In other words: given the situation and setup in the LV grids studied, how well can we predict and estimate critical states such as end-node voltages and transformer temperature? Focus has also been given to not develop isolated methods, but rather methods that can be combined and have a common installation that is scalable. Furthermore, the aim is to develop data-driven models such that the operation can rely on real-time information. In summary, the research objectives to develop methods for adaptive DSO smart grid operation is:

- To develop models that can estimate and predict real-time states of the grid, while providing uncertainty information.
- Using available data, that can be practically measured or fetched from existing data bases.
- Achieving a practical and scalable installation.

- Developing methods/tools that gives information for managing the uncertainty of the grid state and DR in the context of the SE-OS.

Moreover, the aim of paper B and C are:

- To develop a forecasting model for transformer temperature to apply dynamic transformer rating (paper B)
- To develop a voltage observability method for radial low voltage grids for real-time grid state monitoring (paper C).

5.2 Data-driven modeling for DSO observability and dynamic ratings

This section gives an overview of applied methods related to paper B and C. A special focus is given to time series modeling in DSO grids to provide increased observability and dynamic transformer rating (DTR) for the adaptive DSO operation framework described in Figure 5.1. The overview covers both the development of the transformer thermal model for DTR in paper B and the observability method in paper C.

5.2.1 Time series modeling techniques

5.2.1.1 ARMA type models

Auto Regressive - Moving Average (ARMA) type models have been widely used in literature to develop various forecasting models. E.g. in [82] to predict electricity prices or in [83] to predict electricity demand in Turkey. Thus, they have also been explored in the approach towards finding suitable forecasting models for the DSO forecasting tools.

ARMA type models have also been especially adapted in thermal forecasting models. For instance, [84] developed an Integrated ARMA (ARIMA) model to forecast thermal ratings of transmission lines and [85] used an ARIMA approach to forecast temperature in residential premises. For the thermal transformer model, data from weather forecasts, electrical load and other variables were available, which could possible be useful input variables to model the temperature at interest. Hence, ARMAX models (ARMA with exogenous input) were explored as a possible solution to build the thermal model for DTR. ARMAX models have the following structure:

$$\varphi(B)Y_t = \omega(B)X_{t-b} + \theta(B)\epsilon_t \tag{5.1}$$

where Y is the observed variable, X is the exogenous input variables, b is a time delay, ϵ is white noise residuals and $\varphi(B)$, $\omega(B)$ and $\theta(B)$ are polynomials in the back shift

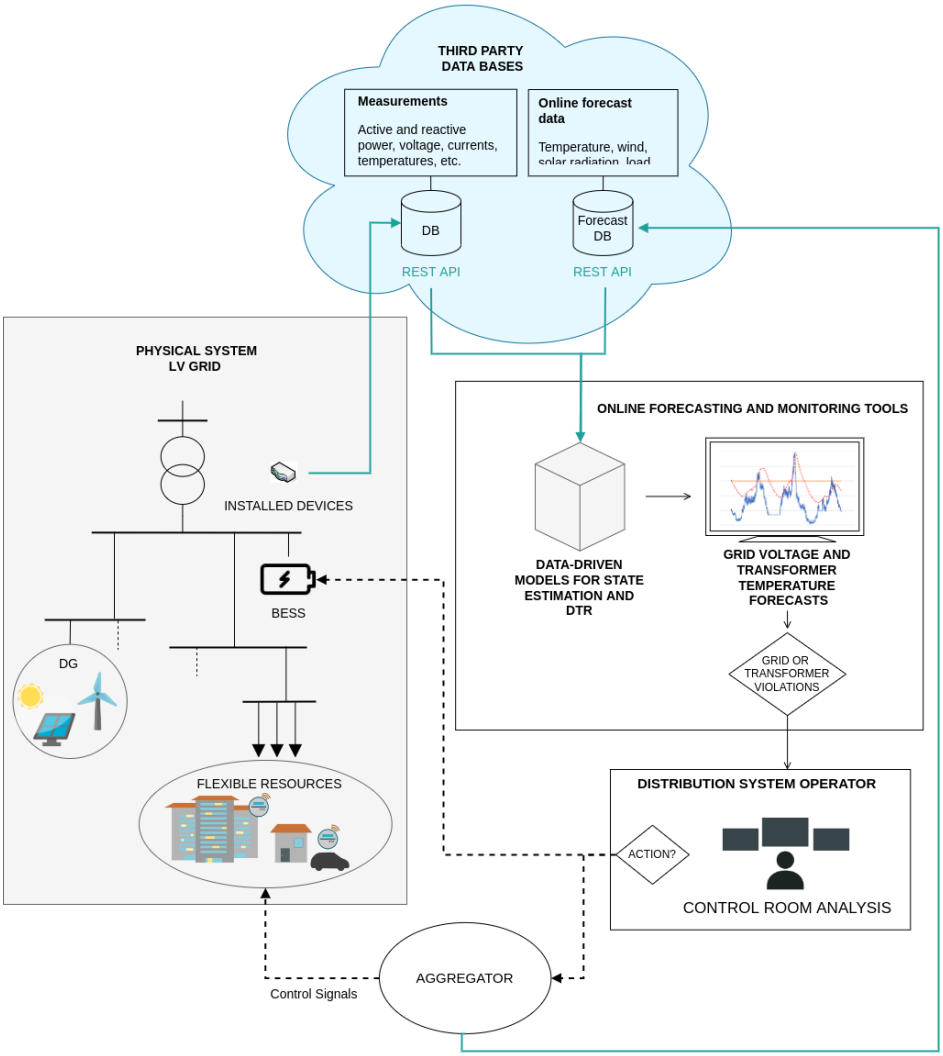


Figure 5.1: Operational framework for adaptive DSO smart grid operation. Turquoise lines indicate data flows and dotted lines indicate communication signals.

operator [86]. This model structure was explored to find a transformer thermal model (aim of paper B). Although models developed using this approach showed reasonable one step ahead predictions, their ability to forecast further ahead was very limited. This could possibly be due to the limited possibility to describe the thermal dynamics in relation to the exogenous input given the model structure. To successfully apply DTR in the DSO operational framework (Figure 5.1), predictions with a horizon of at least a few hours are crucial for the DSO's ability to request flexibility ahead in time before transformer limits are violated. Therefore, other modeling approaches were explored.

To construct the voltage observability model Multivariate ARIMA (MARIMA) models were also explored. MARIMA models have a similar formula as in Eq. (5.1), however, observation variables are also allowed as input to describe other observation variables, thereby being multivariate. [87]. As node voltages affect other node voltages in a power grid, a multivariate model structure could be useful and MARIMA models were investigated using the *marima package* in R [88, 87]. Yet, the model showed poor performance in estimating the voltages when applying it to a data set from fewer electric measurement devices. One of the aims for the observability model was to have a modeling approach with a scalable installation, relying on few measuring devices and thus, MARIMA models were excluded as a solution.

5.2.1.2 Grey-box models

In the SE-OS framework grey-box models have proven to be an effective way to model dynamical systems. For instance, they proved useful in constructing models for wastewater treatment plants in the context of the SE-OS [25] as well as for wind power forecasting [89]. They have also proven to effectively model complex thermal systems (e.g. in [90]), and have thus been explored to build useful models in paper B andC. Grey-box models are physics-informed, in the sense that a first proposal of the model structure is derived from theoretical knowledge. The models are also data-driven, since statistics and information from data is used to identify the relationship between input and output variables. Thereby, grey-box models are at the intersection between white-box approaches (deterministic models), and black-box approaches for which no theoretical knowledge of the process is used [91].

In grey-box modeling, first order stochastic differential equations (SDE) are often used to describe the dynamics of the system, which can be expressed as follows:

$$dX_t = f(X_t, U_t, t, \theta)dt + \sigma(X_t, U_t, t, \theta)dw_t \tag{5.2}$$

where X_t is a state vector, U_t is an input vector, θ are the model parameters and dw_t is a standard Wiener processes (i.e. a Gaussian stochastic process with mutually independent increments) [92]. Here, the first term, the drift term ($= f(X_t, U_t, t, \theta)dt$), represent the deterministic part of the model, while the diffusion term ($\sigma(X_t, U_t, t, \theta)dw_t$)

allows to model noise and disturbances in the system that are not captured by the drift term. This description of the system is linked to the observed states through:

$$Y_k = g(X_{t_k}, U_{t_k}, t_k, \theta) + e_{t_k} \tag{5.3}$$

where Y_k is a vector of observed states and e_{t_k} are the measurement errors, i.e. Gaussian white noise, $\mathcal{N}(0, \sigma^2)$. Note that as the system equation is defined in continues time the observation equation is formulate in discrete time, with time steps k, allowing for such observations. The parameters to the model, θ, are optimized to the data by, for instance, using maximum likelihood. In paper B and C model parameters are optimized using CTSM-R [91]. For a more in depth description of grey-box models the reader is referred to [92, 55, 56]. The presented model formulation here is equivalent to the continuous-discrete time state-space representation in paper B and C.

5.2.1.3 Generalized Additive Models

In the power system, seasonal behaviour is commonly seen in the demand load behaviour [93, 94]. One way to model seasonal behaviour is through using cubic seasonal splines and has for instance been implemented in [95] to model daily electricity load profiles. However, not all electrical load behaviour can be represented using cubic splines. Generalized Additive Models (GAMs), derived from generalized linear models, offer a flexible modeling approach in the possibility to combine various smooth functions. The have for instance been applied to model fault In [96], GAMs are highlighted as a promising candidate for online fault assessment of electrical machines with a low computational burden and in [97] GAMs proved computationally feasible to a big data set to model U.K black smoke network data. Thus, GAMs was explored as a modeling approach for the voltage observability model in paper C, that is required to be scalable in data throughput.

The general formula for GAMs is:

$$g(\mu_i) = \boldsymbol{A_i\theta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + ... \tag{5.4}$$

where μ_i is the expected value of a response variable Y_i. The parametric part of the model is represented in $\boldsymbol{A_i\theta}$, with parameters, θ, and model smooth functions of variables x_j are represented by f_k [98]. In paper C, GAMs for voltage observability were developed using R package *mgcv version 1.8-40* [98, 99, 100, 101, 102]. For further details on GAMs the reader is referred to [98].

5.2.2 Application in forecasting and monitoring tools for DSOs

5.2.2.1 Thermal model for dynamic transformer rating

To develop useful thermal transformer models for DTR in the DSO adaptive operational framework (Figure 5.1) there are e few important considerations in addition

to the background information presented in Section 2.2.2. Firstly, the models need to be able to provide reasonable forecasts, as outlined in Section 5.1. The forecasts further, need to have a time horizon that allows the DSO to request flexibility with a time time margin such that the transformer limits are not violated. To ensure that enough time margin is provided in the predictions, a cross-correlation analysis is implemented in paper B.

It is further becoming increasingly relevant to account for harmonic distortion in distribution grids, due to non-linear loads [36, 103]. Some works can be found on thermal transformer models in literature, investigating the impact of harmonics on the loading capacity of transformers, e.g. [104] introducing an algorithm to estimate loading capability under harmonic conditions, [105] suggesting a 3-D finite element approach, and [106] proposing an extension to the IEC Std. 60076-7 thermal models to account for harmonic effects. The presented works indicate a considerable impact from harmonic distortion in the transformer temperature, and it is thus investigated in paper B. It should be noted that the mentioned studies are focused on monitoring and do not have forecasting capabilities. Thus, they cannot be applied in the DSO operational framework, for which other methods are required.

The models within the presented framework are further focused on secondary distribution grids and thus the transformer thermal models primarily concerns secondary transformers. Considering the size of DSO areas and the supposedly large number of secondary transformers, the applied methods need to be scalable. In thermal modeling a commonly applied method is 3-D finite element method also used for transformer modeling, for instance applied in [105] or in [107] using the approach for model validation. However, these type of methods were instantly excluded due to the computational burden and therefore low scalability.

Another aspect is that the secondary transformers are in general small transformers. As stated in Section 2.2.2, temperatures of such small transformers might be affected by environmental factors, for example wind and solar radiation [35]. In paper B, environmental factors are thus included in the model.

5.2.2.2 Voltage observability method in low voltage radial grids

A few remarks can also be made on important factors to have in mind, in the approach to develop the voltage observability model in paper C.

The developed operational concept concerns unbalanced radial LV distribution grids. Therefore, it is interesting to develop a model for estimating phase voltages. Furthermore, the unbalance might lead to the neutral carrying current and thus a voltage drop or rise along the neutral wires [108, 109] (see also Section 2.2.3). Significant neutral currents were observed in both studied grids of the experimental setup and therefore their impact was investigated in paper C.

Although, several studies can be found on distribution system state estimation (DSSE) for medium voltage grids, there are less studies to be found on radial LV grids at 0.4 kV, which is the primary focus in this thesis. At this voltage level real-time observability is typically nonexistent or very low. In DSSE for higher voltage levels neural network (NN) have been extensively applied, for instance in [110] applying a NN approach to an IEEE 34-bus system at 24.9/4.16 kV or in [111], where the authors apply a NN approach for system monitoring in a 20 kV grid. Since the aim for the observability model is to develop a method with some degree of explainability based on inputs from few devices, it is questionable if an NN approach would provide enough explainability and if the measuring points would provide enough information to optimize NN models.

Many DSSE methods are also based on pseudo-measurements. An approach for a radial LV system state estimation can be found in [112] and used pseudo-measurements to fit the state estimation model through power flow calculations and polynomial regression. Relying on synthesized data requires relevant and updated pseudo-measurements and could potentially result in maintenance work for the DSO to find and maintain such data sets. In paper C another approach was therefore taken to avoid such pseudo-measurement data, to rather rely on the measured data instead.

In LV grids, smart meters broadly being installed (in Europe for instance through the EU directive 2009/72/EC [113]). It could be tempting to use the large amount of smart meter data for voltage estimation. However, the authors in [42] highlight that the required communication infrastructure is expensive and could be subject to cyber attacks. Therefore, phase management units (PMUs) might be a better choice for data input to observability models, as they generally also have a higher sampling frequency [14]. The choice of data input is also important in terms of scalability of the model.

Finally, in radial LV grids the end-nodes are of more critical concern for the DSO. As they are the furthest away from the transformer this is commonly where the largest voltage drops and rises occurs.

5.3 Discussion of results - insights on online tools and DSO smart grid operation

This section firstly discusses the separate results and contributions the models presented in paper B and C, followed by the results and contribution in the context of the DSO smart grid operation framework described in Section 5.1.

5.3.1 Transformer thermal model

In paper B, a thermal model was developed for transformer temperature forecasting using a grey-box modeling approach. Through statistical analysis of the model parameters it was found that relevant inputs were the 3-phase current squared, neutral current squared (for transformer 2 only), ambient temperature, solar radiation and wind speed. The final model had the following formulation:

$$dT_i = \frac{1}{C_i}\left(\frac{1}{R_{ti}}(T_t - T_i) + \frac{1}{R_{ib}}(T_b - T_i)\right)dt + \sigma_1 dw_1 \tag{5.5}$$

$$dT_t = \frac{1}{C_t}\left(\Phi_h + \frac{1}{R_{ti}}(T_i - T_t)\right)dt + \sigma_2 dw_2 \tag{5.6}$$

$$dT_b = \frac{1}{C_b}\left(\frac{1}{R_{ib}}(T_i - T_b) + \frac{1}{R_{ba}}(T_a - T_b) + \omega\Phi_{wind} + \Phi_{sol}\right)dt + \sigma_3 dw_3 \tag{5.7}$$

$$T_{lid} = T_i + e \tag{5.8}$$

where R and C represent thermal resistance and capacitance, respectively. Three states are represented in this model; the state at the lid of the transformer, T_i, the state in the housing, T_b, and the thermal state within the transformer, T_t. The latter two states are hidden states. Environmental impact is model represented in Φ_{wind}, which is the wind speed and Φ_{sol} representing the impact from solar radiation, such that $\Phi_{sol}(t) = \sum_{j=1}^{n} sc_j B_j(t) Gh(t)$, where B_j are B-splines (see Section 4.3.1), $\hat{sc}_j$ is the corresponding estimated parameter and Gh is the solar radiation. Φ_h represent the internal heating in the transformer and is described by the three-phase current squared ($\Phi_h = aI_{3ph}^2 + b$), or alternatively also neutral current for transformer 2 ($\Phi_h = aI_{3ph}^2 + b + cI_N^2$). For further details see Section 5.2.1.2 and paper B.

Differences between the final models of the two transformers were observed for solar radiation B-splines and the neutral current proved to only be significant for transformer 2. As mentioned in paper B, this is probably due to the smaller size of transformer 2 (200 kVA compared to 400 kVA for transformer 1), rather than a higher harmonic content since the neutral current to total phase current ratio is quite similar for the two transformers. Since the loading of transformer 2 is higher, it could mean that the electrical input variables have a higher impact in relation to the environmental inputs compared to transformer 1, which is only loaded at approximately 50 %.

For scalability of the thermal model in the DSO network, ideally, the final formulation of the thermal model should be the same for all transformers. Thus, when applying the model to the operational framework, it is suggested that neutral current is included as an input variable and only removed if proven to be insignificant (for a better model calibration). Furthermore, automizing the optimizations of the B-splines could improve scalability. The final model further suggests that not all

measurements are required and the final setup is suggested in Figure 5.2. Note that only the measurement at the lid, T_{lid} is required to run the model, and the measurement at the radiator T_{rad} is only kept in the installation, because it is suggested by the manufacturer of the transformer. In paper B, it is suggested that heat-run test from the manufacturer could further improve the accuracy of the model, and thus this measurement is kept in the installation to ensure future model alignment with the manufacturer.

The forecasts are provided with a 6 hour horizon. A cross-correlation analysis

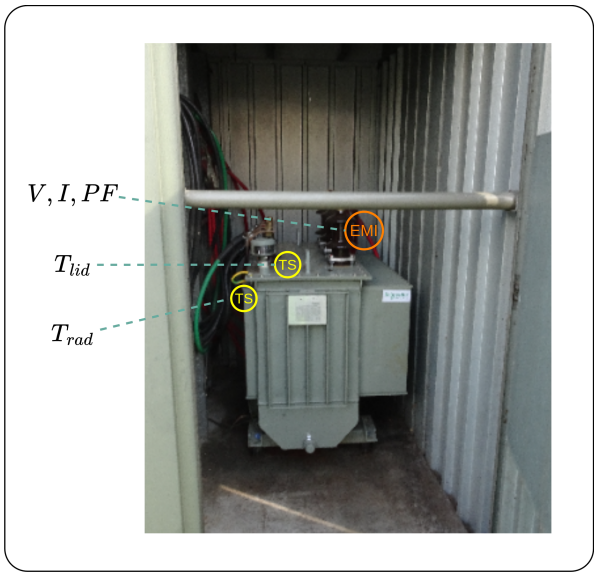
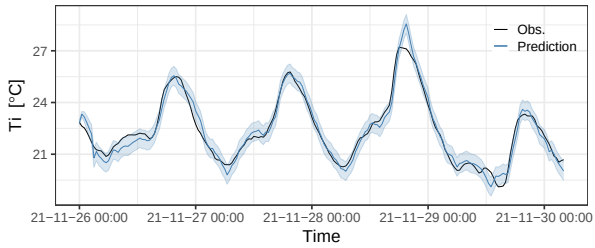


Figure 5.2: Suggested final setup for the transformers, with temperature sensor (TS) and electronic measurement instruments (EMI).

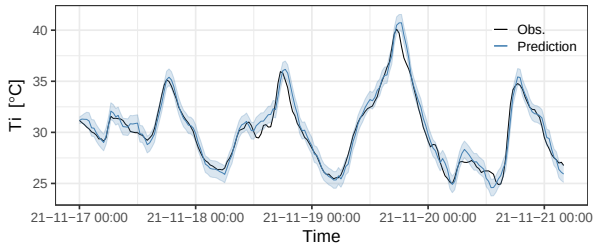
in paper B, indicates that the lid temperature correlates to the inner hot-spot temperature with an approximately 2.5 and 2 hours time delay for transformer 1 and 2, respectively. The 6 hour time horizon, thus gives 3.5 and 4 hours, respectively, for the DSO to take action in the SE-OS framework and request flexibility. The forecasts related to temperature monitoring, i.e.e with time horizons of 2.5 and 2 hours are seen in Figure 5.3, whereas the forecasts to be utilized for flexibility requests in the operational framework, i.e. with 6-hour horizons are seen in Figure 5.4.

5.3.2 Voltage Observability Model

In paper C, a model for voltage observability is developed. The final model for observability along a radial includes a grey-box model using input parameters from one



(a) 5 step ahead predictions for the test data set, using the one state model for transformer 1.



(b) 4 step ahead predictions for the test data set, using the extended two state model for transformer 2.

Figure 5.3: Predictions of the transformer lid temperature for 5 and 4 time steps ahead, respectively, corresponding to the current state inside the transformer. Black line – observations, Blue line – predictions, Light blue area – 95% prediction interval. Figures from paper B..

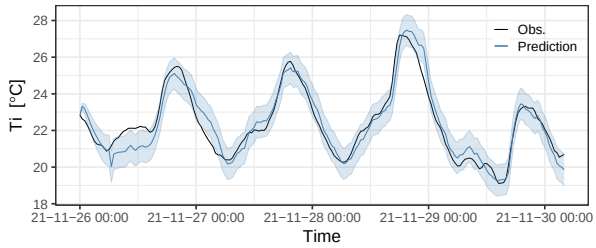
end-node device to estimate a node voltage along the feeder, V_{GM} and the voltage drop between the nodes. The estimated voltage drop is then used as an input variable to estimate other end-node voltages. The final grey-box model is formulated as follows:

$$d\Delta V_{R,GM-GE1} = a \cdot R_{GM-GE1} \cdot \frac{dI_{GE1}}{dt} \cdot PF_{GE1} + \sigma_1 dw_1 \tag{5.9}$$

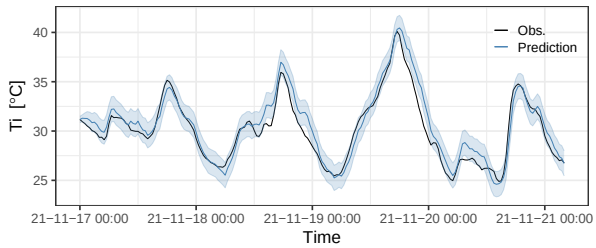
$$d\Delta V_{X,GM-GE1} = b \cdot X_{GM-GE1} \cdot \frac{dI_{GE1}}{dt} \cdot \sin\left(\arccos\left(PF_{GE1}\right)\right)) + \sigma_2 dw_2 \tag{5.10}$$

$$V_{GM} = c \cdot V_{GE1} + d \cdot (V_{R,GM-GE1} + V_{GE1}) \tag{5.11}$$

where $\Delta V_{R,GM-GE1}$ is the state for the voltage drop related to the wire resistance, R_{GM-GE1}, between the nodes with measuring devices, G_M and G_{E1}, and $\Delta V_{X,GM-GE1}$ is the state for the voltage related to the reactance, X_{GM-GE1}, along the same line. See Section 2.2.3 for further details on voltage drop calculations used to derive the model structure and paper C for further details on the final model. Further, the final



(a) Thermal forecast for transformer 1



(b) Thermal forecast for transformer 2.

Figure 5.4: Forecasts with a horizon of 12 step ahead (6 hours), using the final model in paper B. Black line – observations, Blue line – predictions, Light blue area – 95% prediction interval. Figures from paper B..

GAMs to estimate end-node voltage is formulated as follows:

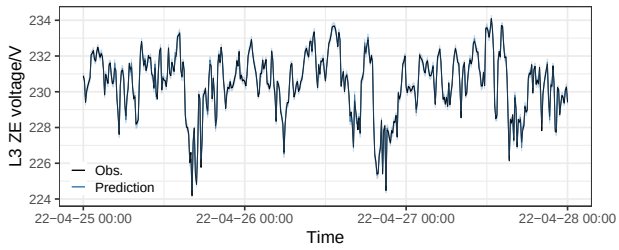
$$V_{end} = s(V_{est,GM}) + s(\Delta V_{est,R,GM-GE1}, I_{ZT}) + s(t_{day}) \tag{5.12}$$

where $\Delta V_{est,R,GM-GE1}$ and $V_{est,GM}$ are the estimated states from the grey-box model, and $s(t_{day})$ are seasonal splines for the daily variation using B-splines of degree 3 with 144 knots. It was investigated to incorporate the other state from teh grey-box model,$\Delta V_{X,GM-GE1}$, but it was insignificant to the model structure. All voltage estimations are in a 10 minute time resolution and per phase.

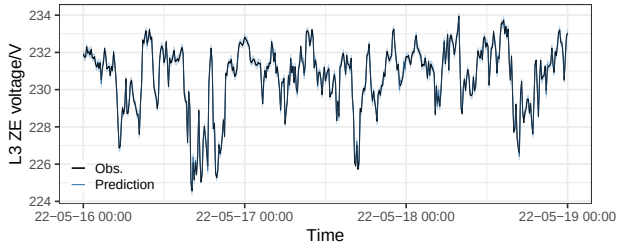
Voltage estimations for the middle node using the grey-box model are seen in Figure 5.5. The predictions have a low confidence interval due the low standard deviation in the model output. The estimations further have a low root mean squared error (rmse) of approximately 0.1 V for both the training data set (18th to 30th April 20022) and the test data set (1st to 31st May 2022). An example of end-nod estimations are seen in Figure 5.6. The end-node estimations of V_{10H} had a rmse of 0.22 V and 0.24 V for the training and test data sets, respectively while for V_{10H} the found rmse were 0.39 V and 0.49 V. For comparison with other works the 0.1 V error is

0.0004 in p.u. and the 0.49 V error is corresponds to 0.002 p.u. Given this, the model provides output with reasonable accuracy for online monitoring and can be utilized in the operational framework.

Nonetheless, the model selection process and data analysis in paper C suggested



(a) Estimations on the training data set

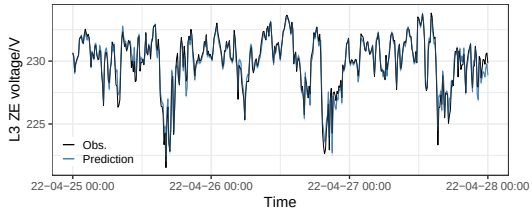


(b) Estimations on the test data set

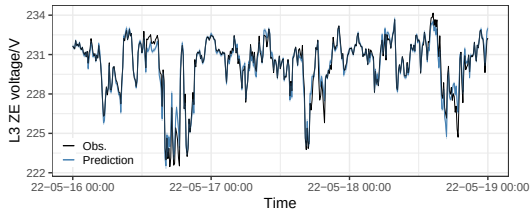
Figure 5.5: Grey-box model estimations on the training and test data sets, zoomed in on three days. The black line represents the observations and the blue line the model predictions. There is also a 95 % confidence interval indicated by a blue area, but is visually difficult to see in the graph due to a low standard deviation in the model. Figures are from paper C.

some improvements for future installations to further. For instance, the end-node voltage measurements should be improved by also measuring the neutral wire as it might have an impact on the phase voltage in an unbalanced network. In the model selection process, focus was given to selecting few input devices to achieve a scalable and practical installation. A reduced installation for the observability online monitoring tool is proposed in Figure 5.7 and Figure 5.8. These setups should however, be validated for other feeders in the future.

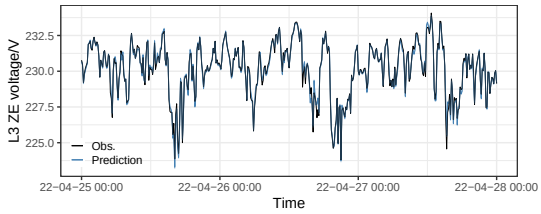
Finally, the developed method does not rely on pseudo-measurements limiting extra work to apply the model in operation. It also provides some explainability in the



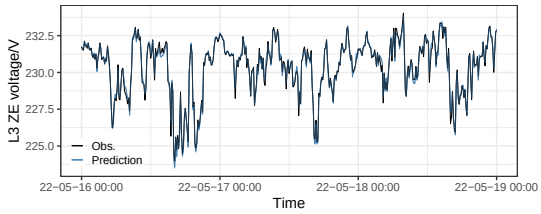
(a) Estimations on the training data set for 10V



(b) Estimations on the test data set for 10V



(c) Estimations on the training data set for 10H



(d) Estimations on the test data set for 10H

Figure 5.6: GAM model end-node estimations on the training and test data sets for V_{10V} and V_{10H}, zoomed in on three days. The black line represents the observations and the blue line the model predictions. There is also a 95 % confidence interval indicated by a blue area. Figures are from paper C.

grey-box model estimating the voltage drop along the radial, which can be used to explain other voltage drops. Furthermore, one end-node device is chosen as input and thus, one end-node voltage will be know to the DSO with high certainty which is beneficial for the grid operation.

5.3.3 DSO Operational Framework

his section discuss the thesis contribution of developing the concept of the adaptive operational framework in Section 5.1. It further discusses the contributions from paper B, C; and D from this perspective.

5.3.3.1 Combined installation

In the DSO operational framework, focus is given to digitizing the LV grids and the setup to run the online tools is thus of interest. The combination of input variables required for the developed methods for DTR and voltage observability in paper B and C, results in a combined installation for the two studied grids respectively. The final setup to run the monitoring and forecasting models can be seen in Figure 5.7 for LL1 and in Figure 5.8 for LL2. Here, the final placement of the devices for the voltage observability model are a suggestion, to be validated in future studies. Nevertheless, the number of devices should be the same. The final setup includes one device at the end of each feeder which is required to run the voltage observability model. It also includes a device at middle nodes for branched feeders, which is needed for model validation and could possibly be removed at a latter stage if the model calibration is successful. For feeders that are not branched, for example the brown, yellow and green feeder in grid 2 (Figure 5.8), it is suggested to rather measure the end-node voltage directly. This is supposedly the most crucial node voltage for the respective feeders and the voltage drop along the feeder could then be estimated with the grey-box model in paper C.

Throughout the model selection process for the methods, it has been actively chosen to keep the required number of devices low. Thus, the combination of the voltage monitoring and DTR tools result in a combined installation that is practical and possible to scale in DSO grids, consisting of many small LV grids.

5.3.3.2 Daily operation of observability tools

The developed methods in paper B and C deal with two different physical phenomena, with regard to dynamics. The transformer model in paper B considers a thermal system, typically with slow dynamics, whereas paper C considers nodal voltages, characterized by considerably faster dynamics. This is reflected in the model output as the thermal model requires a longer forecasting horizon, and thus a forecasting horizon of 6 hours is given such that the DSO has time to mitigate any transformer capacity violations. For the observability model, it is significantly harder to provide

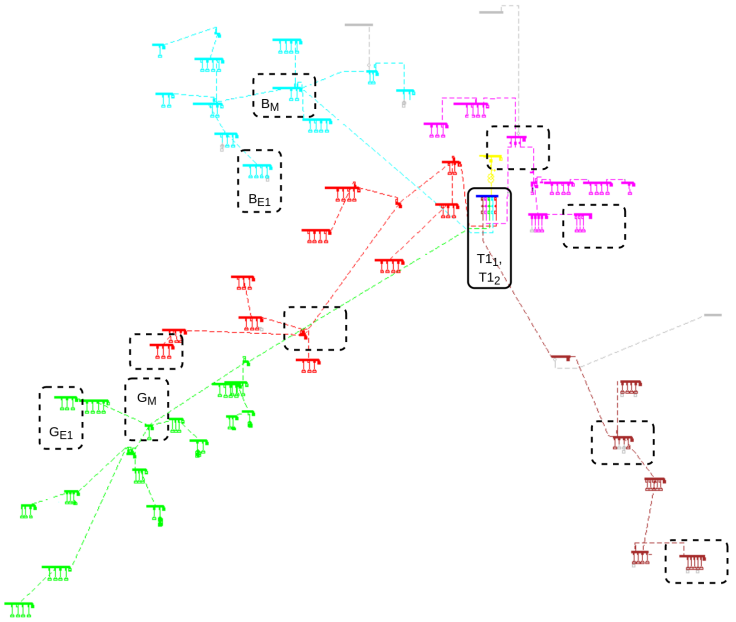


Figure 5.7: Suggested final setup for grid 1/LL1.

forecast with such long horizons and reasonable accuracy. Further, such forecasts has potentially little value as the system dynamics would most likely change within those 6 hours. Instead focus is given to achieve a monitoring tool better suited for voltage observability purposes.

Moreover, the DSO can use both models in paper B and C to investigate future scenarios in the grid as well as analysing historical output from the model to evaluate grid equipment replacement. For instance, if the voltage or transformer temperature has such behaviour that limitations are often ”about to be violated” and flexibility is required for a majority of the time during the daily operation, it might be time for a replacement. Such analysis is possible due to explainability in the models. This extends the dimension for the usage of the DSO tools to - *analysis*, monitoring and forecasting.

The different dynamics also has implications for the required characteristics of flexibility in the operational framework. As the temperature changes more slowly, a more constant response might be required from the flexible devices in Figure 5.1, whereas the voltage is more volatile and thus, the requested flexibility to solve voltage related

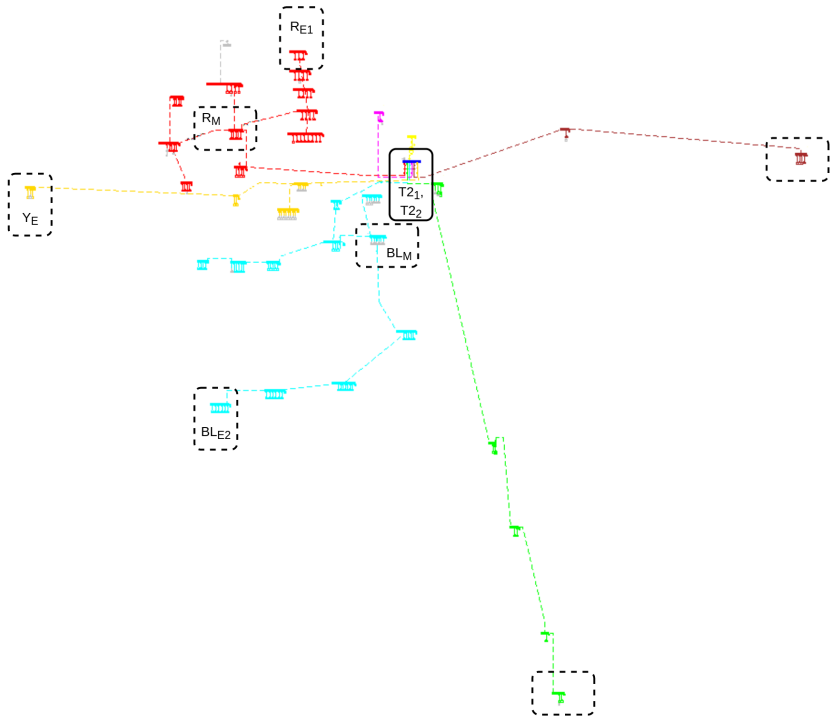


Figure 5.8: Suggested final setup for grid 2/LL2.

issues might be needed with a faster ramping and of a more volatile nature. Therefore, it is very important in the framework that the electricity load forecasts are updated for the input to the online forecasting tools. For instance, the faster changing flexibility for voltage control will also affect the loading of the transformer and thus, the input and output of the DTR tools. This, further highlights the need for combined solutions for DSOs, rather than having separate solutions for different grid issues. As the solutions for each of the grid issues, most likely is a change in power flow, their solutions and control signals will affect each other as they are in the same power system. Hence, the forecasting tools and control systems need to communicate in a combined framework.

Furthermore, the computation time for both methods in paper B and C allows for daily updates of the models. The computation time to optimize the parameters for the thermal model is about 6 minutes and for the observability model the first iteration takes 14 seconds and for each node and additional 1-2 seconds (see paper B and C for further details). Since smart meters have been or are about to be installed in many LV grids the DSO can utilize this in the daily operation. Meter measurements

are often collected once a day [14], and simultaneously the model parameters can be updated to maintain the accuracy of the predictions, as the short computation time would allow this.

Finally, real-time observability has through this project increased from practically zero to achieving models that can be used for online grid monitoring and forecasting. In summary, both models provide the DSO with tools to take action in the SE-OS framework and request flexibility ancillary services.

5.3.3.3 Uncertainty mitigation

As previously stated, the models in paper B and C are capable of providing prediction or confidence intervals (PIs or CIs). Since the associated uncertainty for the predictions are reported in the model output, the DSO can apply risk averse operation strategies in the operational framework. For instance, the DSO might choose to base the DTR on the upper interval for the temperature predictions, as too high temperature are preferably avoided. A similar approach can also be adapted for the voltage observability, but here the DSO need to change between upper and lower CIs as both drops and rises are unfavourable.

BESS
While the above considers the uncertainty in the grid states there are another dimension of uncertainty through the unidirectional communication flow, namely in the DR. As seen in paper A, the certainty of the DR will vary for customer pool composition and time of the day. That is, there will be periods with higher and lower certainty in the DR. For the periods with low certainty the DSO might need a complementary strategy to the unidirectional communication flexibility to safely mitigate local grid issues. Paper D, Battery-Energy Storage Systems (BESS) in LV networks are highlighted as a solution to mitigate local grid issues. Examples of grid relief and ancillary services through utilization of BESS can also be found in other studies. For instance, in [114] the authors investigated a system operator owned BESS to provide ancillary services, finding that it could improve grid management and delay grid investments, while the authors in [115] found that a BESS for grid relief reduced local demand and power peak stress. Furthermore, potential benefits of communal BESS rather than individual are highlighted in [116] and [117]. It might thus be of interest for the DSO to incorporate a BESS to mitigate local grid issues in the operational framework.

In [118] (also attach as paper D in this thesis), we studied the placement of a BESS in three different LV grids (city, suburban and village) in the context of an energy community with different operation strategies. Energy communities compromise cooperatives sharing renewable DERs, while minimizing their energy consumption through flexibility of active consumers. The BESS placement was studied through a series of power flow calculations of various scenarios, including different operating strategies of the energy community as well as different energy community configurations. In

total 180 scenarios were investigated in the paper. In paper D, the first configuration (EC1 in the paper) only includes households and is comparable to grid 1, while the other configurations (EC2 and EC3 in the paper) are comparable to grid 2 as they also include a commercial customer (i.e. with larger load). The grids are comparable to grid 1 and 2 described in Chapter 3, although the village and suburban grids are probably the most relevant. We used household load profiles representing detached houses without electric heating, detached houses with heat pumps and apartments without electric heating.

In paper D we found that the energy community and BESS strategy of peak-shaving improved the line loading. As an example the LV line loading was up to 58 % lower for the city grid (with configurations EC2 and EC3) compared to the strategy to maximize economic benefits. This indicates promising results to reduce the loading of the transformer in LV grids, with potential benefits for the application of DTR (relevant for paper B). Whereas, paper D was focused on energy community owned BESS, El-Batawy [119] studied optimal parameters for a BESS from a DSO perspective, finding voltage improvements for two of their four investigated scenarios. This suggest that a BESS could also be utilized to mitigate voltage violations, during times with low certainty in the DR and is therefore incorporated in the operational framework (see Figure 5.1). However, in paper D, where the study is conducted from energy communities perspective, no particular voltage improvements were found. Due to this, it is suggested that the BESS is owned and operated by the DSO. Nevertheless, paper D showed a clear advantage for the grid voltage impact, in having the BESS at the beginning of the feeder in the LV networks.

5.3.3.4 Flexibility coordination

As seen in Chapter 4, the aggregator have probabilistic price elasticity information that might be useful for the DSO to design their grid issues mitigation strategies. For instance, if the probability of overloading the transformer is high, while the certainty of DR is low it is useful for the DSO to have this information such that the risk of system failure can be avoided. For instance, the situation can be avoided ahead in time by planning to activate a BESS instead. Vice versa, the DSO has useful information for the aggregator to build their customer pools. For instance, the thermal transformer models shows a daily peak in temperature in the afternoon/evening (Figure 5.4). Such information can be combined with the DR price elasticity box plots, in Figure 4.2, Figure 4.3 and Figure 4.4, to evaluate which customer clusters are likely to respond to DR signals at that time with high certainty. Therefore, exchange of information between the aggregator and the DSO is suggested to effectively allocate flexible resources for ancillary services in the SE-OS framework. It should be noted that, although a soft version of two-way communication is introduced here, the communication to the customer flexible resources is kept unidirectional.

Energy communities

As previously stated, energy communities with communal BESS offers an alternative approach for the DSO can gain support to solve local grid issues. However, communication between the energy community and DSO is encouraged, to avoid unsuccessful configurations which might undermine the energy community goals of easing their loading and impact on the power grid. This is highlighted in paper D, where it was seen that the energy community operation strategy and placement of the communal BESS had different impacts on line loading and voltage violations in the studied LV grids. Coordination is thus suggested to avoid grid violations by adjusting operational strategies and placement of communal BESS. The results further suggest that energy communities can provide ancillary services, for instance through peak-shaving (i.e. related to congestion problems). To incorporate this in the DSO operational framework it is, however, suggested that the flexibility coordination is operated by an aggregator to simplify communication flows. Thus, energy communities would also be included as "flexible resources" in the framework described in Figure 5.1.

5.4 Chapter Summary

In this chapter model development, results and contributions from paper B, C and D have been discussed. The chapter presents an adaptive DSO smart grid operation framework. A thermal transformer model for DTR have been presented in paper B and a voltage observability method for radial unbalanced LV grids have been presented in paper C. Both models contributes to the online monitoring and forecasting tools in the described operational framework. Such methods are crucial for the DSOs to adapt to a new distribution power system situation with increasing levels of DGs and changing demand behaviour. The operational framework further suggests adaptive measures to arising grid issues during operation by requesting flexibility form an aggregator in the SE-OS. Furthermore, energy communities are highlighted as flexible resources to provide DR in paper D and BESS is suggested as an alternative addition to times with high uncertainty in the DR ancillary services. To avoid intensified grid impact, communication between energy communities and DSOs is also encouraged.

The contributions form this chapter can be outlined as follows:

- A grey-box transformer thermal model, capable of providing 6-hour temperature forecasts for the application of DTR.
- A method combining grey-box and GAM models for node voltage estimations in unbalanced LV grids, with high accuracy and low computational burden.
- Data-driven methods relying on a combined setup that is practical and scalable.
- An adaptive DSO smart grid operation framework, including online forecasting and monitoring tools as well as solving local grid issues through DR ancillary services, energy communities and BESS.

CHAPTER 6
Conclusions and Perspectives

The focus of this thesis has been to develop an operational framework for adaptive DSO smart grid operation in the context of SE-OS including data-driven modeling approaches supporting the framework.

For this purpose, an experimental setup compromising two living labs has been designed. The living labs consist of two real-world low voltage grids operated by a Danish DSO, allowing for the developed methods to be anchored in a real-world installation and the compromises and challenges that inevitably come with it. Thereby, the developed methods (in paper B and C) have an advantage towards simulation studies, since they are based on data describing a real system.

Furthermore, a framework explaining how a DSO can operate with adaptive measures to solve local grid issues in the context of the SE-OS has been developed and described. Through the framework the DSO gain information about the LV grid and transformer through online monitoring and forecasting tools, which is becoming increasingly important due to increasing levels of DG and changes in demand behaviour.

Within the framework models for two grid evaluating tools are developed through paper B and C. A thermal model for dynamic transformer rating is proposed in paper B, by defining a grey-box model, meaning it is both physics-informed and data-driven. The model further provides forecasts accompanied by prediction intervals (PIs) with reasonable accuracy. The model is evaluated with a 6-hour prediction horizon giving the DSO a time margin to take action within the SE-OS framework if the transformer is predicted to be congested. Furthermore, paper B incorporates both electrical and environmental input variables in a data-driven approach, which is scarcely seen in existing literature.

A model for nodal voltage observability is defined in this thesis through paper C. The model aims to serve the DSO in real-time monitoring of node voltages in unbalanced radial LV grids within the framework. The method relies on a combination of

grey-box modeling and GAMs, providing node voltage estimations per phase and 10 minutes, while also providing confidence intervals. The model achieved high accuracy when being evaluated on a data set twice as long as the training data set, with root mean squared errors of 0.002 – 0.0004 p.u. depending on the node. The model only uses input data from one device per feeder and has a low computational burden.

Both models developed for online monitoring and forecasting in paper B and C, rely on measurements that are practical and possible to measure by the DSO. The combined installation setup that is required is thus supposedly practical and scalable in DSO grids consisting of many radial LV grids. By providing probabilistic output in both models, it further allows he DSO to apply risk averse strategies within the SE-OS framework. Moreover, the short computation times allow the DSO to calibrate the models through optimizing the parameters on a regular basis to maintain the accuracy of the models (for instance daily).

The DSO operational framework, further describes how the DSO can request flexibility from an aggregator in the SE-OS to solve local grid issues. To appropriately support DSOs with relevant ancillary services through activating flexible resources the methods in chapter 4 were developed. In paper A, a flexible and probabilistic price elasticity model for various types of DR was proposed. The model proved capable of extracting cumulative density functions for the DR given hourly electricity prices to analyze how many customers are required to meet a certain request of flexibility services. Through the model the aggregator can both evaluate the DR potential in an LV grid given the customer composition and compose flexibility customer pools that are predicted to be capable of providing required flexibility in the ancillary services. Here, it is also recommended that voltage adjusting flexibility is locally prioritized above frequency regulating flexibility, as active power regulation for voltage control is more effective than reactive power regulation due the low X/R ratio in the LV grids.

As an alternative, energy communities could also provide flexibility to the DSO. An advantage of the energy communities are that they can design the operation strategy to support the DSO if it aligns with their own goals. For instance, in paper D, it was seen that the energy communities could significantly reduce the maximum line loading through a peak-shaving strategy. In the operational framework suggested in the thesis it is, however, suggested that control signals and communication is handled by the aggregator to simplify communication flows and account for other flexible resources in the grid.

An exchange of information of daily patterns in both DR and needed ancillary services between the aggregator and the DSO is further suggested. The communication flow could further help the aggregator to construct relevant customer pools and the DSO could gain information on when complementary measures might be needed to solve local grid issues. For instance, the DSO could use a complementary battery energy storage system (BESS) to account for the times with high uncertainty in the

DR. In paper D it is suggested that such a BESS should be placed at the beginning of a feeder in a LV grid. It is further highlighted that a DSO owned and operated BESS could better improve the voltage situation in the LV grid.

As underlined in Chapter 1, there is a need for new methods to increase the observability to support adaptive measures and control for the DSOs. The PhD project involved communication with DSOs and a general concern has been noted, that applying control algorithms and new methods to the grid would impose a risk to the daily operation. However, as highlighted in this thesis, observability and forecasting methods are key for the DSOs to adapt to a new and future grid situation. The consumption and production behaviours are changing through increasing levels of distributed generation and new technologies in the demand such as electric vehicles and energy management systems. The methods developed in this thesis show that it is possible to estimate and predict grid and transformer states, with reasonable accuracy while reporting uncertainty, equipping the DSOs with tools to apply risk averse strategies in their real-time operation.

In conclusion, the models serve the purpose of using data driven methods to enable adaptive operation of DSO smart grids. The developed framework further enables utilization of renewable distributed generation and flexibility - necessary in the transition towards a carbon neutral power system. The developed framework and methods can further support full utilization of grid equipment, avoiding unnecessarily early replacement, which has clear economic and environmental benefits.

6.0.1 Outlook (Future work)

The methods developed in this thesis can be used to apply an adaptive operational framework for DSO smart grid operation. Nevertheless, a few improvements are suggested here to further improve the methods and applicability.

If price elasticity data becomes available for relevant LV grids in the operational framework, it is additionally suggested to apply the price elasticity model to such data sets. Thereby, the flexibility potential of relevant customer pools can be evaluated.

Although parts of the online tool have been developed, they need to be assembled and implemented in order to verify the models during real-time operation.

Further improvements to the installation setup for the voltage observability tool are suggested, consisting of measuring neutral wire currents at all measuring point as well as all outgoing and ingoing cables for middle node devices. It is then recommended to validate the model on more feeders with the new installation setup.

Bibliography

[1] Hans-O. Pörtner et al. "Climate Change 2022 - Impacts, Adaptation and Vulnerability - Summary for Policymakers." In: *Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (2022).

[2] European Commission. *A Clean Planet for all - A European long-term strategic vision for a prosperous, modern, competitive and climate neutral economy.* Brussels, 2018.

[3] European Commission. *Clean energy for all Europeans.* Technical report. 2019. DOI: 10.2833/9937. URL: https://ec.europa.eu/energy/en/topics/energy-strategy-and-energy-union/clean-energy-all-europeans.

[4] William Lilley, Jennifer Hayward, and Luke Reedman. "Realizing the Potential of Renewable and Distributed Generation." In: *Smart Grid*. Elsevier Inc., 2012, pages 161–183. ISBN: 9780123864529. DOI: 10.1016/B978-0-12-386452-9.00007-3. URL: http://dx.doi.org/10.1016/B978-0-12-386452-9.00007-3.

[5] Benjamin Biegel et al. "Value of flexible consumption in the electricity markets." In: *Energy* 66.2014 (2014), pages 354–362. ISSN: 03605442. DOI: 10.1016/j.energy.2013.12.041. URL: http://dx.doi.org/10.1016/j.energy.2013.12.041.

[6] F. D'Ettorre et al. "Exploiting demand-side flexibility: State-of-the-art, open issues and social perspective." In: *Renewable and Sustainable Energy Reviews* 165 (2022), page 112605. ISSN: 13640321. DOI: 10.1016/j.rser.2022.112605. URL: https://linkinghub.elsevier.com/retrieve/pii/S1364032122005007.

[7] Oktoviano Gandhi et al. "Review of power system impacts at high PV penetration Part I: Factors limiting PV penetration." In: *Solar Energy* 210.June (2020), pages 181–201. ISSN: 0038092X. DOI: 10.1016/j.solener.2020.06.097. URL: https://doi.org/10.1016/j.solener.2020.06.097.

[8] Anderson Hoke et al. "Steady-state analysis of maximum photovoltaic penetration levels on typical distribution feeders." In: *IEEE Transactions on Sustainable Energy* 4.2 (2013), pages 350–357. ISSN: 19493029. DOI: 10.1109/TSTE.2012.2225115.

[9] Reigh A. Walling et al. “Summary of distributed resources impact on power delivery systems.” In: *IEEE Transactions on Power Delivery* 23.3 (2008), pages 1636–1644. ISSN: 08858977. DOI: 10.1109/TPWRD.2007.909115.

[10] N. Jenkins, J.B. Ekanayake, and G Strbac. *Distributed Generation*. The Institution of Engineering and Technology (The IET), 2010, pages 1–20.

[11] Qin Wang and Bri Mathias Hodge. “Enhancing power system operational flexibility with flexible ramping products: A review.” In: *IEEE Transactions on Industrial Informatics* 13.4 (2017), pages 1652–1664. ISSN: 15513203. DOI: 10.1109/TII.2016.2637879.

[12] Carlos Madina et al. “Technologies and Protocols: The Experience of the Three SmartNet Pilots.” In: *TSO-DSO Interactions and Ancillary Services in Electricity Transmission and Distribution Networks: Modeling, Analysis and Case-Studies*. 2019, pages 141–183. ISBN: 9783030292034. DOI: 10.1007/978-3-030-29203-4.

[13] Henrik Madsen et al. *Recommendations from the CITIES project - Centre for IT-Intelligent Energy Systems*. 2021.

[14] Kaveh Dehghanpour et al. “A survey on state estimation techniques and challenges in smart distribution systems.” In: *IEEE Transactions on Smart Grid* 10.2 (2019), pages 2312–2322. ISSN: 19493053. DOI: 10.1109/TSG.2018.2870600.

[15] Giulia De Zotti. “Leveraging Consumers’ Flexibility for the Provision of Ancillary Services.” PhD thesis. Technical University of Denmark (DTU), 2019.

[16] Giulia De Zotti et al. “Ancillary Services 4.0: A Top-To-Bottom Control-Based Approach for Solving Ancillary Services Problems in Smart Grids.” In: *IEEE Access* 6 (February 2018), pages 11694–11706. ISSN: 21693536. DOI: 10.1109/ACCESS.2018.2805330.

[17] Henrik Madsen et al. “Control of Electricity Loads in Future Electric Energy Systems.” In: *Handbook of Clean Energy Systems*. 2015.

[18] Peter Nystrup et al. “Temporal hierarchies with autocorrelation for load forecasting.” In: *European Journal of Operational Research* 280.3 (2020), pages 876–888. ISSN: 03772217. DOI: 10.1016/j.ejor.2019.07.061.

[19] Rasmus Halvgaard et al. “Economic Model Predictive Control for building climate control in a Smart Grid.” In: *2012 IEEE PES Innovative Smart Grid Technologies, ISGT 2012* (2012), pages 1–6. DOI: 10.1109/ISGT.2012.6175631.

[20] Nikita Zemtsov et al. “Economic nonlinear MPC for a population of thermostatically controlled loads.” In: *Computer Science - Research and Development* 33.1-2 (2018), pages 157–167. ISSN: 18652042. DOI: 10.1007/s00450-017-0354-7.

[21] Jacopo Parvizi, John B. Jorgensen, and Henrik Madsen. "Robust Model Predictive Control with Scenarios for Aggregators in Grids with High Penetration of Renewable Energy Sources." In: *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, SmartGridComm 2018* (2018), pages 1–7. DOI: 10.1109/SmartGridComm.2018.8587603.

[22] Olivier Corradi et al. "Controlling electricity consumption by forecasting its response to varying prices." In: *IEEE Transactions on Power Systems* 28.1 (2013), pages 421–429. ISSN: 08858950. DOI: 10.1109/TPWRS.2012.2197027.

[23] Rune Grønborg Junker et al. "Stochastic nonlinear modelling and application of price-based energy flexibility." In: *Applied Energy* 275.November 2019 (2020), page 115096. ISSN: 03062619. DOI: 10.1016/j.apenergy.2020.115096. URL: https://doi.org/10.1016/j.apenergy.2020.115096.

[24] *Smart CITIES Centre*. visited on 2022-07-14. URL: https://smart-cities-centre.org/.

[25] Peter Alexander Stentoft et al. "Towards model predictive control: Online predictions of ammonium and nitrate removal by using a stochastic ASM." In: *Water Science and Technology* 79.1 (2019), pages 51–62. ISSN: 02731223. DOI: 10.2166/wst.2018.527.

[26] Amos Schledorn et al. "Optimising block bids of district heating operators to the day-ahead electricity market using stochastic programming." In: *Smart Energy* 1 (2021), page 100004. ISSN: 26669552. DOI: 10.1016/j.segy.2021.100004. URL: https://doi.org/10.1016/j.segy.2021.100004.

[27] Juan M. Morales et al. *Integrating Renewables in Electricity Markets*. 2014. ISBN: 0884-8289. DOI: Doi10.1007/978-1-4419-1640-2{_}1.

[28] William H. Kersting. *Distribution System Modeling and Analysis*. CRC Press, 2001. ISBN: 0849308127.

[29] Ali Asghar Taheri et al. "Investigation of Mineral Oil-Based Nanofluids Effect on Oil Temperature Reduction and Loading Capacity Increment of Distribution Transformers." In: *Energy Reports* 7 (2021), pages 4325–4334. ISSN: 23524847. DOI: 10.1016/j.egyr.2021.07.018.

[30] Barry H Ward. "A Survey of New Techniques in Insulation Monitoring of Power Transformers." In: *IEEE Electrical Insulation Magazine* 17.3 (2001), pages 16–23. DOI: https://doi.org/10.1109/57.925299.

[31] C J Wallnerstrom et al. "Potential of dynamic rating in Sweden." In: *2014 International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2014 - Conference Proceedings* July (2014), pages 7–10. DOI: 10.1109/PMAPS.2014.6960605.

[32] Thahirah Syed Jalal, Numan Rashid, and Ben Van Vliet. "Implementation of Dynamic Transformer Rating in a distribution network." In: *2012 IEEE International Conference on Power System Technology, POWERCON 2012* (2012), pages 2–6. DOI: 10.1109/PowerCon.2012.6401328.

[33] Tahereh Zarei et al. "Reliability considerations and economic benefits of dynamic transformer rating for wind energy integration." In: *International Journal of Electrical Power and Energy Systems* 106.September 2018 (2019), pages 598–606. ISSN: 01420615. DOI: 10.1016/j.ijepes.2018.09.038. URL: https://doi.org/10.1016/j.ijepes.2018.09.038.

[34] Romain Dupin and Andrea Michiorri. "Dynamic line rating forecasting." In: *Renewable Energy Forecasting: From Models to Applications*. Elsevier Ltd, 2017, pages 325–339. ISBN: 9780081005057. DOI: 10.1016/B978-0-08-100504-0.00013-5. URL: http://dx.doi.org/10.1016/B978-0-08-100504-0.00013-5.

[35] IEC. *DS/IEC 60076-7:2018, Power transformers – Part 7 : Loading guide for mineral-oil-immersed power transformers*. Technical report. 2018.

[36] IEEE. *IEEE Std C57.110-1998 Recommended Practice for Establishing Transformer Capability When Supplying Nonsinusoidal Load Currents*. Technical report August. 1998. URL: http://ieeexplore.ieee.org/iel5/59/10801/x0247870.pdf.

[37] IEEE Standards Board. *IEEE Std. 141-1993, Recommended Practice for Electric Power Distribution for Industrial Plants*. Technical report. 1993, pages 61–108. DOI: 10.1049/pe:19880018.

[38] Shanmuga S. Pandian. "Various considerations for estimating steady-state voltage drop in low voltage AC power distribution systems." In: *Conference Record - Industrial and Commercial Power Systems Technical Conference* (2006). DOI: 10.1109/icps.2006.1677307.

[39] Energinet. *Technical Regulation for Electricity- generating Facilities of 11 kW or lower*. Technical report June. 2008.

[40] Emma Blomgren. "Grid impact study of frequency regulation with EVs Grid impact study of frequency regulation with EVs." PhD thesis. Lund University, Faculty of Engineering, 2018.

[41] J. Duncan Glover, Thomas J. Overbye, and Mulukutla S. Sarma. *Power System Analysis and Design*. Sixth Edit. Cengage Learning, 2012, pages 309–390.

[42] István Táczi et al. "The challenges of low voltage distribution system state estimation—an application oriented review." In: *Energies* 14.17 (2021). ISSN: 19961073. DOI: 10.3390/en14175363.

[43] Lennart Söder et al. "A review of demand side flexibility potential in Northern Europe." In: *Renewable and Sustainable Energy Reviews* 91 (August 2018), pages 654–664. ISSN: 18790690. DOI: 10.1016/j.rser.2018.03.104.

[44] James M Gillan. *Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption*. 2017. URL: https://are.berkeley.edu/sites/are.berkeley.edu/files/job-candidates/paper/JMP_Gillan_most_recent_draft.pdf.

[45] Paul Damien et al. "Impacts of day-ahead versus real-time market prices on wholesale electricity demand in Texas." In: *Energy Economics* 81 (2019), pages 259–272. ISSN: 01409883. DOI: 10.1016/j.eneco.2019.04.008.

[46] Soumyadip Ghosh, Xu Andy Sun, and Xiaoxuan Zhang. "Consumer profiling for demand response programs in smart grids." In: *2012 IEEE Innovative Smart Grid Technologies - Asia, ISGT Asia 2012* (2012), pages 1–6. DOI: 10.1109/ISGT-Asia.2012.6303309.

[47] Mercedes Vallés et al. "Probabilistic characterization of electricity consumer responsiveness to economic incentives." In: *Applied Energy* 216 (2018), pages 296–310. ISSN: 03062619. DOI: 10.1016/j.apenergy.2018.02.058.

[48] Bryan K Bollinger and Wesley R Hartmann. "Information vs. Automation and implications for dynamic pricing." In: *Management Science* 66.1 (2020), pages 290–314. ISSN: 15265501. DOI: 10.1287/mnsc.2018.3225.

[49] Emma M. V. Blomgren et al. *Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty: Data Presentation*. Technical report. 2021. DOI: 10.1109/ISGTEurope52324.2021.9640146.

[50] Ruben Sanchez et al. "Observability of low voltage grids: Actual DSOs challenges and research questions." In: *2017 52nd International Universities Power Engineering Conference, UPEC 2017* 2017-Janua (2017), pages 1–6. DOI: 10.1109/UPEC.2017.8232008.

[51] Izudin Kapetanović, Jasna Hivziefendić, and Majda Tešanović. "Different Approaches for Analysis of Harmonics Impact on the Transformer Losses and Life Expectancy." In: *Lecture Notes in Networks and Systems* 28 (2018), pages 392–408. ISSN: 23673389. DOI: 10.1007/978-3-319-71321-2{_}36.

[52] Yuanyuan Sun et al. "Hotspot temperature prediction of dry-type transformers based on particle filter optimization with support vector regression." In: *Symmetry* 13.8 (2021). ISSN: 20738994. DOI: 10.3390/sym13081320.

[53] D P Rommel, D Di Maio, and T Tinga. "Transformer hot spot temperature prediction based on basic operator information." In: *International Journal of Electrical Power & Energy Systems* 124 (2021), page 106340. ISSN: 0142-0615. DOI: https://doi.org/10.1016/j.ijepes.2020.106340. URL: https://www.sciencedirect.com/science/article/pii/S014206152030867X.

[54] Anggoro Primadianto and Chan Nan Lu. "A Review on Distribution System State Estimation." In: *IEEE Transactions on Power Systems* 32.5 (2017), pages 3875–3883. ISSN: 08858950. DOI: 10.1109/TPWRS.2016.2632156.

[55] Niels Rode Kristensen, Henrik Madsen, and Sten Bay Jørgensen. "Parameter estimation in stochastic grey-box models." In: *Automatica* 40.2 (2004), pages 225–237. ISSN: 00051098. DOI: 10.1016/j.automatica.2003.10.001.

[56] Niels Rode Kristensen, Henrik Madsen, and Sten Bay Jørgensen. "A method for systematic improvement of stochastic grey-box models." In: *Computers and Chemical Engineering* 28.8 (2004), pages 1431–1449. ISSN: 00981354. DOI: 10.1016/j.compchemeng.2003.10.003.

[57] *Flexible Energy Denmark*. visited on 2022-07-10. URL: https://www.flexibleenergydenmark.dk/.

[58] *Uni-lab.dk, Living labs*. visited on 2022-07-10. URL: https://www.uni-lab.dk/en/living-labs/.

[59] L. Thurner et al. "pandapower — An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems." In: *IEEE Transactions on Power Systems* 33.6 (November 2018), pages 6510–6521. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2018.2829021.

[60] *DMI, Frie Data*. visited on 2022-07-12. URL: https://www.dmi.dk/frie-data/.

[61] *Danish Meteorological Institute - Open Data*. visited on 2022-07-12. URL: https://confluence.govcloud.dk/display/FDAPI.

[62] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. "Exploring Network Structure, Dynamics, and Function using NetworkX." In: *Proceedings of the 7th Python in Science Conference*. Edited by Gaël Varoquaux, Travis Vaught, and Jarrod Millman. Pasadena, CA USA, 2008, pages 11–15.

[63] Xavier Labandeira, José M Labeaga, and Xiral López-Otero. "A meta-analysis on the price elasticity of energy demand." In: *Energy Policy* 102 (2017), pages 549–568. ISSN: 03014215. DOI: 10.1016/j.enpol.2017.01.002.

[64] Mark Miller and Anna Alberini. "Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data." In: *Energy Policy* 97 (2016), pages 235–249. ISSN: 03014215. DOI: 10.1016/j.enpol.2016.07.031.

[65] Giulia De Zotti et al. "Analysis of rebound effect modelling for flexible electrical consumers." In: *IFAC-PapersOnLine* 52(4) (2019), pages 6–11.

[66] Homa Rashidizadeh-Kermani et al. "A Regret-Based Stochastic Bi-Level Framework for Scheduling of DR Aggregator under Uncertainties." In: *IEEE Transactions on Smart Grid* 11.4 (2020), pages 3171–3184. ISSN: 19493061. DOI: 10.1109/TSG.2020.2968963.

[67] Magnus Brolin and Claes Sandels. "Controlling a retailer's short-term financial risk exposure using demand response." In: *IET Generation, Transmission and Distribution* 13.22 (2019), pages 5160–5170. ISSN: 17518687. DOI: 10.1049/iet-gtd.2018.6708.

[68] Giulia De Zotti et al. "Consumers' Flexibility Estimation at the TSO Level for Balancing Services." In: *IEEE Transactions on Power Systems* 34.3 (May 2019), pages 1918–1930. ISSN: 08858950. DOI: 10.1109/TPWRS.2018.2885933.

[69] Giulia De Zotti et al. "A Control-based Method to Meet TSO and DSO Ancillary Services Needs by Flexible End-Users." In: *IEEE Transactions on Power Sys- tems* 35.3 (2019), pages 1868–1880.

[70] Malcolm Sabin. "B-Splines." In: (2010), pages 11–16. DOI: 10.1007/978-3-642-13648-1{_}3.

[71] Gernot Beer, Benjamin Marussig, and Christian Duenser. *Basis functions, B-splines*. Volume 90. 2020, pages 35–71. ISBN: 9783030233396. DOI: 10.1007/978-3-030-23339-6{_}3.

[72] Roger Koenker and Gilbert Bassett. "Regression Quantiles Author (s): Roger Koenker , Gilbert Bassett and Jr . Published by : The Econometric Society Stable URL : https://www.jstor.org/stable/1913643 The Econometric Society is collaborating with JSTOR to digitize , preserve and extend acce." In: *Econometrica* 46.1 (1978), pages 33–50.

[73] Ludwig Fahrmeir et al. "Quantile Regression." In: *Regression: Models, Methods and Applications*. Springer Berlin Heidelberg, 2013, pages 597–620. ISBN: 9783642343339. DOI: 10.1007/978-3-642-34333-9.

[74] Kaatje Bollaerts, Paul H.C. Eilers, and Marc Aerts. "Quantile regression with monotonicity restrictions using P-splines and the L1-norm." In: *Statistical Modelling* 6.3 (2006), pages 189–207. ISSN: 1471082X. DOI: 10.1191/1471082X06st118oa.

[75] Wenjie Wang and Jun Yan. *Regression Spline Functions and Classes, Package 'splines2' Version 0.4.5*. 2021. DOI: 10.6339/21-jds1020. URL: https://wwenjie.org/splines2.

[76] Roger Kroenker et al. *Handbook of Quantile Regression*. Chapman and Hall/CRC, 2018. ISBN: 9781498725286.

[77] Alfred Galichon. "Quantile and Probability Curves Without Crossing." In: *Econometrica* 78.3 (2010), pages 1093–1125. ISSN: 0012-9682. DOI: 10.3982/ecta7880.

[78] Vito M.R. Muggeo et al. "Estimating growth charts via nonparametric quantile regression: A practical framework with application in ecology." In: *Environmental and Ecological Statistics* 20.4 (2013), pages 519–531. ISSN: 13528505. DOI: 10.1007/s10651-012-0232-1.

[79] Vito M.R. Muggeo. "Using the R package quantregGrowth: some examples." In: March (2018). DOI: 10.13140/RG.2.2.12924.85122.

[80] Vito M.R. Muggeo et al. "Multiple smoothing parameters selection in additive regression quantiles." In: *Statistical Modelling* x (2020), pages 1–21. ISSN: 14770342. DOI: 10.1177/1471082X20929802.

[81] Vito M.R. Muggeo. *quantregGrowth*. URL: https://cran.r-project.org/web/packages/quantregGrowth/quantregGrowth.pdf.

[82] Javier Contreras et al. "ARIMA models to predict next-day electricity prices." In: *IEEE Transactions on Power Systems* 18.3 (2003), pages 1014–1020. ISSN: 08858950. DOI: 10.1109/TPWRS.2002.804943.

[83] Erkan Erdogdu. "Electricity demand analysis using cointegration and ARIMA modelling: A case study of Turkey." In: *Energy Policy* 35.2 (2007), pages 1129–1146. ISSN: 03014215. DOI: 10.1016/j.enpol.2006.02.013.

[84] Afshin Salehian. "ARIMA Time Series Modeling for Forecasting Thermal Rating of Transmission Lines." In: *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference* 3 (2003), pages 875–879. DOI: 10.1109/tdc.2003.1335052.

[85] Snezhinka Lubomirova Zaharieva et al. "Arima Approach for Forecasting Temperature in A Residential Premises Part 2." In: *2021 20th International Symposium INFOTEH-JAHORINA, INFOTEH 2021 - Proceedings* (2021), pages 17–19. DOI: 10.1109/INFOTEH51037.2021.9400674.

[86] Henrik Madsen. *Time Series Analysis*. Chapman and Hall/CRC, 2007. ISBN: 9781420059687. DOI: doi-org.proxy.findit.cvt.dk/10.1201/9781420059687.

[87] Henrik Spliid. "Estimation of Multivariate Time Series with Regression Variables A time series program in R." In: *Journal of Time Series Analysis* 8.1 (2016), pages 95–109. ISSN: 14679892. DOI: 10.13140/RG.2.2.35854.36165. URL: www.imm.dtu/DSCC.

[88] Henrik Spliid. "A fast estimation method for the vector autoregressive moving average model with exogenous variables." In: *Journal of the American Statistical Association* 78.384 (1983), pages 843–849. ISSN: 1537274X. DOI: 10.1080/01621459.1983.10477030.

[89] Jan Emil Banning Iversen et al. "Leveraging stochastic differential equations for probabilistic forecasting of wind power using a dynamic power curve." In: *John Wiley and Sons Ltd* (2017). ISSN: 10991824 and 10954244. DOI: 10.1002/we.1988.

[90] Peder Bacher and Henrik Madsen. "Identifying suitable models for the heat dynamics of buildings." In: *Energy and Buildings* 43.7 (2011), pages 1511–1522. ISSN: 03787788. DOI: 10.1016/j.enbuild.2011.02.005. URL: http://dx.doi.org/10.1016/j.enbuild.2011.02.005.

[91] Rune Juhl et al. "Modeling and Prediction Using Stochastic Differential Equations." In: *Lecture Notes in Bioengineering*. Springer, 2016, pages 183–209. ISBN: 9783319259130. DOI: 10.1007/978-3-319-25913-0{_}10.

[92] Henrik Madsen and Jan Holst. "Modelling Non-Linear and Non-Stationary Time Series." In: *Lecture Notes*. December. Kgs. Lyngby, Denmark: Technical University of Denmark, Dpt. of Informatics and Mathematical Modeling, 2006.

[93] Jiajun Lyu et al. “Seasonal variation of residential cooling use behaviour derived from energy demand data and stochastic building energy simulation.” In: *Journal of Building Engineering* 49.December 2021 (2022), page 104067. ISSN: 23527102. DOI: 10.1016/j.jobe.2022.104067. URL: https://doi.org/10.1016/j.jobe.2022.104067.

[94] D. Stenzel and R. Witzmann. “Modelling the Aggregate Load Behaviour of European Regions and Its Influence on Dynamic and Static Voltage Stability.” In: *CIRED 2021 - The 26th International Conference and Exhibition on Electricity Distribution, 2021*. The Institution of Engineering and Technology, 2021, pages 2391–2395. DOI: 10.1049/icp.2021.1824.

[95] Abdelmonaem Jornaz and V. A. Samaranayake. “A multi-step approach to modeling the 24-hour daily profiles of electricity load using daily splines.” In: *Energies* 12.21 (2019). ISSN: 19961073. DOI: 10.3390/en12214169.

[96] Adrienn Dineva and Istvan Vajda. “Insulation State Assessment of Rotating Electrical Machines by Employing Generalized Additive Models.” In: *2020 11th International Conference on Electrical Power Drive Systems, ICEPDS 2020 - Proceedings* (2020), pages 8–12. DOI: 10.1109/ICEPDS47235.2020.9249328.

[97] Simon N. Wood et al. “Generalized Additive Models for Gigadata: Modeling the U.K. Black Smoke Network Daily Data.” In: *Journal of the American Statistical Association* 112.519 (2017), pages 1199–1210. ISSN: 1537274X. DOI: 10.1080/01621459.2016.1195744. URL: https://doi.org/10.1080/01621459.2016.1195744.

[98] Simon N. Wood. *Generalized Additive Models : An Introduction with R*. Second Edi. CRC Press, 2017. DOI: 10.1201/9781315370279-4.

[99] S. N. Wood. “Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models.” In: *Journal of the Royal Statistical Society (B)* 73.1 (2011), pages 3–36.

[100] S.N. Wood et al. “Smoothing parameter and model selection for general smooth models (with discussion).” In: *Journal of the American Statistical Association* 111 (2016), pages 1548–1575.

[101] S. N. Wood. “Stable and efficient multiple smoothing parameter estimation for generalized additive models.” In: *Journal of the American Statistical Association* 99.467 (2004), pages 673–686.

[102] S. N. Wood. “Thin-plate regression splines.” In: *Journal of the Royal Statistical Society (B)* 65.1 (2003), pages 95–114.

[103] M. S. Dalila, M. N. Khalid, and M. Md Shah. “Distribution transformer losses evaluation under non-linear load.” In: *AUPEC’09 - 19th Australasian Universities Power Engineering Conference: Sustainable Energy Technologies and Systems* (2009), pages 1–6.

[104] Shamsodin Taheri et al. “Effect of power system harmonics on transformer loading capability and hot spot temperature.” In: *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering: Vision for a Greener Future, CCECE 2012*. IEEE, 2012. ISBN: 9781467314336. DOI: 10.1109/CCECE.2012.6334834.

[105] Zhiming Huang et al. “Study on the Influence of Harmonics on the Magnetic Leakage Field and Temperature Field of 500 kV Connected Transformer.” In: *ICEMPE 2019 - 2nd International Conference on Electrical Materials and Power Equipment, Proceedings*. IEEE, 2019, pages 596–600. ISBN: 9781538684344. DOI: 10.1109/ICEMPE.2019.8727246.

[106] Bhaba Das, Thahirah Syed Jalal, and Fiona J.Stevens McFadden. “Comparison and extension of IEC thermal models for dynamic rating of distribution transformers.” In: *2016 IEEE International Conference on Power System Technology, POWERCON 2016*. IEEE, 2016, pages 1–8. ISBN: 9781467388481. DOI: 10.1109/POWERCON.2016.7753896.

[107] Edgar Alfredo Juarez-Balderas et al. “Hot-spot temperature forecasting of the instrument transformer using an artificial neural network.” In: *IEEE Access* 8 (2020), pages 164392–164406. ISSN: 21693536. DOI: 10.1109/ACCESS.2020.3021673.

[108] Lieven Degroote et al. “Neutral-point shifting and voltage unbalance due to single-phase DG units in low voltage distribution networks.” In: *2009 IEEE Bucharest PowerTech: Innovative Ideas Toward the Electrical Grid of the Future* (2009), pages 1–8. DOI: 10.1109/PTC.2009.5281998.

[109] Tack Hyun Jung et al. “Voltage Regulation Method for Voltage Drop Compensation and Unbalance Reduction in Bipolar Low-Voltage DC Distribution System.” In: *IEEE Transactions on Power Delivery* 33.1 (2018), pages 141–149. ISSN: 08858977. DOI: 10.1109/TPWRD.2017.2694836.

[110] Doyun Kim, Justin Migo Dolot, and Hwachang Song. “Distribution System State Estimation Using Model-Optimized Neural Networks.” In: *Applied Sciences (Switzerland)* 12.4 (February 2022). ISSN: 20763417. DOI: 10.3390/app12042073.

[111] Jan Hendrik Menke, Nils Bornhorst, and Martin Braun. “Distribution system monitoring for smart power grids with distributed generation using artificial neural networks.” In: *International Journal of Electrical Power and Energy Systems* 113 (December 2019), pages 472–480. ISSN: 01420615. DOI: 10.1016/j.ijepes.2019.05.057.

[112] Valentin Rigoni, Alireza Soroudi, and Andrew Keane. “Use of fitted polynomials for the decentralised estimation of network variables in unbalanced radial LV feeders.” In: *IET Generation, Transmission and Distribution* 14.12 (2020), pages 2368–2377. ISSN: 17518687. DOI: 10.1049/iet-gtd.2019.1461.

[113] European Union. "Directive of 2009/72/EC of the European Parliament and of the Council of 13 July 2009 Concerning Common Rules for the Internal Market in Electricity and Repealing Directive 2003/54/EC." In: *Official Journal of the European Union* L211.August (2009), pages L 211/55 –L 211/93. ISSN: 0036-8075.

[114] Marco Di Clerico et al. "Prosumers' Battery Electrical Storage Systems: new ancillary services, impact on network planning and operation." In: *CIGRE Session 46*. 2016. ISBN: 9782858730803. DOI: 10.1109/PESGM.2012.6345361. URL: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5211198.

[115] Marcus Müller et al. "Evaluation of grid-level adaptability for stationary battery energy storage system applications in Europe." In: *Journal of Energy Storage* 9 (2017), pages 1–11. ISSN: 2352152X. DOI: 10.1016/j.est.2016.11.005.

[116] Matthias Resch et al. "Technical and economic comparison of grid supportive vanadium redox flow batteries for primary control reserve and community electricity storage in Germany." In: *International Journal of Energy Research* 43.1 (2019), pages 337–357. ISSN: 1099114X. DOI: 10.1002/er.4269.

[117] Edward Barbour et al. "Community energy storage: A smart choice for the smart grid?" In: *Applied Energy* 212 (2018), pages 489–497. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.12.056.

[118] Tilman Weckesser et al. "Renewable Energy Communities: Optimal sizing and distribution grid impact of photo-voltaics and battery storage." In: *Applied Energy* 301.June (2021), page 117408. ISSN: 03062619. DOI: 10.1016/j.apenergy.2021.117408. URL: https://doi.org/10.1016/j.apenergy.2021.117408.

[119] Shady A. El-Batawy and Walid G. Morsi. "Optimal design of community battery energy storage systems with prosumers owning electric vehicles." In: *IEEE Transactions on Industrial Informatics* 14.5 (2018), pages 1920–1931. ISSN: 15513203. DOI: 10.1109/TII.2017.2752464.

Part II

Publications

1 Paper A

Emma M. V. Blomgren, Razgar Ebrahimy, Seyyed Ali Pourmousavi Kani, Jan Kloppenborg Møller, Francesco D'Ettorre, Mohsen Banaei, Henrik Madsen,
"Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty",
Proceedings of 2021 IEEE PES Innovative Smart Grid Technologies Conference Europe (2021)

Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty

1st Emma M. V. Blomgren,
2nd Razgar Ebrahimy, 4th Jan Kloppenborg Møller,
5th Francesco D'Ettorre, 6th Mohsen Banaei, 7th Henrik Madsen
Department of Applied Mathematics and Computer Science
DTU Compute, Copenhagen, Denmark
emvb@dtu.dk, raze@dtu.dk, jkmo@dtu.dk,
frade@dtu.dk, moban@dtu.dk, hmad@dtu.dk

3rd Ali Pourmousavi Kani
School of Electrical and Electronic Engineering
University of Adelaide,
Adelaide, Australia
a.pourm@adelaide.edu.au

***Abstract*—Aggregators are expected to become an inevitable entity in future power system operation, playing a key role in unlocking flexibility at the edge of the grid. One of the main barriers to aggregators entering the market is the lack of appropriate models for the price elasticity of flexible demand, which can properly address time dependent uncertainty as well as non-linear and stochastic behavior of end-users in response to time varying prices. In this paper, we develop a probabilistic price elasticity model utilizing quantile regression and B-splines with penalties. The proposed model is tested using data from residential and industrial customers by assuming automation through energy management systems. Additionally, we show an application of the proposed method in quantifying the number of consumers needed to achieve a certain amount of flexibility through a set of simulation studies.**

***Index Terms*—Flexibility, data-driven modelling, quantile regression, B-splines, industrial and residential consumers**

I. Introduction

To ensure a green transition of the energy system and enable further integration of renewable energy resources into the power grid, new and green flexibility resources will be necessary for the day-to-day grid operation [1]. Advancements in behind-the-meter controllable technologies along with automation, i.e., Energy Management Systems (EMS), enable both traditional electricity consumers and prosumers to provide flexibility. However, the flexibility provided by individual consumers/prosumers is too small for direct participation in the wholesale electricity market. Thus, there is a need for aggregation of these resources.

While there is general agreement that aggregators will become an inevitable part of the future power system, several issues must be addressed before their full potential can be realised in practice. One of the main barriers is the uncertainty associated with Demand Response (DR) of diverse types of consumers, especially through indirect load control by time-varying price signals. Aggregators will need information and methods to evaluate DR from flexible resources to develop strategies on how to optimally manage these resources. This calls for probabilistic price elasticity models of prosumers, specifically developed for aggregators.

This work was supported by the Flexible Energy Denmark (FED) project funded by Innovation Fund Denmark under Grant No. 8090-00069B.

As of today, several papers have investigated the long-term price elasticity of electricity consumption, e.g., [2] in the US. Unfortunately, the proposed approaches for long-term elasticity cannot be used for day-to-day aggregator operations and are substantially different in nature. In [3], elasticity values and calculation methods are summarised from several studies. The authors found 538 observations of price elasticity estimations in literature, where the average short-term price elasticity was -0.201. However, the authors did not provide a definition of the short-term price elasticity; thus, the granularity of the price elasticity is unclear. Furthermore, Miller et al. [4] used three data sets for residential consumers in the US to estimate price elasticity, which varies between -0.2 and -0.8 for the given data sets. The variation in price elasticity estimation is an indicator of the uncertainty associated with the underlying phenomenon, in which point estimations may not be able to explain it sufficiently. Hence, there is a need for dynamic and probabilistic elasticity models. Furthermore, a probabilistic consumer flexibility model is essential to apply risk-based methods (e.g., CVaR in [5]) and scenario-based stochastic programming (e.g., [6]) for optimal operation of aggregators. In [7], the DR from households responding to economic incentives for critical load and peak shifting is investigated. The authors concluded that automation strongly increases the price responsiveness, while manual DR can only make long-run adjustments. Thus, automation, e.g., EMS, is necessary to fully realise the potential of demand-side flexibility.

There are also a few papers on short-term price elasticity. J. M. Gillan [8] investigated the short-term price elasticity for residential electricity consumers in California. However, the study only concerns explicit DR. In [9], the authors looked at the price elasticity in the real-time and day-ahead market. These studies consider all customers in the data set with possible effects of customer fatigue in manual DR. Furthermore, in [10], the authors developed an artificial neural networks solution to learn the behavior of a single electricity consumer. The models developed in these studies are deterministic, which is less useful for the aggregators' operation in an uncertain environment with low profit margin. Probabilistic DR models are presented in the literature, e.g., [11] in which conditional

Preprint submitted to
Proceedings of 2021 IEEE PES Innovative Smart Grid Technologies Conference Europe

probability density for future demand is used for prediction of demand. However, the model is based on two-way communication and customers react manually to the electricity price signals. A quantile and linear regression-based model of DR is proposed for a single customer in [12] based on the Spanish ADRESS project, in which 260 residential consumers with Home Energy Management Systems (HEMS) participated. Nevertheless, these studies only considered individual customers from a specific sector, explicit mechanism for DR, and linear modeling approaches.

Furthermore, numerous studies have developed DR models using consumer price elasticity, e.g., [13] in which the short-term price elasticity is assumed to be -0.2 in their probabilistic demand curve model. Also, [14] utilized a function of elasticity to describe DR customer behavior. These studies did not develop a model for the elasticity in particular nor did they discuss the elasticity values/models. Analysing the potential of flexibility to reduce power peak consumption in Northern Europe, [15] found that the flexibility potential varies over different sectors (residential, commercial and industrial). To address the aggregators' interest, the price-load models should therefore be applicable to consumers from different sectors.

This paper aims to provide a methodology for quantifying the consumer/prosumer price elasticity and associated uncertainty. Our research contributes to this field by making hourly price elasticity models using quantile regression and B-splines with penalties. Thereby, we develop a non-linear probabilistic model, reflecting the uncertainty of consumer/prosumer flexibility. We then use the proposed model to quantify the number of consumers required to achieve ± 1 MW flexibility from different sectors in a probabilistic manner. To create a price elasticity model suitable for the aggregators' operation, price-load flexibility data are required. However, due to the scarcity of experimental data involving consumers equipped with automation, synthesized data from [16]–[18] is used for modeling and simulation studies. The contributions of the paper can be summarized as follows:

- We develop hourly flexibility models of consumers that represent non-linearity of the flexibility resources and consumer behavior.
- A probabilistic approach is adopted using quantile regression to model the uncertainty of DR resources, which does not require a probability density function (pdf) as a priori. Therefore, its application is not limited to a certain type of pdf.
- We utilize data from residential and industrial sectors, which are assumed to be equipped with automation through EMS, but still allow for customers' preferences and stochastic behavior.

To the best of our knowledge, such a price-response elasticity model of consumers from different sectors has not been presented in the literature.

The rest of the paper is organized as follows. Section II presents the methodology of applying quantile regression and B-splines with penalties, as well as the methodology to quantify the number of consumers required to obtain a certain amount of flexibility. Section III reports and analyzes the most important findings and Section IV concludes the paper.

II. Methodology

In this section, we describe the methodology of applying quantile regression (QR) and B-splines with penalties (also known as P-splines) to the price-load data set. Thereafter, we describe the methodology to estimate the number of activated customers required to reach ± 1 MW bid size using the QRs.

In this work, up-regulation refers to a reduction in consumption which is assumed to be a result from positive price on top of the baseline price. Hence, down-regulation means an increase in consumption, which is assumed to be a response to negative incentives. The flexibility behavior appears to be quite different for up- and down-regulation (i.e., to positive and negative price deviations). Therefore, we apply the regression methodology for up- and down-regulation separately.

A. Quantile regression and B-splines with penalties

The magnitude of flexibility, L_h, at a certain hour, h, given a price deviation π, can be described by some unknown pdf, $f(L_h)$ with cumulative density function (cdf) (1).

$$P(L_h|\pi_h) = \int_{-\infty}^{L_h} f(x;\pi_h)dx \tag{1}$$

where the flexibility magnitude, L_h, is the change in load from the baseline at a certain hour, h, while the price, $\pi_h \in [-0.75, 0.75]$ is the deviation in electricity price from a certain baseline electricity price (assumed 2.25 DKK/kWh as in [16]) for the same hour h. For a specific data set C, we can define the conditional distribution as $P(L_h|L_{h,c}, \pi_{h,c})$.

One approach to find the conditional distribution from a data set is to utilize QR. The pdf of DR is unknown and can vary depending on various factors such as type of customers, time of the day and weather. Using QR avoids parametric assumptions on the pdf. Thus, QR is especially well suited for this purpose, since it does not assume any distribution a priori [19].

When applying QR to find the cdf, quantiles may cross. If this occurs, it would imply negative probability according to the definition of quantiles. Hence, a method that does not result in crossings should be applied. One methodology is proposed in [19], in which constraints are applied on the fitted parameters in the linear programming to ensure non-crossing regressions. In this work, we utilize this method through the R package *quantregGrowth version 0.4.3* [19]–[21]. The QR applied can be described as follows (2).

$$Q_{\tau,h}(\hat{L}_h|L_{c,h}\pi_{c,h}) = \sum_{j=1}^{n} \hat{a}_{j,h} B_{j,h}(\pi_h; q) \tag{2}$$

where $Q_\tau(\hat{L}_h|L_{c,h}, \pi_{c,h})$ is the estimated function for a quantile τ, given the data set C for hour h. The dimension of the problem is n and is less or equal to the number of price signals. In the loss function for the QR, there is a penalty

term $\lambda \sum_j |\Delta_j^d|$, as defined in the R package *quantregGrowth version 0.4.3* [19]–[21], where the order of the difference operator (d) is set to 3. The penalty term penalizes overfitting; thus, it affects the smoothness of the regression. We test for several values of the weight of the penalty, λ, and allow the algorithm to choose the best value through cross validation as described in [19]. We also set the degree of the B-splines to 3. Equation (2) is initially applied to quantiles $\tau = [0.1, 0.2, \cdots, 0.9]$

From analyzing the load versus price data set $(L_{c,h}, \pi_{c,h})$, the consumer flexibility appears to be non-linear. Thus, we apply B-splines with penalties using the *ps()* function from the *quantregGrowth version 0.4.3* package in R as in [19].

We apply further modifications in the *ps()* function by limiting the B-splines function to be monotonically non-increasing. This is a reasonable implementation given that the higher the price the lower consumption is expected. If the customers are equipped with EMS, it is not expected that a higher price will give a higher consumption. It should be mentioned that, due to the stochastic behavior of flexible electricity consumers, it can happen that a higher price yields a higher consumption although a lower consumption is expected. However, this would not be driven by the price, but rather happen due to other reasons, e.g., rebound effect of critical loads. Thus, it should not be reflected in the fitted spline functions, but rather by the uncertainty, i.e., the shape of the cdf. Alternatively, applying more explanatory variables, such as the rebound effect, may fix this issue, although this has not been investigated in this article.

B. Estimation of activated customers

As most markets require a minimum bid size to participate in the market, aggregators are interested in estimating the required number of customers to reach a minimum bid size. This estimation can be utilized to determine the number of customers in the pool. Alternatively, it can be used to evaluate the participation of customers from different categories or sectors by estimating the number of customers that should be activated at a certain hour from an already existing customer pool. Here, we set the required bid size of a hypothetical market to ± 1 MW and describe how to make such estimates from the proposed QR models.

As described above, the cdf for a certain price can be extracted from the QR. Here, we demonstrate this by extracting the cdf for 0.5 DKK/kWh and -0.5 DKK/kWh price variations, which are the medians for the positive and negative price deviations. The extracted values are samples from the cdf for their respective τ. To get a full cdf, we make a piecewise linear regression between the extracted points. Since the uncertainty is higher in the tail probabilities of the cdf, where the cdf tends to vary more, we add quantiles $\tau_k = [0.01, 0.05, 0.95, 0.99]$ to the QR model. To deal with the end points of the QRs, i.e., from $\tau = 0.01$ to 0 and $\tau = 0.99$ to 1, a different approach is needed. Since the data in this case behave well and do not reach negative loads for negative prices, it is possible to make a linear regression between the QR for $\tau = 0.01$ and zero for down regulation. The same is valid for up-regulation and $\tau = 0.99$, since the load at $\tau = 1$ does not go above zero. For the case of down-regulation and $\tau > 0.99$, there is no natural maximum that can be extracted from the data. Therefore, we fit an exponential function to describe the cdf from $\tau = 0.99$ to $\tau = 1$. For the up-regulation, it is assumed that the total consumption will not be negative and a linear regression is made from $Q_{\tau=0.01}(L_h)$ to the negative value of what they are already consuming at that hour, i.e., the base load, $-L_{base,h}$.

Using a uniform distribution, $\mathcal{U}(0, 1)$, we then simulate from the cdf until 1 MW or -1 MW is reached. In other words, we find i such that (3) for down-regulation is satisfied for $L_{bid} = 1MW$; similarly for up-regulation that (4) is satisfied for $L_{bid} = -1MW$.

$$L_{bid} \geq \sum_{j=1}^{i} F^{-1}(r_j) \tag{3}$$

$$L_{bid} \leq \sum_{j=1}^{I} F^{-1}(r_j) \tag{4}$$

where $r_j \sim \mathcal{U}(0, 1)$. This is repeated 1000 times, giving an estimate of mean and variance for how many customers are required to reach 1 or -1 MW, respectively, for a certain customer cluster and certain hour, given a price deviation.

III. Results and Analysis

In this section, we present some of the findings from applying the methodology described in Section II. In addition, we present the results of the simulations of activated customers from the presented cases.

In the data utilized for this study, the baseline electricity price is assumed to be 2.25 DKK/kWh with a variable price component, $\pi_h \in [-0.75, 0.75]$. Consumers categories are aggregated to form 3 clusters (residential, light industry, heavy industry). Further details of the data are described in [22], accompanied by [16]–[18].

A. Flexibility model of demand with QR

In this work, we apply 9 QRs such that τ_k takes on values from 0.1 to 0.9, linearly spaced with 0.1 interval. Of course, when applying the methodology to a specific case, this could be changed depending on the aggregator's needs in terms of uncertainty analysis.

In Fig. 1, the results for the residential cluster at hour 12 are provided, while the results for the heavy industry cluster can be seen in Fig. 2. In both figures, it is shown that a point estimation would not be sufficient and that a confidence interval would not give the full picture of the possible outcomes of flexibility. Meanwhile, the QR curves span over the entire data set. It can also be seen that the response function is not linear, which supports the necessity of a non-linear method. Additionally, it can be observed that crossings of

the QR curves are avoided through the application of the quantregGrowth package [19], [21]. From both graphs (Figs. 1 and 2), cdfs can be extracted by simply fixing the electricity price deviation and obtaining the flexibility magnitude from the QRs, arranging the values of τ_k in an increasing order.

For cluster 1 (Fig. 1), the variance is quite high for up-regulation compared to cluster 3. For the case of down-regulation, the flexibility has more dense QR curves in the upper region of the data set. This implies a more skewed distribution function, with higher probabilities of the customers actively responding. As a result, the widely-used normal distribution in such studies does not properly describe the flexibility. This justifies the application of QRs with which a priori pdf is not required.

Looking at cluster 3 (Fig. 2), however, it can be noticed that the variance is higher for the down-regulation. It is also seen that the pdf is skewed towards non-responsive behavior, unless the price deviation is high. For the up-regulation, on the other hand, the observed reactions are smaller in magnitude but more certain. Here, the response also appears to have a bimodal distribution that is well captured by the QR curves by being more dense around these lines. If simulations were made from the up-regulation case, two most probable scenarios would be observed.

One reason for this bimodal distribution for up-regulation in cluster 3 could be that two categories in the cluster, "Non-metallic" and "Other industries", have quite similar behavior, whereas the behavior of the category "Chemical" differs significantly from the other two. This can be seen in the assumptions made for the categories (Table I in [16]).

Furthermore, the lack of flexibility to down-regulation prices can be explained by the industrial customers' strict technical and operational constraints compared to the residential ones. However, they are business-driven: for the right price deviation at the right time, they react to optimize their energy consumption expenditure. In Fig. 2, it can be seen that for hour 12, it

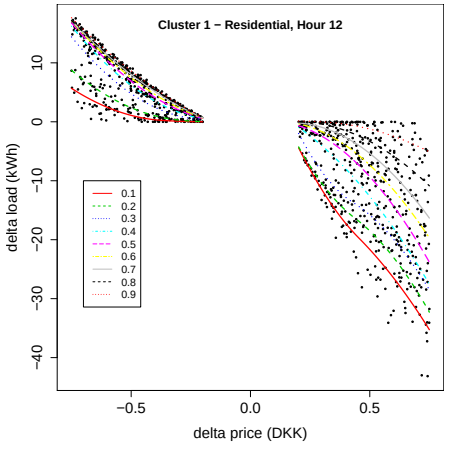


Fig. 1. Flexibility of 280 residential customers in cluster 1. The graph shows hour 12 with 9 quantile regressions.

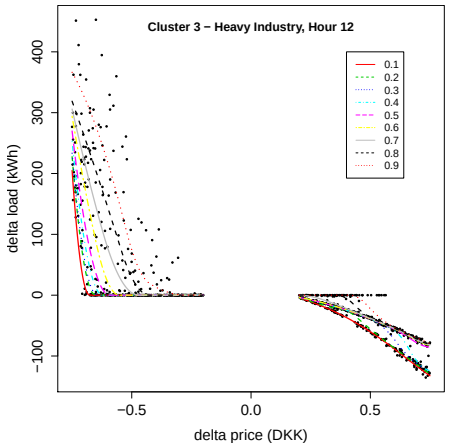


Fig. 2. Flexibility of 210 heavy industry customers in cluster 3. The graph shows hour 12 with 9 quantile regressions.

is not favorable for the industrial consumers to increase their consumption, unless the price deviation is very high.

The fact that industries have technical constraints and generally operate in more safety-driven ways than residential consumers can also be seen in the comparison between clusters 1 and 3. Thus, the response from the residential cluster has a higher variance, while for the heavy industry cluster, the patterns are more recognizable, probably due to their routine day-to-day operation.

B. Customer activation

In this section, we present the results from the 1000 simulations for up- and down-regulation respectively, to achieve ± 1 MW in load deviation through the flexibility models, as described in Section II. To better account for the uncertainty in the lower and upper end of the cdf, we also include QRs for $\tau_k = [0.01, 0.05, 0.95, 0.99]$, as described in Section II.

The results of the simulations for cluster 1 (residential), are presented in Fig. 3. In general, it can be seen that the uncertainty in the number of required customers is larger for up-regulation compared to down-regulation. This is a direct result from the larger variance on the up-regulation side in Fig. 1. It can also be observed that significantly more customers are needed for down-regulation compared to up-regulation for hours 2, 3, and 4, as well as hours 12, 13, and 14. This could be due to the rebound effect and the fact that the residential customers are asleep in the earlier hours or not at home in the middle of the day. Thus, there is no need for increasing their electricity consumption.

For cluster 2 (light industry), the results are visualized in the box-plot of Fig. 4, where it can be observed that significantly fewer customers are required to obtain -1 MW than 1 MW of flexibility. The number of electricity customers required for down-regulation in hours 1, 8, 17, and 24 is high, which means that there is not much flexibility (or willingness by the consumers to provide flexibility) to be activated; thus lower

commitment for the aggregator in the wholesale market can be suggested. For hours 6, 10, 11, 13, 20, 21, 22, and 23, we stopped the simulations at 2.8 million customers without reaching 1 MW. It should, however, be noted that for a larger price deviation, the 1 MW flexibility could be achieved.

For cluster 3 (heavy industry) in Fig. 5, it can be seen that fewer consumers are required to achieve -1 MW than 1 MW. This is in line with what is shown in Fig. 2. For a price deviation of -0.5 DKK/kWh, there is a 70-80% probability that there will be no reaction or a very small reaction from the customer cluster. On the other hand, for 0.5 DKK/kWh, the probability that the cluster is not responsive to the incentives is less than 10%. It can also be seen that the number of required customers for both up- and down-regulation decreases over the day. The base-load consumption is lower during the early hours of the day and higher in the later hours. Seeing higher activity in the later part of the day suggests that there are more active loads to be offered for both up- and down-regulation.

Overall, these results could be further used by the aggregators to target the right consumers for flexibility provision. Aggregators can get an insight into how large a customer pool from different clusters they need and their associated uncertainty. In other words, they can estimate the number of activated customers that are required at each hour in the best and worst case scenarios. Such results can be used in operational risk assessments and meeting the minimum bid requirements. As another advantage of the proposed approach, the aggregators receive valuable insights into cluster integration. For instance, while more customers in cluster 2 are required for up-regulation in hours 19, 20, and 21, consumers in cluster 3 can provide a lot more flexibility in the same period of time. These complementary effects can be exploited by the aggregators.

IV. Conclusion

This paper presents an hourly price load model for implicit flexibility provision. QR and B-splines with penalties were applied to achieve a non-linear probabilistic model to capture the variance and uncertainty in the data. A future scenario was assumed, in which the customers are equipped with EMS and the model was applied to both residential and industrial electricity customers. From the QRs, the cdf for ± 0.5 DKK/kWh was extracted. From this study, we observed a higher uncertainty for up-regulation compared to down-regulation from the residential cluster. We also discovered that fewer customers were required for up-regulation compared to down-regulation from both light and heavy industry clusters.

From an aggregator's perspective, the simulation results can be further employed for risk assessments. For instance, scenarios can be defined, such as best, worst and most probable scenarios of required activated customers from an already existing customer pool. Additionally, the results may assist aggregators in determining the customer segments with the highest flexibility and willingness for their business.

Furthermore, QR and extracted cdfs could be used for

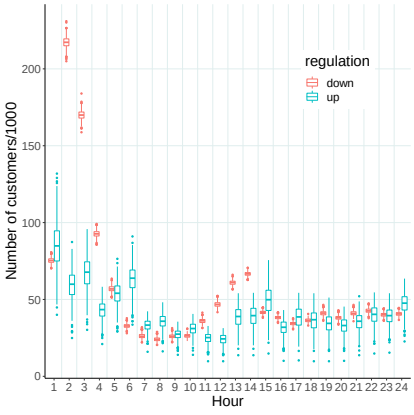


Fig. 3. Number of activated electricity customers required to achieve ± 1 MW in DR from cluster 1 (residential cluster).

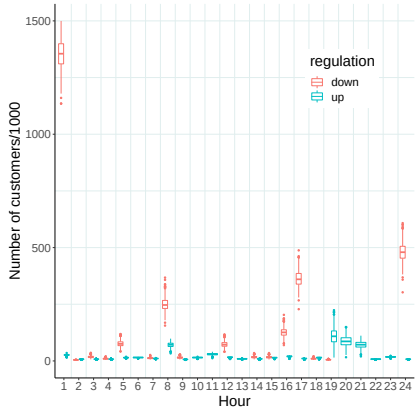


Fig. 4. Number of activated electricity customers required to achieve ± 1 MW in DR from cluster 2 (light industry cluster).

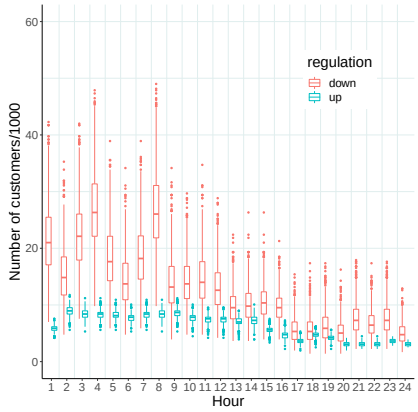


Fig. 5. Number of activated electricity customers required to achieve ± 1 MW in DR from cluster 3 (heavy industry cluster).

scenario generation for flexibility. Alternatively, they can also be utilized to generate inputs for CVaR models.

Acknowledgements

The authors thank Dr. Giulia De Zotti for providing the data utilized in this study.

References

[1] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets*. 2014.

[2] C. K. Woo, Y. Liu, J. Zarnikau, A. Shiu, X. Luo, and F. Kahrl, "Price elasticities of retail energy demands in the United States: New evidence from a panel of monthly data for 2001–2016," *Applied Energy*, vol. 222, pp. 460–474, Jul. 2018.

[3] X. Labandeira, J. M. Labeaga, and X. López-Otero, "A meta-analysis on the price elasticity of energy demand," *Energy Policy*, vol. 102, pp. 549–568, 2017.

[4] M. Miller and A. Alberini, "Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data," *Energy Policy*, vol. 97, pp. 235–249, Oct. 2016.

[5] H. Rashidizadeh-Kermani, M. Vahedipour-Dahraie, M. Shafie-khah, and P. Siano, "A Regret-based Stochastic Bi-level Framework for Scheduling of DR Aggregator under Uncertainties," *IEEE Transactions on Smart Grid*, vol. 3053, no. c, pp. 1–1, 2020.

[6] M. Brolin and C. Sandels, "Controlling a retailer's short-term financial risk exposure using demand response," *IET Generation, Transmission and Distribution*, vol. 13, no. 22, pp. 5160–5170, 2019.

[7] B. K. Bollinger and W. R. Hartmann, "Information vs. Automation and implications for dynamic pricing," *Management Science*, vol. 66, no. 1, pp. 290–314, 2020.

[8] J. M. Gillan, *Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption*, 2017. [Online]. Available: https://are.berkeley.edu/sites/are.berkeley.edu/files/job-candidates/paper/JMP_Gillan_most_recent_draft.pdf.

[9] P. Damien, R. Fuentes-García, R. H. Mena, and J. Zarnikau, "Impacts of day-ahead versus real-time market prices on wholesale electricity demand in Texas," *Energy Economics*, vol. 81, pp. 259–272, 2019.

[10] T. A. Nakabi and P. Toivanen, "An ANN-based model for learning individual customer behavior in response to electricity prices," *Sustainable Energy, Grids and Networks*, vol. 18, p. 100 212, 2019.

[11] S. Ghosh, X. A. Sun, and X. Zhang, "Consumer profiling for demand response programs in smart grids," *2012 IEEE Innovative Smart Grid Technologies - Asia, ISGT Asia 2012*, pp. 1–6, 2012.

[12] M. Vallés, A. Bello, J. Reneses, and P. Frías, "Probabilistic characterization of electricity consumer responsiveness to economic incentives," *Applied Energy*, vol. 216, pp. 296–310, Apr. 2018.

[13] Z. Zhou and A. Botterud, "Price responsive demand for operating reserves and energy in electricity markets with wind power," *IEEE Power and Energy Society General Meeting*, pp. 2–6, 2013.

[14] A. Niromandfam, A. S. Yazdankhah, and R. Kazemzadeh, "Modeling demand response based on utility function considering wind profit maximization in the day-ahead market," *Journal of Cleaner Production*, vol. 251, p. 119 317, 2020.

[15] L. Söder, P. D. Lund, H. Koduvere, *et al.*, "A review of demand side flexibility potential in Northern Europe," *Renewable and Sustainable Energy Reviews*, vol. 91, pp. 654–664, Aug. 2018.

[16] G. De Zotti, S. A. Pourmousavi, J. M. Morales, H. Madsen, and N. K. Poulsen, "Consumers' Flexibility Estimation at the TSO Level for Balancing Services," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1918–1930, May 2019.

[17] G. De Zotti, D. Guericke, S. A. Pourmousavi, J. M. Morales, H. Madsen, and N. K. Poulsen, "Analysis of rebound effect modelling for flexible electrical consumers," *IFAC-PapersOnLine*, vol. 52(4), pp. 6–11, 2019.

[18] G. De Zotti, S. A. Pourmousavi, J. M. Morales, H. Madsen, and N. K. Poulsen, "A Control-based Method to Meet TSO and DSO Ancillary Services Needs by Flexible End-Users," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1868–1880, 2019.

[19] V. M. Muggeo, M. Sciandra, A. Tomasello, and S. Calvo, "Estimating growth charts via nonparametric quantile regression: A practical framework with application in ecology," *Environmental and Ecological Statistics*, vol. 20, no. 4, pp. 519–531, 2013.

[20] V. M. Muggeo, *quantregGrowth*. [Online]. Available: https://cran.r-project.org/web/packages/quantregGrowth/quantregGrowth.pdf (visited on 12/01/2020).

[21] V. M. Muggeo, F. Torretta, P. H. Eilers, M. Sciandra, and M. Attanasio, "Multiple smoothing parameters selection in additive regression quantiles," *Statistical Modelling*, no. x, pp. 1–21, 2020.

[22] E. M. V. Blomgren, G. De Zotti, R. Ebrahimy, S. A. Pourmousavi Kani, and H. Madsen, *Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty: Data presentation*, 2021. [Online]. Available: https://orbit.dtu.dk/en/publications/behind-the-meter-energy-flexibility-modelling-for-aggregator-oper (visited on 07/28/2021).

2 Supplementary Information for paper A

Emma M. V. Blomgren, Giulia de Zotti, Razgar Ebrahimy, Ali Pourmousavi Kani*, Henrik Madsen,
"Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty - Data presentation",
Technical University of Denmark (2021)

Behind-the-Meter Energy Flexibility Modelling for Aggregator Operation with a Focus on Uncertainty - Data presentation

Emma M. V. Blomgren,
Giulia de Zotti, Razgar Ebrahimy, Ali Pourmousavi Kani*, Henrik Madsen
DTU Compute, Copenhagen, Denmark
**University of Adelaide,* Adelaide, Australia
emvb@dtu.dk, raze@dtu.dk, a.pourm@adelaide.edu.au, hmad@dtu.dk

Data Presentation

The data presented here is based on data from the Danish Elforbrugspanel project, which was a collaboration between the Danish TSO Energinet and Dansk Energi. The flexibility of consumers in response to different electricity prices is then synthesized using the method proposed in [1]. The electricity price is a combination of a baseline electricity prices, 2.25 DKK/kWh, and variable price component π_h, such that the total price is $2.25 + \pi_h$ DKK. The variable electricity price component is different for each hour and has a cap, such that $\pi_h \in [-0.75, 0.75]$. There are 29 different customer categories in the data with different responsiveness to dynamic prices (see Table I in [1] for details). In our work, 70 unique customers of each load category are exposed to 1000 price-sets for one day in a simulation study. As previously mentioned, a future scenario is assumed where the electricity customers have EMS, in which they can set their personal preferences. A static rebound effect is assumed when generating the data according to [1]. For further details on the data and the method used for synthesizing, the reader is referred to [1]. The reason for synthesizing the flexibility of consumers is the lack of available data for research on price responsive demand with automation. It is assumed that the methodology and analytical approach can be applied to real world data when available.

From the aggregators' perspective, the baseline consumption is not of primary concern, but rather the deviation from the baseline due to a given price deviation is important. Thus, we build our models for load deviation versus price deviation data in this study. An example of the data is plotted in Fig. 1, showing the flexibility from 70 customers of category 3 (house with heating) in hours 3, 6, 12 and 19. It can be seen that category 3 is not so flexible at late night (hour 3), but as the day starts there is more active energy consumption to be utilized as a flexible resource. Fig. 1 shows the aggregated flexibility for 70 consumers of categories 1, 2, 3 and 4 respectively. The four categories form the "residential cluster", as they all belong to residential categories including "Apartment without heating" (cat. 1), "House without heating" (cat. 2), "House with heating" (cat. 3) and "Cottage" or alternatively "Summer house" (cat. 4). It is clear that the variance of the flexibility increases when aggregating these consumers together. In the aggregated response in Fig. 1, it can be observed that the response to the negative price changes (i.e., down-regulation) are smaller than the flexibility achieved for up-regulation (i.e., positive price deviations), possibly due to a 'natural maximum energy consumption' for residential consumers. For example, a consumer would probably not start the dishwasher just because it is possible, but rather when it is needed.

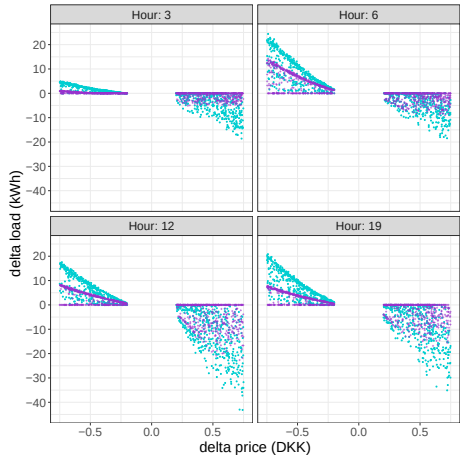


Fig. 1. **Purple**: flexibility of 70 customers in category 3 in hours 3, 6, 12 and 19. **Turquoise**: flexibility of consumers in cluster 1.

As previously mentioned, categories 1-4 are aggregated to create a residential cluster (cluster 1). Likewise, 70 consumers of categories 8-12 are aggregated to form a light industry cluster, called cluster 2 (including Food, Basic metal, Wood, Textile and Paper). 70 consumers of categories 13-15 are aggregated to form the heavy industry cluster (cluster 3), which includes Non-metallic, Chemical, and Other Industries.

References

[1] G. De Zotti, S. A. Pourmousavi, J. M. Morales, H. Madsen, and N. K. Poulsen, "Consumers' Flexibility Estimation at the TSO Level for Balancing Services," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1918–1930, May 2019.

3 Paper B

Emma M.V. Blomgren, Francesco D'Ettorre, Olof Samuelsson, Mohsen Banaei, Razgar Ebrahimy, Magne E. Rasmussen, Nicolaj H. Nielsen, Anton R. Larsen, Henrik Madsen,
"Grey-box modeling for dynamic rating of oil-immersed transformers in power distribution networks"
(submitted to) *Electric Power Systems Research*

Grey-box modeling for dynamic rating of oil-immersed transformers in power distribution networks

Emma M.V. Blomgren[a], Francesco D'Ettorre[a], Olof Samuelsson[b], Mohsen Banaei[a,*], Razgar Ebrahimy[a], Magne E. Rasmussen[a], Nicolaj H. Nielsen[a], Anton R. Larsen[a], Henrik Madsen[a]

[a] *Technical University of Denmark, Department of Applied Mathematics and Computer Science, Lyngby, Denmark*
[b] *Lund University, Faculty of Engineering, Industrial Electrical Engineering and Automation, Lund, Sweden*

Abstract

Power transformers are one of the most costly assets in power grids. Due to increasing electricity demand and levels of distributed generation, they are more and more often loaded above their rated limits. Transformer ratings are traditionally set as static limits, set in a controlled environment with conservative margins. Through dynamic transformer rating, the rating is instead adapted to the actual working conditions of the transformers. This can help distribution system operators (DSOs) to unlock unused capacity and postpone costly grid investments. To this end, real-time information of the transformer operating conditions, and in particular of its hot-spot and oil temperature, is required. This work proposes a grey-box model that can be used for online estimation and forecasting of the transformer temperature. It relies on a limited set of non-intrusive measurements and was developed using experimental data from a DSO in Jutland, Denmark. The thermal model has proven to be able to predict the temperature of the transformers with a high accuracy and low computational time, which is particularly relevant for online applications. With a six-hour prediction horizon the mean average error was 0.4 – 0.6 °C. By choosing a stochastic data-driven modeling approach we can also provide prediction intervals and account for the uncertainty.

Keywords: Grey-box modeling, Dynamic transformer rating, Distribution grid flexibility, Thermal model, Data-driven modeling

1. Introduction

The decarbonization pathways towards a carbon-neutral Europe are deeply reshaping the power system. Increasing levels of distributed generation (DG) together with the electrification of the heating and transport sectors (e.g. use of heat pumps and electric vehicles)

*Corresponding author
Email address: moban@dtu.dk (Mohsen Banaei)

enable a shift from passive to active distribution grids. The increased peaks and congestion that come with the increasing demand have resulted in transformers being more frequently loaded above their rated limits, i.e. nameplate ratings. The overloading could reduce their life expectancy, and also jeopardize the reliability of the entire network. This highlights the need for more dynamic grid operations.

Over the past decades, capacity expansion and replacement of existing grid assets have been the main measures taken by distribution system operators (DSOs) to keep the network running smoothly, while handling an increasing number of new connections and increasing levels of DG and low-carbon technologies (e.g. heat pumps). However, the traditional "connect and reinforce" operating model is economically and environmentally costly and takes time to implement. As a result, new methods to increase transformer capacity, while limiting their aging are needed to reduce the need for expensive network upgrades.

Power transformers are one of the most expensive assets in a power grid infrastructure [1]. Loading a transformer beyond its nameplate capacity increases the leakage flux to the core and outside, which heats the metallic parts of the transformer. This might further affect the internal thermal dynamics as the composition of the insulation oil might change and gas content increase [2]. As a consequence of the transformer losses (i.e. ohmic winding, core and stray losses), the temperature increases. If the transformer hot-spot (the area with the highest temperature) and oil temperature rise above the recommended thermal limits (given by manufacturer or see e.g. [2]), it could increase the insulation's aging rate and reduce transformer lifetime [1].

These limits vary according to the transformer type and cooling strategy. According to IEEE Std. C57.12.00-2015 the average and maximum (hottest-spot) winding temperature rise above ambient temperature shall not exceed 65 °C and 80 °C, respectively, at rated kVA when tested in accordance with IEEE Std. C57.12.90 (i.e. continuous ambient temperature of 40 °C for air-cooled transformers and of 30 °C for water-cooled transformers) [3]. However, such constant conditions are quite unusual during normal operation [4], as the environment within which the transformer is operated constantly changes. When the ambient temperature is lower than the rated temperature, a higher load can be allowed without increasing the transformer aging rate. Therefore, by adapting the transformer rating to its actual working conditions (i.e. dynamic rating), it is possible to unlock extra capacity, without violation of the safety margins [5], and achieve monetary saving by deferring the investment in new transformers. This strategy is especially useful in colder climates, since the peak demand coincides with low temperatures [6].

The application of dynamic ratings requires increased awareness of the transformer operating conditions to safely increase the load above the nameplate rating, without increasing the risk of failures and safety breaches [7]. More specifically, it requires information on the critical temperatures within the transformer, namely winding, inner core and other metallic hot-spot temperatures as well as top oil temperature. For normal cyclic loading, i.e. normal daily operation, the maximum recommended temperatures are 120 °C, 130 °C, 140 °C and 105 °C, respectively [2]. These temperatures can be either monitored by fiber optics or estimated using transformer thermal models.

The International Electrotechnical Commission (IEC) Standard 60076-7 [2] proposes

two different calculation methods for the transformer hot-spot temperature: an exponential equation and a difference equation method. Both methods provide the hot-spot temperature for arbitrarily time-varying load factor and ambient temperature. However, the former method is more suited to be used by manufacturers during tests for the determination of the transformer heat transfer parameters, while the latter method is more suitable for real time monitoring.

In [8], the IEC Std. 60076-7 thermal model was used for dynamic transformer rating for wind energy applications. The model was applied to an existing transformer to assess its reduction in lifespan and reliability when overloaded, and on that basis to provide design and financial considerations. Results showed that by applying dynamic transformer rating, the transformer size can be safely reduced (-20%), thus reducing future investment costs, or that the wind farm can be expanded up to 60%.

In [7], Jalal et al. proposed an extended version of the calculation methods proposed in [2], by developing a dynamic rating algorithm which also includes measurements of the top oil temperature, tap position, and cooling operation in the evaluation of the transformer hot-spot temperature and reduction in lifespan of the insulation.

Similarly, in [9], the differential approach of the IEC Std. 60076-7 was modified to incorporate the dependency of oil viscosity and winding loss on temperature. The proposed improved model was then validated using temperature measurements from a 40 MVA, 21/115 kV, oil forced air forced (OFAF) transformer. Results showed that the improved model outperformed the thermal model proposed by the IEC Std. 60076-7 in estimating the hot-spot temperature for short-time dynamic loading.

Also Annex G of the IEEE C57.91 Standard provides a thermal model that takes in to account the effects on transformer losses of temperature and oil viscosity. However, the requirement of complex and many input parameters is a downside of the methods that rely on the above-mentioned standards, as it complicates the practical implementation. In view of this, simplified models have been proposed as an alternative to the standards' calculation methods.

Arabul and Senol [10] proposed a regression model for hot-spot temperature calculation based on experimental measurements from fiber optic temperature sensors, while aiming to reduce the error rate without increasing input data. Results showed that the proposed method provided more accurate lifetime calculations, by significantly reducing the error in the reduction in lifespan estimation.

A further aspect affecting the transformer hot-spot temperature is the presence of harmonic conditions caused by nonlinear loads [11]. In [12], Das et al. proposed an extension of the two IEC Std. 60076-7 thermal models that account for unbalanced loading with different harmonic distortions in each phase. The model was then incorporated into a dynamic transformer rating algorithm to help utilities to minimize the risk of transformer failure.

A 3-D finite element model was used by Huang et al. [13] to investigate the impacts of harmonics on the magnetic flux leakage and hot-spot temperature rise. Simulation results showed that high-order harmonics can produce a hot-spot temperature rise of around 7 °C. A similar approach was used in [14] for evaluating the transformer losses and estimating the lifetime of oil-filled and dry-type transformers under harmonic loads. Based on numerical

simulation, Zang et al. [15] identified a quantitative relationship for the winding temperature rise under different harmonic content and harmonic frequency. As in [13], their results showed that the hot-spot temperature increases from 0.3 °C to 18.7 °C when the harmonic content increases from 1% to 10%.

While the works above focused on the development of thermal models for real-time monitoring and reduction in lifespan estimation, only a few works in the literature turned their attention on thermal models for dynamic rating applications with predictive capabilities. For dynamic rating applications, predictions are a crucial requirement.

Juarez-Balderas et al. [16] developed a prediction model for forecasting the transformer hot-spot temperature based on Artificial Neural Networks (ANN). The model was validated with both finite element method (FEM) simulations and experimental data, and showed accurate prediction with respect to the latter (average error of 2.71%). However, the model was developed for medium voltage/low voltage (MV/LV) transformers and tested indoors. The model thereby does not consider ambient temperature, while using many inputs and computationally heavy FEM calculations. Therefore, the model has little applicability to many small MV/LV transformers placed outdoors. Moreover, the prediction horizon is not reported in the paper, further limiting the applicability to dynamic operation.

Unlike [16], Bracale et al. [17] proposed a probabilistic stress-strength framework to predict the probability of load not exceeding the transformer rating, and formulated an alarm-setting strategy based on this probability. However, the proposed model only detects the transformer status, i.e. either overload or not, but does not provide any quantitative information regarding the overloaded status of the transformer. Thus, the model is not well suited for real-time monitoring and predictions to be used by a grid operator in a dynamic operation setting. Sun et al. [18] used a support vector regression method to predict the hot spot temperature of a distribution grid transformer. The input data for the method includes the ambient temperature, load rate, historical hot spot temperature and cooling fan status. While the simulations results shows high accuracy of results, its implementation requires hot spot temperature data and can be applied only to the dry-type transformers.

Rommel et al. [19] proposed a method to predict hot spot temperature of transformers when limited information is available. The method uses voltage and current measurement to estimate losses and proposed a simple virtual twin of the transformer to estimate the winding hot spot temperature. The virtual twin is created based on only the transformer nameplate data. The method does not consider the impacts of ambient temperature and environmental data such as solar radiation on the results.

Zang et al. [20] developed a prediction model for transformer winding hot-spot temperature fluctuation based on fuzzy information granulation and the chaotic particle swarm optimized wavelet neural network. The model shows a high prediction accuracy, but the author suggests more research to make it more applicable to engineering practices.

In this context, the transformer winding hot-spot temperature has been rarely studied for time series prediction [20]. In particular, stochastic models for distribution transformer temperature forecasting which take into account prediction of future disturbances have to date rarely been investigated. To fill this gap, the present paper focuses on presenting a methodology to develop a transformer thermal model for temperature prediction by using

stochastic grey-box modeling. The main contribution of this paper is the development of a thermal model for transformers in power distribution grids. The proposed model has the following novelties:

- The proposed model can provide both estimation and predictions. Hence, the model can be used for online forecasting, enabling dynamic rating of distribution grid transformers.
- We use a grey-box modeling approach, meaning that the model accounts for the physics informed and stochastic behavior in the system. This results in the ability to estimate the uncertainty in the predictions and laying the ground for a risk informed dynamic control strategy.
- We developed the proposed model considering input data accessible through non-intrusive measurements. We further chose a model selection process that minimizes the amount of input data (and sensors). Thus, the installation required to apply the model for online forecasting is practical and affordable.
- The proposed model is developed based on experimental data collected during field-trials in the context of the Flexible Energy Denmark (FED) project [21]. This gives the opportunity to investigate and evaluate the transformer temperature predictions in a real world scenario.

The rest of this paper is structured as follows: Section 2 defines the problem and describes the framework within which the thermal model will be applied. Next, Section 3 presents the experimental setup and data acquisition. Section 4 describes the grey-box modeling approach and Section 5 presents and discusses the results. Section 6 concludes the paper.

2. Model application framework

The use of dynamic transformer ratings can help DSOs to unlock extra capacity at distribution grid level and postpone costly grid investments. The dynamic rating should be set relative to the temperature of the transformer, which is the actual limiting factor for the power flow through the transformer. If the temperature is below the thermal limit, a higher power transmission can be allowed. If instead the temperature is above the limit, less power has to be transmitted to avoid transformer failure and aging.

Figure 1 shows the proposed operational framework of a temperature-based dynamic transformer rating (DTR). In this framework, temperature sensors and power measurement devices monitor the transformer operating conditions. An online forecasting tool fetches the data from the transformer sensors as well as the latest environmental data and load forecasts through APIs. In the online forecasting tool the data is then fed to the thermal model, giving predictions of the transformer temperature in a requested time horizon. The DSO can use these predictions along with load forecasts to determine whether more or less power can be transmitted. The adjustments in power flow could be done through activation of flexible

resources. In such a scenario, the DSO could request flexibility from a flexibility operator (for example an aggregator or other actor in a flexibility market). However, determining how this should be done is not within the scope of this work. Through the activation of flexible resources the power flow is changed, and will be visible in the transformer data collected by the online forecasting tool. The electricity demand forecasts are also updated accordingly and the procedure is repeated for each time step.

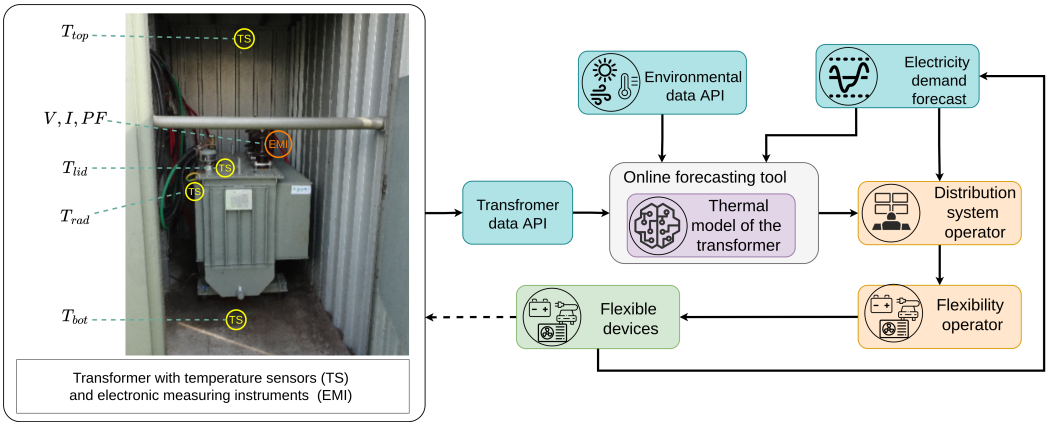


Figure 1: The proposed general framework of data transactions (solid arrow) and changed power flow (dotted arrow) for dynamic rating of transformers considering flexible devices.

Successful implementation of this framework requires accurate prediction of the transformer temperature. The goal of this paper is to develop a thermal model that can be used for both parameter and state estimations as well as predictions of the transformer temperature in the online forecasting tool. According to the transformer manufacturer, monitoring the transformer lid temperature is sufficient for avoiding critical overloading conditions as it relates to the top oil temperature. Hence, the proposed model will be used to predict the lid temperature of the transformers. For this purpose we chose the approach of grey-box modeling. Why this approach is chosen and how it is applied is further explained in Section 4.

To develop a model that can be seamlessly applied in the framework of dynamic rating, we aim to find input variables that can be measured in a non-intrusive way or that can easily be fetched using existing APIs. Nevertheless, the input variables should be chosen such that acceptable prediction results are obtained. This should result in a solution that is affordable and practical, not requiring any interruption in power delivery or replacement of grid equipment, which has clear economic and environmental benefits.

3. Experimental setup and data acquisition

The installation setup includes two 3-phase 10/0.4 kV oil cooled transformers at two separate low voltage (LV) grids owned by a Danish DSO in Jutland. Both transformers are

situated in living labs (LLs) [22] in the Flexible Energy Denmark project [23], meaning they are real world grids that are used for testing new technologies. Transformer 1 (TRF 1) is rated at 400 kVA and serves around 170 residential customers. Transformer 2 (TRF 2) is rated at 200 kVA and serves around 140 residential customers and a small industry. Each transformer supplies 5 - 10 customers with electric vehicles (EVs) and 15 - 20 customers with photovoltaic (PV) panels. Heating occurs through a mixture of heat pumps and district heating. The peak load of both transformers is in the range of 200 to 250 kVA. TRF 2 has a relatively larger base load, compared to TRF 1, which could be due to the industrial load. The transformers are installed in ventilated metal housings that are placed outdoors.

Electrical metering devices (EMDs) are installed in a non-intrusive way (i.e. magnetic mounting and clamp-on current/voltage sensors) on the low-voltage side of the two transformers to collect current, voltage, harmonic and power factor data. Four temperature sensors are installed inside the metal housing as shown in Figure 1. Two sensors measure the housing temperature 10 - 20 cm below the ceiling (T_{top}) and 10 - 20 cm above the floor (T_{bot}), while the other two measure the temperatures of the lid of the transformer case (T_{lid}) and of the transformer radiator (T_{rad}). Table 1 summarizes all the measured data.

Thus, the solution is practically simple and can take place without any interruption in power delivery. A sampling rate of one second was used; however, measurements can be filtered to other resolutions considering the mean value of the per second values. The data from the measuring devices is from a third party company, and we fetch the data using an API.

Variable	**Notation**	**Unit**
Solar radiation	Gh	W/m^2
Wind speed	Φ_{wind}	m/s
Wind speed South	$\Phi_{wind,S}$	m/s
Wind speed North	$\Phi_{wind,N}$	m/s
Wind speed East	$\Phi_{wind,E}$	m/s
Wind speed West	$\Phi_{wind,W}$	m/s
Ambient temperature	T_a	$^\circ C$
Transformer lid temperature	T_{lid}	$^\circ C$
Transformer radiator temperature	T_{rad}	$^\circ C$
Housing top temperature	T_{top}	$^\circ C$
Housing bottom temperature	T_{bot}	$^\circ C$
Apparent three-phase power	S_{3ph}	VA
Phase current	I_{ph}	A
Neutral current	I_N	A

Table 1: Measured input variables.

The ambient temperature, wind speed, wind direction and solar radiation data are retrieved from the open data provided by the Danish Meteorological Institute (DMI open data) [24]. All the data is available at 10 min resolution and can be filtered for other resolutions, e.g. 30 minutes resolution, by considering the appropriate mean value of the 10 min measurements.

The entire data set available is from November 2021 to June 2022. However, if using the entire data set, there would be seasonal effects in the data that the model would need to describe. To properly describe seasonal effects we would need a minimum of two years of

data and associated parameters estimated with the data at hand will be unreliable. As an initial step of the model development we thus choose a shorter time period, namely November 2021, and thereby we can neglect the seasonal effects. A similar approach is seen in e.g. [25]. An example of the time series data for November 1st 2021 to November 8th 2021 is shown in Figure 2. The data has been filtered to 30 min time resolution and this resolution is also used in the model development. This time resolution was chosen as a compromise between smooth data where behavior in a longer time scale is seen (typically systems with high inertia) and a more volatile data set where variation in a shorter time resolution is seen (typically systems with low inertia). It is seen in the figure that the peak in transformer temperature generally occurs after peaks in the other data inputs. For the model development, the data set was divided into a training and a testing data set. The training data set is roughly 80% of the entire data set.

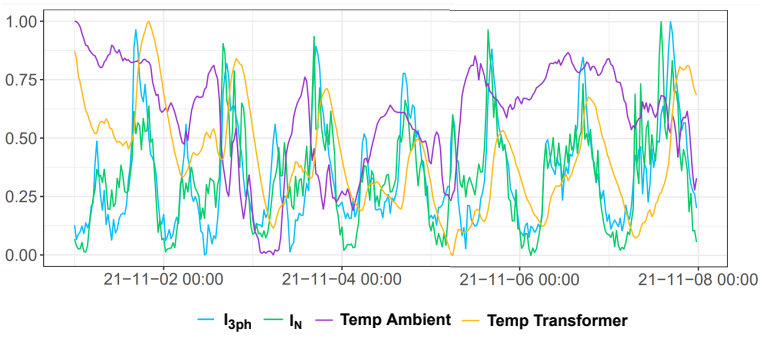


Figure 2: Example of normalized time series data for three-phase current (I_{3ph}), neutral current (I_N) ambient temperature and transformer temperature, measured on the lid. The data has been filtered to a 30 min time resolution.

4. Modeling approach

The present section describes the stochastic grey-box modeling approach used to develop the thermal model. Grey-box models are introduced first, together with the related mathematical framework, followed by a presentation of model structure and the tested models. Statistical methods are used in the model selection.

4.1. Grey-box modeling

Grey-box models have proven to be an effective way to model the dynamics of complex thermal systems [26]. As their name suggests, grey-box approaches are at the intersection between white box approaches, where the model is derived from the theoretical knowledge of the systems (e.g. continuity, momentum and energy equations), and black box approaches,

where statistics, and hence information from data, is used to identify the relationship between the model inputs and outputs, without exploiting any knowledge of its internal process. In grey-box models, the theoretical knowledge of the system is used to suggest a first proposal for the model structure, i.e. a set of first-order stochastic differential equations. Grey-box models are also data-driven in the sense that statistics and input data is used to optimize the parameters of the model. [27] The models are usually written in a continuous-discrete time state-space representation with system equation and observation equation as follows:

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)dt + \mathbf{B}\mathbf{U}(t)dt + \boldsymbol{\sigma}d\mathbf{w}(t) \tag{1}$$

$$\mathbf{Y}(t) = \mathbf{C}\mathbf{X}(t) + \mathbf{e}(t) \tag{2}$$

$\mathbf{X} \in \mathbb{R}^n$ is the state vector, $\mathbf{U} \in \mathbb{R}^p$ is the input vector, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$ and $\mathbf{C} \in \mathbb{R}^{m\times n}$ are the state-space matrices, $\mathbf{Y} \in \mathbb{R}^m$ is the vector of measured outputs, $\mathbf{w}$ standard Wiener processes with incremental covariance matrix $\boldsymbol{\sigma} \in \mathbb{R}^{n\times n}$, and $\mathbf{e} \in \mathbb{R}^m$ are the measurement errors, each assumed to be Gaussian white noise $\mathcal{N}(0, \sigma_{e_k}^2)$ to the kth measured output. We also assume that the measurement errors for the different measurements are uncorrelated. In this work we assume that the Wiener processes are independent, and thus, the diagonal covariance matrix consists of the corresponding variances, σ_i^2, to each ith Wiener process. Finally, we assume that the Wiener processes and the measurement error are independent.

In this work, an iterative model-selection strategy similar to that described by Bacher and Madsen in [26] was adopted to identify the best dynamic model to estimate and predict the transformer temperature. It consists in a forward selection procedure that starts from the simplest model structure, and then iteratively extends the model by adding new states and/or input variables. Model parameters are found by maximizing the joint probability of the observed data given the model structure (see [28] for a detailed discussion). This was done by using the R-package CTSM-R [27], which is a tool for developing stochastic state space models in R. Given the maximum likelihood estimates of the model parameters, each model was then evaluated by analyzing the corresponding residual auto-correlation function (acf) and cumulated periodogram to verify the model assumption of white noise residuals. If the residuals are not white noise this reflects that the model does not describe all the systematic variation in the data, and hence the model has to be expanded. Moreover, the visual inspection of the inputs, outputs, and residuals time series was used to detect what effects the model did not capture, and hence to provide insights for the subsequent model extension. Through the expansion of the model, the significance levels of the estimated parameters were also evaluated, aiming for a p-value lower than 5%. If higher p-values were detected, the model was reduced.

Since the transformers considered in this study are located in two different geographical areas and present different loading conditions, as discussed in Section 3, two different models have been investigated.

4.2. Model structure: transformer heat balance

The model structure was derived from the first law of thermodynamics. By considering the transformer as a closed system that exchanges energy with its surroundings, i.e. external environment, the dynamics of the transformer temperature directly follow from the transformer heat balance:

$$CdT(t) = \Phi_{gain}(t)dt - \Phi_{loss}(t)dt \tag{3}$$

where C and T are the transformer thermal capacity and temperature, respectively, Φ_{gain} is the internal heat gains due to the transformer power losses, and Φ_{loss} is the heat losses towards the surrounding environment, such as those due to convective heat transfer between the transformer case and external air. Power losses are due to the dissipative effects that take place within the transformer, i.e. load and no-load losses, further described in 4.2.1. Heat losses towards the environment account for the heat removed by the transformer cooling system, and the convective and radiative heat transfer between the transformer and its surroundings, namely the environment inside the metal housing. Since the latter is affected by outdoor conditions, the impact of local weather data on the transformer temperature was also taken into account. Figure 3 shows the correlation analysis among the measured variables (Table 1). It can be seen clearly that the transformer temperature is correlated to the external environmental conditions, thus confirming the rationale behind the inclusion of ambient temperature, wind speed and global solar radiation in the modeling process. This also agrees with the IEC Std. 60076-7, stating that environmental factors have a larger impact on smaller transformers, however, the factors are not included in the standard calculation methods [2]. A positive correlation can be noted between the transformer and ambient temperatures: a lower T_a helps to cool down the transformer, while a higher T_a reduces the temperature difference driving the heat transfer, hence the cooling capacity. Conversely, the wind speed and solar radiation input data are inversely correlated to the measurements of T_{lid} (this data behavior and interpretation in terms of the physical system is further discussed in Section 4.2.2).

4.2.1. Transformer power losses

The no-load losses in the transformer are due to the induced voltage in the core. Since the voltage generally has much lower variance than the current, we assumed that the no-load losses are constant in the transformer. Thereby, we simply modeled the no-load losses as a constant, $\Phi_{NL} = b$.

Load losses on the other hand depend on the increased load as:

$$\Phi_{LL} = P + P_{EC} + P_{OSL} \tag{4}$$

where P, P_{EC} and P_{OSL} represent ohmic (I^2R), winding eddy current, and other stray losses respectively [29]. All of these losses increase as the total current squared increases. In the thermal model we used only one parameter to model the effect from the current squared. By this formulation, we assumed that the contribution from all parameters, such as impedance or separate effects in (4), can be summarized to one parameter, a ($\Phi_{LL} \propto aI^2_{3ph}$).

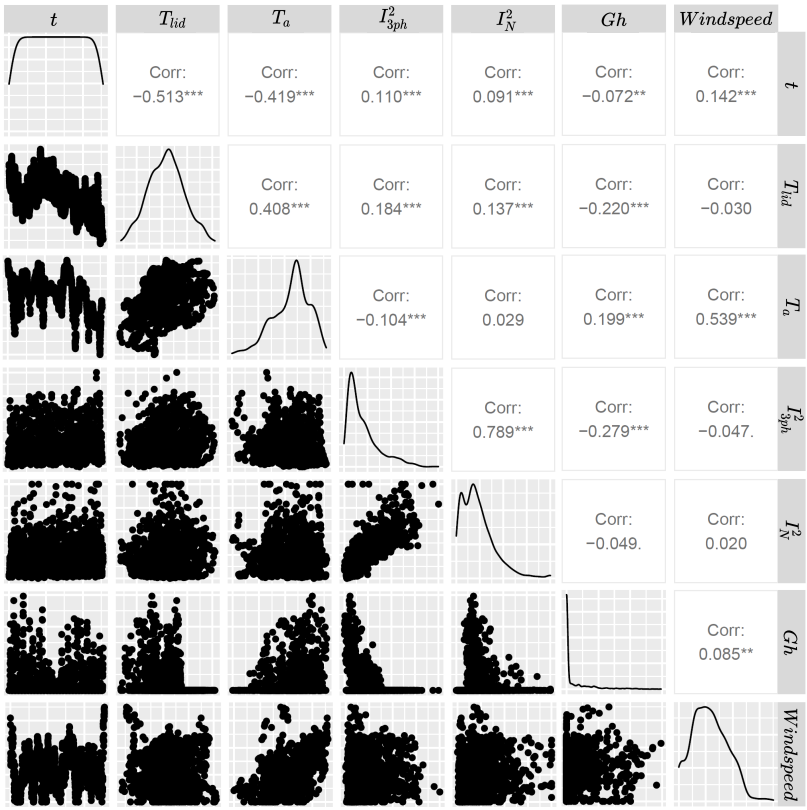


Figure 3: Scatter plots, correlation and data density distribution using data from TRF 1 from November 2021. Here T_{lid} is the transformer lid temperature, T_a is ambient temperature, I^2_{3ph} is three-phase current squared, I^2_N is neutral current squared and Gh is solar radiation.

Moreover, all losses in (4) are also affected by harmonics [29]. Whereas the ohmic loss increases proportionally to increasing harmonic current, eddy-current loss increases proportionally to squared harmonic current and squared frequency. Furthermore, other stray losses increase by a harmonic exponent factor of 0.8 [29]. If taking the approach of modeling the separate effects of harmonic currents, inputs for frequency and each order of the harmonic currents are required. This quickly becomes impractical as data and predictions for harmonic currents are required, while increasing the size and thus the computational time to estimate the model.

To solve this issue, it was suggested to use neutral current as the explaining variable of harmonics effects. Because of the 120 degree symmetry between the phases, it was assumed that harmonic currents pass through the neutral conductor [30]. This behavior was seen and supported in the data, where the 3rd order harmonic at 150 Hz was the second most

apparent frequency component for the three phases, while it was the most apparent in the neutral current. Furthermore, the variance of the frequency is relatively low compared to the variance of harmonic current and thus, we assumed that the effects of frequency can be neglected. Since we used a grey-box modeling approach, we assumed that using the neutral current squared as a linear input is sufficient to describe the losses related to harmonic effects ($\Phi_{LL} \propto cI_N^2$).

In summary, the thermal losses of the transformer can be represented as a function of three-phase current squared and neutral current squared representing the load losses and a constant term representing the no load losses as follows:

$$\Phi_h = \Phi_{LL} + \Phi_{NL} = aI_{3ph}^2 + b + cI_N^2 \tag{5}$$

4.2.2. Environmental factors

The main environmental factors that were taken into account in the thermal model are ambient temperature, wind and solar radiation.

The wind should have a cooling effect on the housing in which the transformer is placed, but could also cause an increase in temperature if it blocks any ventilation in the housing. Thus, we allowed the parameters for wind to be both positive and negative. Both wind speed and wind direction can potentially affect the thermal model. Investigations showed that including wind direction in the model did not give significant parameters while increasing the size of the model. Hence, it was proposed to model the wind using only one input variable, i.e., wind speed.

The solar radiation should have a heating effect on the housing, indirectly resulting in impacts on the heat convection between the temperature in the housing and the temperature of the transformer. This implies less cooling from the surrounding environment and increased transformer temperature. Figure 3 shows no clear relationship and a negative correlation between solar radiation and the transformer temperature (note that a lot of data points are zero). The combination of the physical interpretation of solar radiation impact and the data behavior does not translate well into a linear relationship between the transformer temperature and solar radiation. Therefore, another approach was taken to find a suitable solution to model the solar impact.

The impacts of solar radiation can vary over the day due to the change in solar azimuth angle or shadow effects from, for example, buildings or trees. B-splines offer a flexible fitting of the data to find the function of how one variable affects another. Due to the flexibility, B-splines were used to model the impact of solar radiation on the transformer temperature. The B-splines were applied such that they depend on the time of the day. To reduce the model and avoid unnecessary parameters, the B-splines were designed to be active only during the time intervals for which there is solar radiation, i.e. hours 7 to 16. After applying the B-splines to the model it was discovered that the solar radiation had a significant impact on the temperature of both transformers only between hours 7 to 12. Hence, the interval was reduced to these hours. Using B-splines, the estimated solar radiation impact, Φ_{sol}, can

be formulated as below:

$$\Phi_{sol}(t) = \sum_{j=1}^{n} sc_j B_j(t) Gh(t) \tag{6}$$

where B_j is the jth spline, $\hat{sc_j}$ is the corresponding estimated parameter to the jth spline and Gh is the solar radiation. Since the solar radiation should contribute to an increased transformer temperature we constrained the coefficients, sc_j, to be positive. The best results, i.e., significant parameters and improved estimations of the transformer temperature, for TRF 1 were achieved using a polynomial degree of three and four splines (i.e. n=4 in (6)). It was also discovered that the second spline is insignificant and therefore sc_2 was set to a value close to zero. This simplified the process of finding significant parameters in the model. The behavior could be due to a shadow effect during the time when the second spline is active. For TRF 2 the best results were achieved with five splines (i.e. n=5 in (6)) and again by setting sc_2 to a value close to zero.

4.3. Tested models

Many different models were tested and evaluated in the model development process. As it is infeasible to present all models evaluated within this paper, we will discuss a representative sample from the model selection process. This sample includes a one state model, a two state model, an extended two state model and two three state models, where the latter are the final models for TRF 1 and 2, respectively. The overall model selection process followed the same steps for both TRF 1 and 2. However, there was a slight difference in identifying the final three state model.

4.3.1. One state model

As stated in Section 4.1, we used a forward selection process and thus, we started out with a simple model. The initial model had one state and ambient temperature and current as input variables. This was to represent the simplest model using the most relevant inputs according to the correlation analysis in Section 4.2 (solar radiation was ignored here due to reasons described in Section 4.2.2). The system and observation equations are presented below:

$$dT_i = \frac{1}{C_i}\left(\Phi_h + \frac{1}{R_{ia}}(T_a - T_i)\right)dt + \sigma dw(t) \tag{7}$$

$$T_{lid}(t) = T_i + e(t) \tag{8}$$

where T_i is the corresponding state to the observed lid temperature T_{lid}, C_i is the thermal capacitance at the transformer lid, R_{ia} is the thermal resistance between the lid and the ambient temperature (including the housing) and Φ_h represents the heat generated by the transformer losses. In this model the load and no-load losses were modeled by only using the three-phase current squared as input, i.e. $\Phi_h = aI_{3ph}^2 + b$.

The system can also be described by the circuit model in Figure 4, where the heat generated by load and no-load losses is modeled as a current source and the cooling from ambient temperature as a voltage source.

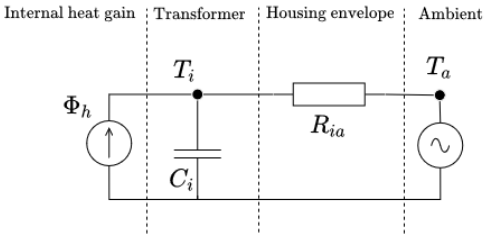


Figure 4: RC circuit of the one state model Ti.

4.3.2. Two state model

We extended the model by adding a second state for the temperature inside the transformer, T_t. Note that this is an arbitrary point inside the transformer and does not aim to identify the hot-spot temperature, but rather model the heat transfer between the inside of the transformer and the lid, i.e. from T_t to T_i. The system equations and observation equation are:

$$dT_i = \frac{1}{C_i}\left(\frac{1}{R_{ti}}(T_t - T_i) + \frac{1}{R_{ia}}(T_a - T_i)\right)dt + \sigma_1 dw_1 \tag{9}$$

$$dT_t = \frac{1}{C_t}\left(\Phi_h + \frac{1}{R_{ti}}(T_i - T_t)\right)dt + \sigma_2 dw_2 \tag{10}$$

$$T_{lid} = T_i + e \tag{11}$$

where T_t is the thermal state within the transformer and generated heat due to power losses is described by the three-phase current squared ($\Phi_h = aI_{3ph}^2 + b$). The Wiener processes are denoted w_1 and w_2 for (9) and (10), respectively. No extra input variables were added in this model compared to the one state model, but parameters for the thermal resistance, R_{ti}, between the states (T_t and T_i) and capacitance, C_t, for the internal state (T_t) were added. This is also visualized in the circuit model in Figure 5.

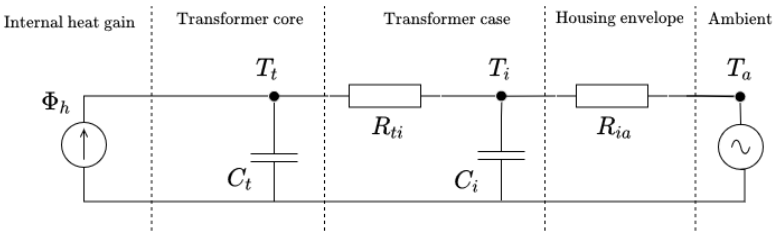


Figure 5: RC circuit of the two state model $TiTt$.

4.3.3. Extended two state model

In this step, we extended the two state model by adding the wind speed as input to the model. This step was done to investigate the impact on the performance of the model that

comes with expanding the model through adding inputs rather than adding states to the model. The resulting system and observation equations are as follows:

$$dT_i = \frac{1}{C_i}\left(\frac{1}{R_{ti}}(T_t - T_i) + \frac{1}{R_{ia}}(T_a - T_i) + \omega\Phi_{wind}\right)dt + \sigma_1 dw_1 \tag{12}$$

$$dT_t = \frac{1}{C_t}\left(\Phi_h + \frac{1}{R_{ti}}(T_i - T_t)\right)dt + \sigma_2 dw_2 \tag{13}$$

$$T_{lid} = T_i + e \tag{14}$$

where Φ_{wind} is the wind speed and ω the corresponding parameter. All other variables and parameters are described in section 4.3.2. The extended two state model is also illustrated in the circuit model in Figure 6, where the wind speed acts as a current source to the transformer temperature state.

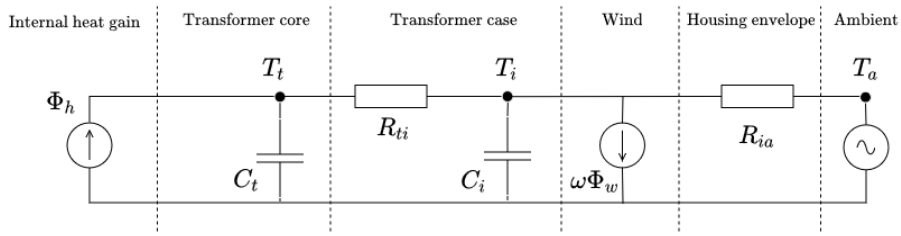


Figure 6: RC circuit of the two state model $TiTt$ with wind contribution.

4.3.4. Three state model

In the three state model, we added a hidden state for the temperature, T_b, representing the temperature inside the metal housing, in which the transformer is placed. It was further explored whether more environmental or electrical inputs should be added to the model. Solar radiation was added to the system equations as an environmental input to increase the performance of the model as described in Section 4.2.2. Due to different locations and properties of the two studied transformers, their three state models are presented separately.

Transformer TRF 1

The three state model for TRF 1 is described by (15) – (18).

$$dT_i = \frac{1}{C_i}\left(\frac{1}{R_{ti}}(T_t - T_i) + \frac{1}{R_{ib}}(T_b - T_i)\right)dt + \sigma_1 dw_1 \tag{15}$$

$$dT_t = \frac{1}{C_t}\left(\Phi_h + \frac{1}{R_{ti}}(T_i - T_t)\right)dt + \sigma_2 dw_2 \tag{16}$$

$$dT_b = \frac{1}{C_b}\left(\frac{1}{R_{ib}}(T_i - T_b) + \frac{1}{R_{ba}}(T_a - T_b) + \omega\Phi_{wind} + \Phi_{sol}\right)dt + \sigma_3 dw_3 \tag{17}$$

$$T_{lid} = T_i + e \tag{18}$$

where R_{ib} is the thermal resistance between the state at the lid, T_i and the state in the housing, T_b. C_b is the thermal capacitance in the housing, Φ_{sol}, is explained in (6) and Φ_h is described by the three-phase current squared ($\Phi_h = aI_{3ph}^2 + b$). All other variables and parameters are explained in Sections 4.3.2 and 4.3.3. The thermal model for TRF 1 is also depicted in the thermal circuit model (Figure 7).

Transformer TRF 2

For the final three state model, the neutral current turned out to be a significant input to TRF 2, but not to TRF 1. Thus, for TRF 2, the neutral current was added to Φ_h in both the system equation (16) and in the RC circuit in Figure 7 (i.e. $\Phi_h = aI_{3ph}^2 + b + cI_N^2$).

It should also be noted that the input for solar radiation differs for TRF 1 and 2 as described in Section 4.2.2.

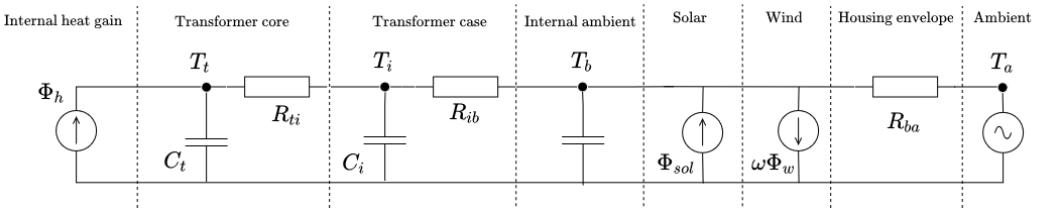


Figure 7: RC circuit of the three state model *TiTtTb*.

4.4. Temperature estimation and prediction

As described in Section 4.2.1, the power losses are directly causing heating inside the transformer. Meanwhile, the measuring point is at the lid of the transformer (see Figure 1) and it is realized that there will be a time delay between the time of the hot-spot temperature inside the transformer and corresponding effects seen at the observed transformer lid temperature. This can also be realized by observing the thermal resistance in the oil as well as thermal capacitance in the RC circuits in e.g. Figure 7.

This time delay is important for the understanding of which prediction horizon of the transformer lid temperature gives an indication of the core conditions at the current time step. To investigate this, a cross-correlation analysis between three-phase current squared and transformer lid temperature was carried out. Assuming that the leakage flux is heating the metallic parts quite instantly, the time span between an increase in current squared and the temperature rise at the lid gives an indication of the time delay. The three-phase current squared was chosen as the main explaining variable for the transformer inner temperature, following the description of transformer losses above and given that it has a higher correlation than the neutral current as seen in Figure 3. The analysis was done by dividing the time series for each transformer into segments of days. The cross-correlation between the three-phase current squared and transformer lid temperature was calculated, while noting the lag

at which the maximum value of the cross correlation was achieved. For TRF 1 the maximum cross correlation occurred on average at a lag of 5 ± 1.5 time steps with a 95% confidence interval. For TRF 2 it instead occurred at 3.5 ± 1 time steps with a 95% confidence interval. Keeping in mind that each time step is 30 minutes, this means that the heat generated from load losses takes 2.5 ± 0.75 hours to transfer to the measuring point at the lid for TRF 1 and for TRF 2 it takes 1.75 ± 0.5 hours. The difference in time delay between the two transformers is most likely due to the smaller size and thus thermal inertia of TRF 2 (200 kVA), which is half the size of TRF 1 (400 kVA).

As seen in the model application framework in Section 2 (Figure 1) the DSO should receive predictions such that adjustments can made to the load of the transformer before the limit is violated. Given that predictions at 5 ± 1.5 and 3.5 ± 1 time steps ahead, relate to the current state inside the transformer, predictions further ahead are required to support the DSO in the model application framework. For this purpose we evaluate the thermal model at 12 step ahead predictions (6 hours), leaving 3.5 and 4 hours respectively to activate flexible resources.

5. Results and Discussion

In this section we present and analyze the results from the models defined in Section 4.3. To avoid repetition, the results are mostly discussed for TRF 1, and the models of TRF 2 are presented for the sake of comparison. We further discuss the results in the context of the model application framework presented in Section 2.

5.1. Residual analysis

Based on the definition of the model in (2), the residuals from an adequate model should be in the form of normally distributed white noise. The residual analysis is done by evaluating the auto correlation function (acf), to ensure that the residuals are independent, and also by looking at the cumulated periodogram, to ensure that no frequencies are left in the residuals. Both evaluations aims to identify whether the residuals are white noise or not. If the residuals are not white noise, there are patterns in the system that the model is not capturing. The principles of using the residuals for a model evaluation are based on the methods suggested in [31].

The acf and cumulated periodograms from the models presented in Section 4.3 are shown in Figure 8. It can be seen that the significant and periodic values in the acf are gradually decreasing throughout the model selection process. However, little improvement is seen between the one and two state models, whereas reduced acf and a shift in the cumulated periodogram can be seen for the extended two state model. This means that extending the model by adding input variables to the model was necessary. It should be noted that adding states such that the input variables could be described in a physics informed manner was also required and improvements for the residual analysis were observed through such extensions. For the presented one and two state models the acf has significant periodic values indicating that there are patterns or behaviors in the system that the models do not capture. It is further seen in the cumulated periodograms that some frequencies of the

residuals are dominating for these models. Thus, it cannot be concluded that the residuals are white noise and neither the one nor the two state models properly describe the system.

Looking at the residual analysis for the three state model in (15) – (18) for TRF 1, it can be seen that the acf (Figure 8d) does not have any significant values. Furthermore, the cumulated periodogram (Figure 8h) shows no dominating frequencies in the residuals, but lies within the interval of the 95% confidence interval. Thereby, the three state model gives residuals that are white noise and we conclude that the three state model provides an adequate description of the system.

The acf and cumulated peridogram for the three state model for TRF 2 show white noise residuals within a 95% confidence interval (Figure 9). Thereby, this three state model with the modification described in Section 4.3 is identified as the final model for TRF 2.

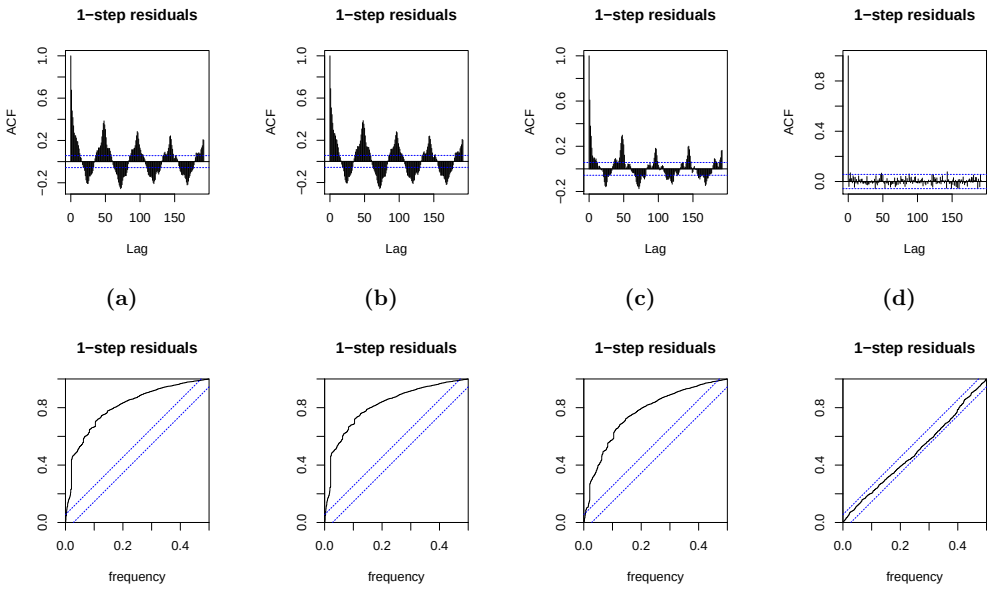
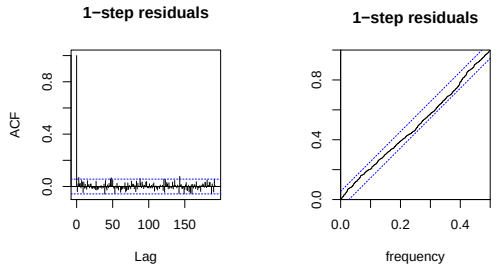


Figure 8: Residual analysis for the model selection process for transformer 1. Graphs for acf are shown in subfigures (a) – (d) and cumulated periodogram in (e) – (h). The results for the different models are presented as follows: One state model in (a) and (e), two state model in (b) and (f), extended two state model in (c) and (g), three state model in (d) and (h). Blue dotted lines indicate 95% confidence bands under the assumption that the residuals are white noise.

5.2. Residual sum of squares and likelihood analysis

We analyzed the residual sum of squares (RSS) values and mean average error (MAE) to evaluate whether the errors in the output given the observed data were reduced in the



(a) **(b)**

Figure 9: Residual analysis for the three state model for TRF 2. The acf is shown in subfigure (a) and cumulated periodogram in (b). Blue dotted lines indicate a 95% significance level.

model selection process. The values of the log likelihood function were used to compare the models. As shown in Table 2, by moving from one state models to three state models the values of RSS, MAE and likelihood values of both transformers are improved, which confirms an increased performance of the model throughout the selection process. Comparing the results for both transformers shows that the proposed three state model for TRF 1 has a better performance than the three state model for TRF 2. Furthermore, in the final models a maximum absolute error of 1.69 and 2.47 °C was seen for TRF 1 and TRF 2, respectively, using the training data. If instead we use the test data, the maximum absolute errors were 5.6 and 3.0 °C, respectively. For TRF 2 this is acceptable for the application of the dynamic rating framework, which is satisfying given that this transformer is the most critically loaded, with a peak above the nameplate rating. For TRF 1, the results could be improved, but it should be mentioned that the MAEs are 0.73 and 0.87 °C, respectively, and the high maximum error for TRF 1 is an exception.

Model	**RSS 1 step**	**RSS 12 step**	**MAE**	**Max Error**	**Log likelihood**	**Computation time**
Trf 1: One state	21.6	1035	0.79	3.82	772	11 secs
Trf 1: Two state	21.4	1061	0.81	3.81	718	1.34 mins
Trf 1: Extended two state	16.2	468	0.54	2.54	885	1.84 mins
Trf 1: Three state	8.5	261	0.39	1.69	1273	5.95 mins
Trf 2: One state	50.4	1179	1.06	5.08	-37.6	9 secs
Trf 2: Two state	50.0	1184	1.07	5.05	-34.4	1.34 mins
Trf 2: Extended two state	42.0	654	0.77	4.40	38.9	2.16 mins
Trf 2: Three state	27.6	442	0.63	2.47	212.8	6.37 mins

Table 2: Summary of RSS for 1 and 12 step ahead, mean average error and maximum absolute error for 12 step ahead predictions, log likelihood from models for TRF 1 and TRF 2 using the training data set. The computation time of estimating the parameters, using a Intel core i7 @ 1.90 Ghz, 16 GB RAM and running on Linux Pop! OS version 21.10, is also presented.

5.3. Estimated parameters

The estimated parameters for the final model are seen in Table 3 for TRF 1 and 2. Normalized inputs were used for all variables, except for the temperature variables, which were kept in measured °C. If including the squared neutral current, I_N^2, for TRF 1, the estimated parameter, c, obtained a p-value higher than 0.1, meaning more than a 10% probability that it should not be included in the model. Thus, there is a slight difference in the final model for the two transformers, meaning that the neutral current is a significant input variable to TRF 2, whereas it was not for TRF 1. Noting that the percentage of neutral current to total phase current is quite similar for both transformers (13% for TRF 1 and 12% for TRF 2), it is concluded that a higher harmonic content in TRF 2 is probably not the reason for this difference. Instead another possible reason is that TRF 2 has a loading that is almost twice as large as the loading condition for TRF 1 (see 3). This could mean that the power related factors have a higher impact compared to environmental factors for TRF 2, and hence, the neutral current is significant during such a loading condition. If this holds, this means that when implementing the model at and above nameplate rating the neutral current should be included as an input variable to the model.

It was also seen that for the parameter for the Wiener process, σ_2 in the third system equation (17) was significant for TRF 2, but not for TRF 1. Thus, in the final model for TRF 1, there is no Wiener process (i.e. $\sigma_2 dw_2$ is not included in (17)).

Moreover, different numbers of splines were needed for the two transformers to model the impact from the solar radiation. This is reasonable given that different shading and angle to direct solar radiation could occur depending on the placement of the transformers. Identification of the splines could potentially to some extent be automized if the model is applied at large scale. It was, however, seen for both transformers that the solar radiation only had an impact during the morning until noon. Seasonal effects could also be incorporated in the future, as the described interval, solar spline function and effects from the wind would probably change throughout the year. However, to investigate seasonal effects, time series of minimum two years are required (as discussed in Section 3). As the model was developed using the available data from November 2021, it was not relevant for this study, but is rather a part of future work.

Although four different temperature measurements were available (see Figure 1), the best results were achieved using only one of them (T_{lid}). This further reduces the number of sensors required in the installation in order to apply the model.

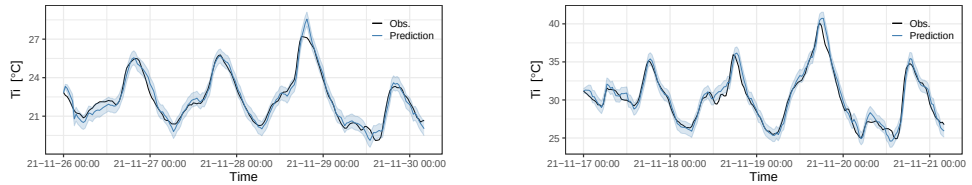
5.4. Application of the proposed method for real time estimation and prediction

As described in Section 4.4, predictions at 5 ± 1.5 and 3.5 ± 1 time steps ahead relate to the current state inside the transformer. The time delay is due to the time constants of the system and the predictions at these horizons are useful for control room purposes. Therefore, we evaluated the predictions at horizons 5 and 4 (rounded off from 3.5) time steps ahead, or 2.5 and 2 hours, respectively. The predictions and PIs can be seen in Figure 10. The predictions follow the observations for most of the time. Nevertheless, there is an overshoot in the estimation of the evening peak on the 28th of November. This could be due to "the memory" (the derivative of previous time steps) in the model structure and the previous

Parameter	TRF 1 Estimate	Std. Dev.	TRF 2 Estimate	Std. Dev.
Initial state T_i	28.5	0.07	28.2	0.15
Initial state T_b	22.7	1.2	20.2	1.9
Initial state T_t	52.0	14.2	41.5	5.1
a	2.61	0.30	3.33	0.26
b	0.607	0.111	0.863	0.113
c	−−	−−	0.158	0.055
C_b	4851	876	2312	403
C_e	412	69	168	40
C_t	413	165	394	79
$ln(\sigma_e^2)$	−45.7	0.09	−45.7	0.07
$ln(\sigma_1)$	−25.1	5.3	−20.7	11.3
$ln(\sigma_2)$	−−	−−	−48.3	26.8
$ln(\sigma_3)$	−5.46	0.12	−4.79	0.14
R_b	3.97	0.79	5.28	0.97
R_{ia}	14.2	1.76	9.76	0.89
R_{ti}	15.8	7.5	8.85	3.08
sc_1	3.67	0.78	14.8	3.64
sc_3	1.86	0.41	1.10	0.60
sc_4	1.03	0.24	1.68	0.58
sc_5	−−	−−	1.41	0.37
ω	−0.817	0.124	−0.902	0.134

Table 3: Parameter estimations and standard deviations (Std. Dev.) in the final models for TRF 1 and 2. All inputs were normalized, except for the initial temperatures, which are presented in °C.

positive trend in the temperature time series. Although the prediction intervals do not fully capture the observations, they give reasonable predictions of the states. Taking in to account that there is a 95% PI in the graph, under and over estimations will occur from time to time. Through adjusting the PIs, the likelihood of having an observation outside the PIs can be as small as desired. Thereby, the DSO can account for the uncertainty in their control strategies by adjusting the PIs. Predictions further ahead than the current state are



(a) 5 step ahead predictions for the test data set, using the one state model for TRF 1.

(b) 4 step ahead predictions for the test data set, using the extended two state model for TRF 2.

Figure 10: Predictions of the transformer lid temperature for 5 and 4 time steps ahead, respectively, corresponding to the current state inside the transformer. Black line – observations, Blue line – predictions, Light blue area – 95% PI.

required to support the DSO in the model application framework in Figure 1. We therefore evaluated 12 step ahead predictions (6 hours), giving 3.5 and 4 hours respectively to adjust the power flow through activating flexible resources.

The 12 step ahead predictions for the test data set are presented in Figure 11. Improvements in the accuracy of the prediction and the prediction intervals (PIs) can be seen as the model is expanded from one to three states. For the one and two state models the PI often misses the observation, which is a sign of a not well enough implemented model. For the final three state model for both transformers, the observations are, however, most of the time inside the PIs. Naturally, some observations will be outside of the PI due to the 95% significance level of the PIs. Most importantly for the safety of the grid operation, the PIs capture the peaks of the temperature, which are the most critical points for the implementation. Thus, we conclude that the predictions at this time horizon are fulfilling the requirements for usage in the model application framework.

Nevertheless, there is also potential for improvements to reduce the uncertainty of the model. The model is developed for, and using data from, normal cyclic operation. However, if the model should be applied under other operating conditions, such as long time emergency loading [2], a reduced uncertainty might be required. For instance, heat run tests could be performed to establish a model, table or such, that can estimate the hot-spot temperature given the transformer lid temperature and loading conditions. This could, for example, be done with optical sensors. With the proposed model, the DSO would need to establish some extra temperature margin that is acceptable from a safety point of view given the uncertainty. Performing such tests would reduce this margin as DSOs would have more knowledge on how to translate the transformer temperature and loading condition to possible scenarios for the hot-spot temperature.

Furthermore, other improvements, such as adding more states to the model, could be investigated. This could, for example, include seasonal effects from solar radiation or having adaptive parameters. It should, however, be noted that such improvements in the accuracy could have a negative impact on the computation time. As seen in Table 2 the computational time increased when increasing the model order.

5.5. Discussion on applicability of the model

To evaluate the applicability of the thermal model, the computation time is a crucial factor. It should be noted that although capturing non-linear behavior, such as heating from the solar radiation, the grey-box models are formulated as linear models. This reduces the computational burden of optimizing the parameters in the models. As seen in Table 2, the computation time increases with increasing order of the model. It is also seen that the time to run and optimize the parameters for the final model is approximately six minutes for TRF 1 and 2, respectively. With this computation time, the model can be updated on a regular basis, enabling the usage of the model in grid operation. How often the parameters of the model should be updated is a compromise between model accuracy and computational burden. It is possible to update the parameters for every time step given that the time resolution is 30 minutes, but it needs to be evaluated if that is feasible. The DSO has to weight the gained capacity from having an accurate thermal model against the economical cost of updating the models, taking into account how many transformers in the system the model will be applied to.

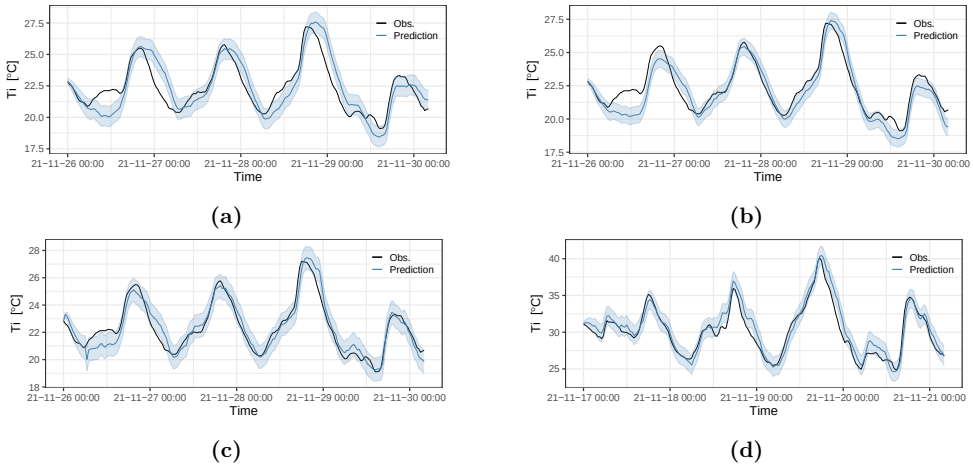


Figure 11: Prediction analysis for 12 step ahead (6 hours) predictions. Subfigures (a)-(c) show predictions for TRF 1 using the one state model (a), extended two state model (b) and the final three state model (c). Subfigure (d) shows predictions for TRF 2 using the final three state model. Black line – observations, Blue line – predictions, Light blue area – 95% PI.

Given the analysis of the predictions, accuracy and the computational time, the model can be applied for online forecasting of the transformer temperature in the model application framework. The manufacturer of the transformer in this study also claims that keeping track of the transformer lid temperature should be enough to safely operate the transformer. However, for full implementation of the model application framework, an algorithm for the online forecasting tool needs to be developed.

In this work we have also found a solution that is both affordable and practical in terms of hardware installation. This is due to the limited number of input variables required to run the models. Weather data can be fetched from the meteorological forecast provider (in our case DMI) and current sensors have become more affordable in recent years. Furthermore, the model automatically includes environmental and electrical characteristics specific to the transformer, potentially unlocking even more capacity than a deterministic solution based on standard conditions. Since we have a stochastic model, we can also provide PIs accounting for the uncertainty.

The model could potentially be extended to larger MV/LV transformers, if data for model calibration is available. As the leakage flux increases with increasing size, electrical input variables will most likely have a larger comparative impact and the model structure for TRF 2 would be recommended for initial model calibration to be reduced if necessary.

6. Conclusion

In this paper, we proposed a framework for obtaining dynamic rating of transformers in power distribution grids. For this purpose, we developed a thermal model for estimating

and predicting transformer temperature. The model was developed using data from a real world experimental installation and can be used for developing an online monitoring and forecasting algorithm for DSOs.

The proposed model is formulated as a grey-box model. This is a physics-informed data-driven model which is optimized for assimilating information from sensors into the model parameters. Furthermore, this approach gives a possibility for providing prediction intervals, and hereby we can specify the risk of violating the temperatures.

The identified input variables in the proposed thermal models result in an affordable and practical solution. The hardware in the solution can be installed without interruption of power supply and without exchanging any equipment.

Furthermore, the proposed thermal model has been proven to give reasonable estimations and predictions. This enables the ability to operate the transformer dynamically, unlocking unused capacity at certain times.

The proposed model also accounts for the electrical and environmental conditions at a specific transformer. This could give a more specifically applied DTR, unlocking more capacity than a DTR or thermal model based on standard conditions.

Moreover, the computational time for the model gives the option to update the parameters of the model on a regular basis. This further enables implementation in grid operation as the accuracy can be maintained.

6.1. Future work

Although the proposed model can be applied for online monitoring and forecasting as it is, it could be complemented or improved to reduce the uncertainty.

As discussed in Section 5, heat run tests with optical sensors could be made to establish how the transformer lid temperature and loading conditions affect the hot-spot temperature. This could provide a better insight into the transformer hot-spot temperature.

Expansions to the model, such as states for seasonal effects, adaptive parameters or improved local weather forecasts could be developed. However, caution needs to be taken to not increase the computation time and burden too much or alternatively find a workaround that does not require us to rerun the model to update the parameters. We are planning on collecting data continuously during the coming years, and within a few years we aim at formulating an extension of the models which includes seasonal effects.

Finally, an algorithm that uses the proposed model and input data to deliver the forecasts online in real time needs to be developed for full implementation of the concept.

7. Acknowledgments

This work was supported by the Flexible Energy Denmark (FED) project funded by Innovation Fund Denmark under Grant No. 8090-00069B and Ebalanceplus project funded by the European Union's Horizon 2020 under the grant agreement number of 864283. The authors also would like to thank Claus Schack Urup for technical support and providing the data for this study.

References

[1] A. A. Taheri, A. Abdali, M. Taghilou, H. Haes Alhelou, K. Mazlumi, Investigation of Mineral Oil-Based Nanofluids Effect on Oil Temperature Reduction and Loading Capacity Increment of Distribution Transformers, Energy Reports 7 (2021) 4325–4334. doi:10.1016/j.egyr.2021.07.018.

[2] IEC, DS/IEC 60076-7:2018, Power transformers – Part 7 : Loading guide for mineral-oil-immersed power transformers (2018).

[3] IEEE Standards Board, IEEE Standard Test Code for Liquid-Immersed Distribution, Power and Regulating Transformers Std C57.12.90, IEEE, 1999.

[4] M. F. Lachman, P. J. Griffin, W. Walter, A. Wilson, Real-time dynamic loading and thermal diagnostic of power transformers, IEEE Transactions on Power Delivery 18 (1) (2003) 142–148. doi:10.1109/TPWRD.2002.803724.

[5] K. Morozovska, P. Hilber, Study of the Monitoring Systems for Dynamic Line Rating, Energy Procedia 105 (2017) 2557–2562. doi:10.1016/j.egypro.2017.03.735.
URL http://dx.doi.org/10.1016/j.egypro.2017.03.735

[6] C. J. Wallnerström, P. Hilber, P. Söderström, R. Saers, O. Hansson, Potential of dynamic rating in Sweden, in: 2014 International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2014 - Conference Proceedings, IEEE, 2014. doi:10.1109/PMAPS.2014.6960605.

[7] T. S. Jalal, N. Rashid, B. Van Vliet, Implementation of Dynamic Transformer Rating in a distribution network, in: 2012 IEEE International Conference on Power System Technology, POWERCON 2012, IEEE, 2012, pp. 2–6. doi:10.1109/PowerCon.2012.6401328.

[8] T. Zarei, K. Morozovska, T. Laneryd, P. Hilber, M. Wihlén, O. Hansson, Reliability considerations and economic benefits of dynamic transformer rating for wind energy integration, International Journal of Electrical Power and Energy Systems 106 (2018) 598–606. doi:10.1016/j.ijepes.2018.09.038.
URL https://doi.org/10.1016/j.ijepes.2018.09.038

[9] F. Josue, I. Arifianto, R. Saers, J. Rosenlind, P. Hilber, Suwarno, Transformer hot-spot temperature estimation for short-time dynamic loading, in: Proceedings of 2012 IEEE International Conference on Condition Monitoring and Diagnosis, CMD 2012, IEEE, 2012, pp. 217–220. doi:10.1109/CMD.2012.6416414.

[10] A. Arabul, I. Senol, Development of a hot-spot temperature calculation method for the loss of life estimation of an ONAN distribution transformer, Electrical Engineering 100 (2018). doi:10.1007/s00202-017-0641-0.

[11] S. Taheri, H. Taheri, I. Fofana, H. Hemmatjou, A. Gholami, Effect of power system harmonics on transformer loading capability and hot spot temperature, in: 2012 25th IEEE Canadian Conference on Electrical and Computer Engineering: Vision for a Greener Future, CCECE 2012, IEEE, 2012. doi:10.1109/CCECE.2012.6334834.

[12] B. Das, T. S. Jalal, F. J. McFadden, Comparison and extension of IEC thermal models for dynamic rating of distribution transformers, in: 2016 IEEE International Conference on Power System Technology, POWERCON 2016, IEEE, 2016, pp. 1–8. doi:10.1109/POWERCON.2016.7753896.

[13] Z. Huang, Y. Zhang, Y. Jing, M. Fu, G. Wang, R. Zhuo, Q. Wang, Study on the Influence of Harmonics on the Magnetic Leakage Field and Temperature Field of 500 kV Connected Transformer, in: ICEMPE 2019 - 2nd International Conference on Electrical Materials and Power Equipment, Proceedings, IEEE, 2019, pp. 596–600. doi:10.1109/ICEMPE.2019.8727246.

[14] I. Kapetanović, J. Hivziefendić, M. Tešanović, Different Approaches for Analysis of Harmonics Impact on the Transformer Losses and Life Expectancy, Lecture Notes in Networks and Systems 28 (2018) 392–408. doi:10.1007/978-3-319-71321-2{_}36.

[15] J. Zhang, L. Cheng, H. Wen, C. Liu, J. Hao, Z. Li, Simulation Analysis of the Influence of Harmonics Current on the Winding Temperature Distribution of Converter Transformer, in: Proceedings - 2021 6th Asia Conference on Power and Electrical Engineering, ACPEE 2021, 2021, pp. 1566–1571. doi:10.1109/ACPEE51499.2021.9436944.

[16] E. A. Juarez-Balderas, J. Medina-Marin, J. C. Olivares-Galvan, N. Hernandez-Romero, J. C. Seck-Tuoh-

Mora, A. Rodriguez-Aguilar, Hot-spot temperature forecasting of the instrument transformer using an artificial neural network, IEEE Access 8 (2020) 164392–164406. doi:10.1109/ACCESS.2020.3021673.
[17] A. Bracale, P. Caramia, G. Carpinelli, P. De Falco, SmarTrafo: A Probabilistic Predictive Tool for Dynamic Transformer Rating, IEEE Transactions on Power Delivery 36 (3) (2021) 1619–1630. doi:10.1109/TPWRD.2020.3012180.
[18] Y. Sun, G. Xu, N. Li, K. Li, Y. Liang, H. Zhong, L. Zhang, P. Liu, Hotspot temperature prediction of dry-type transformers based on particle filter optimization with support vector regression, Symmetry 13 (8) (2021). doi:10.3390/sym13081320.
[19] D. P. Rommel, D. Di Maio, T. Tinga, Transformer hot spot temperature prediction based on basic operator information, International Journal of Electrical Power & Energy Systems 124 (2021) 106340. doi:https://doi.org/10.1016/j.ijepes.2020.106340.
URL https://www.sciencedirect.com/science/article/pii/S014206152030867X
[20] L. Zhang, W. Zhang, J. Liu, T. Zhao, L. Zou, X. Wang, A new prediction model for transformer winding hotspot temperature fluctuation based on fuzzy information granulation and an optimized wavelet neural network, Energies 10 (12) (2017). doi:10.3390/en10121998.
URL https://www.mdpi.com/1996-1073/10/12/1998
[21] Flexible Energy Denmark.
URL https://www.flexibleenergydenmark.dk/
[22] Uni-lab.dk, Living labs, visited on 2022-03-03.
URL https://www.uni-lab.dk/en/living-labs/
[23] Flexible Energy Denmark, visited on 2022-03-03.
URL https://www.flexibleenergydenmark.dk/
[24] DMI, Danish Meteorological Institute - Open Data, visited on 2022-03-03.
URL https://confluence.govcloud.dk/display/FDAPI
[25] P. Nystrup, H. Madsen, E. M. Blomgren, G. de Zotti, Clustering commercial and industrial load patterns for long-term energy planning, Smart Energy 2 (2021) 100010. doi:10.1016/j.segy.2021.100010.
URL https://doi.org/10.1016/j.segy.2021.100010
[26] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy and Buildings 43 (7) (2011) 1511–1522. doi:10.1016/j.enbuild.2011.02.005.
URL http://dx.doi.org/10.1016/j.enbuild.2011.02.005
[27] R. Juhl, J. K. Møller, H. Madsen, ctsmr - Continuous Time Stochastic Modeling in R, arXiv (2016).
URL http://arxiv.org/abs/1606.00242
[28] N. R. Kristensen, H. Madsen, S. B. Jørgensen, Parameter estimation in stochastic grey-box models, Automatica 40 (2) (2004) 225–237. doi:10.1016/j.automatica.2003.10.001.
[29] IEEE, IEEE STD C57.110-1998 Recommended Practice for Establishing Transformer Capability When Supplying Nonsinusoidal Load Currents, Standard (1998).
URL http://ieeexplore.ieee.org/iel5/59/10801/x0247870.pdf
[30] J. B. Noshahr, M. Bagheri, M. Kermani, The Estimation of the Influence of Each Harmonic Component in Load Unbalance of Distribution Transformers in Harmonic Loading Condition, Proceedings - 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2019 (2019). doi:10.1109/EEEIC.2019.8783488.
[31] H. Madsen, Time Series Analysis, Chapman and Hall/CRC, 2007. doi:doi-org.proxy.findit.cvt.dk/10.1201/9781420059687.

4 Paper C

Emma M.V. Blomgren, Mohsen Banaei, Razgar Ebrahimy, Olof Samuelsson, Francesco D'Ettorre, Henrik Madsen,
"Intensive data-driven model for real-time observability in low voltage radial DSO grids"
(submitted to) *Applied Energy*

Intensive data-driven model for real-time observability in low voltage radial DSO grids

Emma M.V. Blomgren[a], Mohsen Banaei[a,*], Razgar Ebrahimy[a], Olof Samuelsson[b], Francesco D'Ettorre[a], Henrik Madsen[a]

[a] *Technical University of Denmark, Department of Applied Mathematics and Computer Science, Lyngby, Denmark*

[b] *Lund University, Faculty of Engineering, Industrial Electrical Engineering and Automation, Lund, Sweden*

Abstract

Increasing levels of distributed generation (DG) and changes in electricity consumption behaviour, are reshaping the power distribution systems. These changes might especially stress the secondary low voltage (LV) distribution systems, that were not originally designed for these power flows. Voltage violations, reverse power flow and congestion constitute some of the concerns for distribution system operators (DSOs), while observability in these grids are typically nonexistent or very low. The present paper addresses this issue by developing a method for node voltage estimations for unbalanced radial LV grids (at 0.4 kV). The proposed method combines a grey-box modeling approach with generalized additive models (GAMs) achieving results with high accuracy and root mean squared errors of 0.002 – 0.0004 p.u.. Furthermore, the developed method relies on data from a real-world LV grid in Denmark, and manages to to use input parameters from only one measuring device per feeder. The method also results in a low computation time achieving applicability to online monitoring in DSO grids.

Keywords: Data-driven modeling, Distribution power systems, Grey-box modeling, Generalized additive models, Phase voltage estimation

1. Introduction

As a necessary means towards carbon neutral energy systems, power systems operation is undergoing a paradigm shift. More devices are becoming electrified, such as electric vehicles (EVs), levels of distributed generation (DG) are increasing, and demand side flexibility is attracting an increasing attention in providing flexibility services for the power system [1]. These developments are leading to changes in consumption and production patterns, which might stress the low voltage (LV) distribution power systems that were originally not designed for these conditions. Meanwhile, operational observability in LV grids is in general

*Corresponding author

Email address: moban@dtu.dk (Mohsen Banaei)

nonexistent or very low. As a result, to ensure reliable operation of grids, distribution system operators (DSOs) require techniques to improve grid observability.

Although smart meters are being installed in large scale in Europe, they offer limited potential for improving observability of distribution grids. In [2], the authors state that it is not the smart meters that are the largest cost, but rather the required communication infrastructure. Moreover, smart meters communication systems could be subject of cyber attacks, are often delayed (e.g., smart meters in household level that collect data only once a day) and have slower sampling frequency compared to phase measurement units (PMUs). Hence, to infer the values of system's state variables using a limited number of data requires distribution system state estimation (DSSE).

While state estimation is common practice in transmission grids there are some factors that complicates the application of the same state estimation methods in distribution grids such as low X/R ratios, unbalanced operation, and fast changes in the configuration of distribution grids [3]. Thus, new methods are proposed in the literature for DSSE that can be categorized from different viewpoints. In general, DSSE problems are voltage-based or branch-current-based. The main focus of this paper is on the voltage-based methods, however several branch-current-based studies can be found in the literature such as [4], [5], and [6].

Among voltage-based studies, many research studies try to apply or modify the weighted least square (WLS) approach, commonly used in transmission system state estimation [7], for DSSE. Lin et. al [8] proposed a fast decoupled DSSE method taking into account the virtual measurements, i.e., perfect information about the grid as equality constraint in the formulation. To this end, a penalty factor was added to the standard WLS problem that enforces satisfying the equality constraint. This method needs no assumptions on voltage magnitudes and phase angles. Chen et. al [9], proposed a methodology for DSSE in case that only the aggregate data of smart meters are available due to the respect to customers' privacy. The variance of measurement errors of the smart meters were used to construct the weight matrix in the WLS optimization problem. Power flow analysis was performed to create time series of active and reactive power data for the study. The problem of DSSE for areas with high penetration of electric vehicles (EVs) was addressed by Nie et. al [10]. To provide more reliable and accurate results, a new quasi-Newton method was used to solve the WLS problem. Effectiveness of the method was evaluated by applying it to a IEEE 14-bus and 30-bus test systems using real travel survey statistics and base load records. Simulation results showed better performance of the proposed method compared to standard WLS and extended Kalman filter methods, especially when the number of EVs increases. The DSSE solvers of the WLS problems may face the issue of numerical instability and high sensitivity to choice of initial values. To address this issue, Yao et. al [11] proposed a semi-definitive programming (SDP) approach for the DSSE problem which obtained by convex relaxation of the original WLS problem. The method is evaluated by applying to IEEE 13-bus, 34-bus, and 123-bus test systems. Similarly, Zhu et. al [12] proposed a distributed SDP approach for formulating the DSSE problem which can be used for areas with several DSOs and minimum data exchange among DSOs due to data confidentiality concerns.

While WLS-based approaches are fast and simple, it could be very sensitive to bad data

[3]. This led to introducing robust state estimation approaches. Some research papers upgrade the WLS-based approaches to improve the robustness of state estimation. For instance, Wu et. al [13] developed a DSSE method for a grid with limited real-time measurements or with delayed information from smart meters. To provide robust results, a machine learning function was used to create inputs for the weight matrix of WLS problem in the state estimator. Test system data were generated using power flow analysis at each time interval and then errors were added to the system intentionally to simulate different measurement errors. Simulation results confirm the robustness of the results against the measurement errors, the type, location and accuracy of measurements, and the temporary failure of the communication system. In contrast, there are other works that introduce new methods for robust DDSE. Liu et. al [14] proposed a methodology based on matrix completion approach to perform a robust DSSE. The matrix completion approach uses the known elements in the matrix to estimate the missing elements by solving the rank minimization problem. In the proposed approach, system information was used to form the system state-measurement matrix, and distribution grid model and Ohm's law are added to the rank minimization problem as constraints. A decentralized PMU-based robust state estimation method for distribution grids including a utility grid and several micro-grids was introduced by Lin et. al [15]. The state estimation problem is formulated as a quadratic optimization problem for utility grid and micro-grids. Each micro-grid is responsible for evaluating its bad data measurements and an iterative algorithm with minimum data exchange between operators is proposed to perform robust DSSE. Fast convergence and scalability are two main features of this method. Dahale et. al [16] proposed robust formulations for four sparsity-based DSSE approaches 1) 1-D compressive sensing, 2) 2-D compressive sensing, 3) matrix completion, and 4) tensor completion. Simulation results highlight the great performance of compressive sensing based approaches, compared to tensor completion and matrix completion based methods.

Data-driven methods are one of the recently introduced approaches in DSSE. These methods can be an auxiliary part for solving the DSSE problem, such as using neural network (NN) method for generating initial points for solving the main optimization problem [17], or applying machine learning for exploiting pseudo measurements[18]. Data driven approaches could also be used to solve the DSSE problem. NN is one of the most common data-driven approaches for DSSE [19], [20], [21]. Kim et. al [19] introduced a modified long short-term NN for state estimation in hybrid DC/AC distribution grids. Zamzam et. al [20] proposed a NN method that utilizes the structure of the power grid for DSSE. The proposed architecture reduces the number of coefficients required for mapping from the measurements to the network state which prevents overfitting and reduces the complexity of the training stage. Among other data driven approaches, Weng et. al [22] introduced a data-driven DSSE approach that uses the power system patterns and physics to clean data. Supervised learning were used to learn the relationship between the measurement and the state of the system using historical data. Using periodic patterns in power systems, an approach was suggested to speed up the estimation up to 1000 times. To benefit from the advantages of both data driven and classical methods, Anubi et. al [23] proposed an enhanced resilient DSSE algorithm, which combines a data-driven model with compressive sensing regression

method. Using this algorithm helps the system estimator to recover the true state of the system if faced with false data injection attacks which mislead regression-based algorithms.

Nevertheless, all of the mentioned studies are focused on medium voltage (MV) distribution grids (here MV refers to voltage levels higher than 0.4 kV). This imposes a considerable difference to the LV systems studied in this work as LV systems are typically more unbalanced than MV systems. As customers loads are connected to different phases, uneven load distribution between the phases often arise, marking the need for per phase voltage estimations. Moreover, MV loads are not as volatile as the aggregation of customer loads from the connected LV networks results in a lower load variance, supposedly easier to estimate. Additionally, there are methods for MV distribution grid can rely on more measurements than what is practically feasible in LV networks. Furthermore, simulation studies and validation through simulation is commonly applied, while real-world implementations are scarce.

With this in mind, it is worth to mention the study in [24], where the authors develop a NN approach to estimate voltages in a 0.4 kV distribution grid. However, it seems that the method is developed based on confidential customer data, and it could be questioned if these data inputs should be used for operational purposes. Meanwhile, the reported root mean squared error (RMSE) is 0.59 V, suggesting better accuracy can be achieved in the voltage estimations. Furthermore, the method further requires retraining after 20 days, which could also be improved. In [25] the authors derive a method for voltage control in LV grids with high levels of PV panels, including a remote voltage estimation technique. However, the method relies on load estimations of customers as well as number of customers to produce a generic feeder and is rather designed for voltage control in networks with on-load-tap-changers, limiting the applicability in the context of this study. Furthermore, in [26] a remote voltage estimation method for radial LV grids is developed, combining a series power flow calculations and polynomial regression. While the model shows good accuracy, it relies on pseudo-measurements, which can complicate the eventual required retraining of the model and applicable and relevant pseudo-measurements need to be ensured. The model is also tested based on simulated data and the computation time to fit the model is 0.79 h, meaning the computational efficiency could be improved.

Reviewing the literature highlights of DSSE, it is realized that methods for estimating remote voltages in radial LV grids is scarce, as opposed to the many methods found for MV grid DSSE. Furthermore, methods with higher accuracy and lower computational burden are crucial for the DSOs to fully realize remote voltage estimation techniques for grid operation. In addition, methods relying on few measurements that are based on and validated for real-world data are needed. In the light of this, the present paper contributes to the field by:

- Proposing a data driven approach for node voltage estimation in unbalanced LV grids.
- Combining a grey-box modeling approach, to gain explainability, and a generalized additive modeling approach, to achieve considerable low computational burden, which makes the method usable for online monitoring
- Deriving the method for a real-world experimental setup, and validating the results

with high accuracy.

In addition, through the real-world setup, it is ensured that the method is based on input variables that are practical for the DSO to measure, and does not rely on pseudo-measurements.

The rest of the paper is structured as follows: Section 2 describes the experimental setup and context for the derived observability model. Section 3 presents the applied method and Section 4 presents and analyzes the results. Section 5 concludes the paper.

2. Framework for model development

This section describes the experimental setup suing a real world LV network as well as the intended use case and goal of the observability model.

2.1. Experimental setup

The studied case is a low voltage grid on Jutland, Denmark, with a 400 kVA 10/0.4 kV transformer serving 170 residential customers. In the network 5-10 customers have electric vehicles (EVs) and 15-20 customers have PV panels with rated sizes 3 - 6 kW. The households use either district heating or heat pump for heating. 12 electronic voltage and current measuring devices have been installed in the grid, placed at the transformer and cable cabinets according to Figure 1. As seen in the figure, the devices were installed on three out of the five feeders. This was deemed the best way to use the limited number of measurement devices giving enough observability on some of the feeders to build models for these feeders, rather than having few (not enough) measurements on all feeders. Measuring all feeders at the transformer required two devices. Due to space limitations in the cable cabinets some of the devices are not collecting current measurements. Table 1 reports which devices that have only voltage measurement or alternatively both voltage and current measurements. The devices deliver per phase data once per second. If measuring current, the devices also measure active power, power factor, harmonic current content and peak current. The DSO operating the grid has full information of cable types and thereby cable impedances and lengths in the grid are known. Active power data from household meters also available and are delivered once per day in a one hour time resolution.

2.2. Observabiltity model for DSO grid operation

The model developed uses data collected from the experimental setup. It is intended to support DSO in diagnosing the state of the grid through increased voltage observability. In Figure 1, the end-nodes are the most critical and the model primarily focuses on estimating such nodes in the system.

To develop a scalable model for DSO grids, which in general are large grids with many nodes and radials, a goal of the method is to use input data from few devices.

The model is further intended to be used in an online monitoring algorithm and thus, priority is given to methods resulting in low computational burden. The final operation

Device	Voltage	Current
T_1,T_2	X	X
BR_M	X	X
BR_E	X	X
G_M	X	X
G_{E1}	X	X
G_{E2}	X	-
G_{E3}	X	-
B_M	X	X
B_{E1}	X	X
B_{E2}	X	X
B_{E3}	X	-

Table 1: Devices installed in the grid (Figure 1), X indicates that voltage or current measurement is available. If current measurements are available, data for harmonic currents, active power and power factor are also delivered.

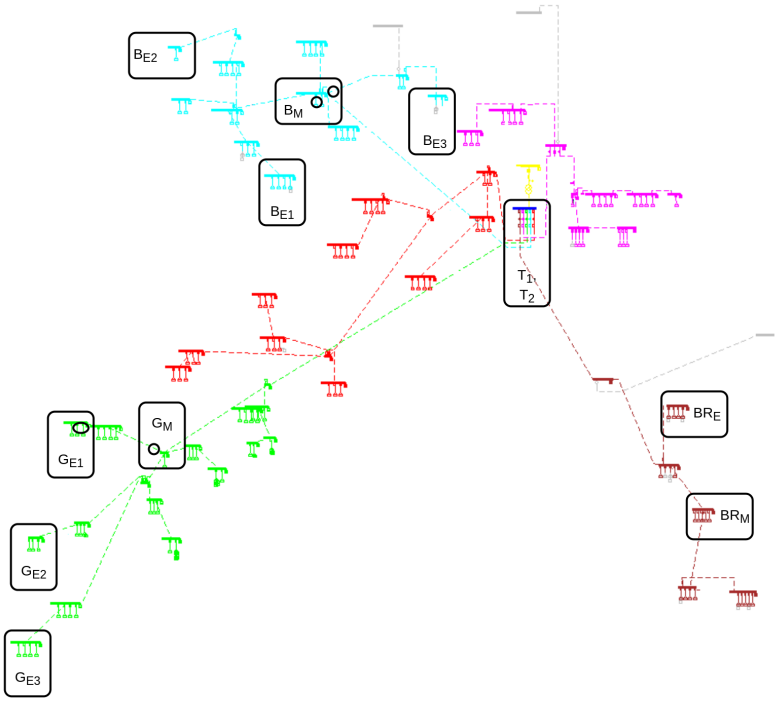


Figure 1: Grid topology and installation of devices. The name of the devices is constructed from the first letter(s) of the color of the feeder, while subscript M indicates that it is a middle node, and subscript E indicates an end-node.

algorithm and architecture can however, be enhanced with smart meter data updates if such data is or would be available.

3. Method

The models derived in this work are physics-informed and data-driven. This means that the physics and known theory is used to build a first structure of the model while the data together with statistical and analytical tools guides the process of finding the final model. The data reflects the system at interest with disturbances that we cannot always measure. Thus, a probabilistic model that does not only reflect a single physical phenomenon, but a real world system and states is achieved.

3.1. Physics information - voltage drop

Since the model was developed for a radial LV network, the equations that guide the first structure of the model are voltage drop equations. For a meshed network, power flow equations might be more suitable, but due to the low X/R ratio in LV grids, this would directly become a set of non-linear equations.

The voltage drop, ΔV, for a feeder is described through:

$$\begin{aligned}\Delta V &= V_S + IR\cos\theta + IX\sin\theta - \sqrt{V_S^2 - (IR\cos\theta + IX\sin\theta)^2} \\ &\approx IR\cos\theta + IX\sin\theta\end{aligned} \tag{1}$$

where R and X are the resistance and reactance of the cable or line, I is the current through the feeder to the load, $\cos\theta$ is the power factor of the load, $\sin\theta$ is the reactive factor of the load, and θ is the phase angle difference between voltage and current at the receiving end [27]. V_S is the sending end voltage (i.e. at the node upstream the network) and V_R is the receiving end voltage (i.e. at the node downstream the network) such that:

$$V_S = V_R + \Delta V \tag{2}$$

The approximation expressed in Eq. (1) is one of the common approximations that is often used in power systems engineering [28]. It should be noted that the equations are used to calculate line-to-neutral voltage drops [27]. Since the LV grids are unbalanced the neutral wire might carry currents and therefore, the neutral might not be at zero potential [29, 30]. Thus, we will investigate if a term for the neutral current voltage drop, ΔV_N, should be included in Eq. (2) such that:

$$V_S = V_R + \Delta V + \Delta V_N \tag{3}$$

Looking at Figure 1 it is realized that if using the nodes with installed devices as sending and receiving end voltage inputs (V_S and V_R) to Eq. (1) there will be loads not only at the receiving end, but also along the feeder. The effective length of the cable to the load center and therefore resistance and reactance parameters to the voltage drop equations was discussed in [28], where the author suggested to calculate the load center distance as it varies with the total ampere distribution along the feeder. However, their suggested method becomes unpractical for a real-time algorithm as it requires real-time input from all households. Instead we will fit a parameter in the model, based on available data, that reflects

the effective resistance and reactance of the feeder. The resistance also increases with temperature. This might lead to seasonal deviations in the model and will be investigated as well.

Since DSO grids are quite large and the installation needs to be repeated many times, another objective of this work is to require as few devices for the model input variables as possible. Turning to classical WLS state estimation is not an option here as it would require more devices that we wish to install to have enough observability or very accurate pseudo-measurements in at least the same time resolution as the model (minimum 10 minute). Instead we use the data at hand to develop a model for that estimates the states that are of most concern, i.e. the states at the customer premises.

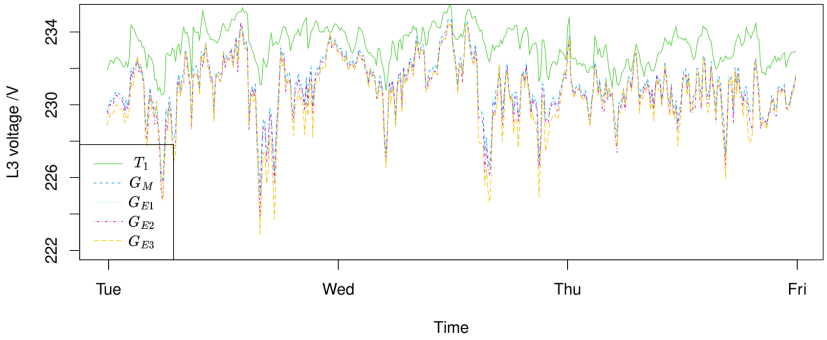
3.2. Data analysis

This section presents a data analysis, supporting the selection of input parameters. Here, the input parameters analyzed were inspired by the voltage drop equations, (eqs. (1) – (3)). To avoid redundant discussion, the data analysis only presents data for the third phase, L3. Model input parameters used in the model selection process are listed in Table 2. The training data set is from April 18th 2022 to April 30th 2022 and the test data set is from May 1st 2022 to May 31st 2022.

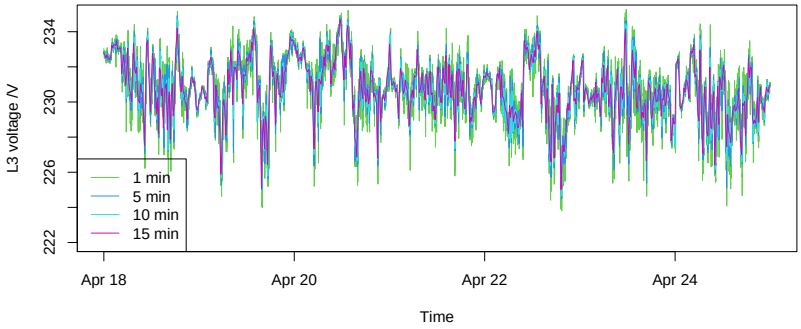
The voltage time series are seen in Figure 2 for the green feeder in the grid (see Figure 1). It is seen that there is a significant voltage drop from the transformer to the devices at the edge of the grid, and that the variations in voltage drop seem quite correlated. This behaviour is further supported by Figure 3, where higher correlations between the edge device voltages compared to the transformer voltage are seen. Figure 4 presents scatter plots, data density and correlation for relevant inputs. It is noteworthy that although G_M voltage has a higher correlation to the other edge voltages, the current for G_{E1} has higher

Variable	**Notation**	**Unit**
V_{T1}	voltage at T_1	V
V_{GM}	voltage at G_M	V
V_{GE1}	voltage at G_{E1}	V
V_{GE2}	voltage at G_{E2}	V
V_{GE3}	voltage at G_{E3}	V
I_{T1}	current at T_1	A
PF_{T1}	power factor at T_1	-
$I_{N,T1}$	neutral wire current at T_1	A
$PF_{N,T1}$	neutral wire power factor at T_1	-
$I_{N,GM}$	neutral wire current at G_M	A
$PF_{N,GM}$	neutral wire power factor at G_M	-
I_{GM}	current at G_M	A
PF_{GM}	power factor at G_M	-
I_{GE1}	current at G_{E1}	A
PF_{GE1}	power factor at G_{E1}	-
solar	solar radiation (from DMI)	W/m^2
T_{amb}	ambient temperature (from DMI)	$^\circ C$

Table 2: Measured input variables used in the model selection process. All electrical measurements from the experimental setup are per phase and their placements in the LV grid are seen in Figure 1.



(a) Phase L3 voltages of all devices on green feeder.



(b) Voltages filtering for device G_M, phase L3

Figure 2: Phase L3 voltage times series

correlation to the edge voltages than the current at G_M. Ambient temperature has very low correlation and will be excluded as a potential input. Solar radiation on the other hand has a higher correlation to the voltages, however, it is not necessarily the explanatory variable as it probably coincides with a higher load, in which case the current should be a better input variable as it is supported by eqs. (1) – (3). Nevertheless, it will be investigated. Interestingly, Figure 5 suggests a high correlation between neutral currents and the voltages, which will be further investigated in the model selection process.

As the raw data from the measuring devices has a time resolution of 1 second, filtering of the times series lower time resolutions is required to achieve a reasonable data throughput. In Figure 2b filtering to 1, 5, 10 and 15 minutes times resolutions can be seen for G_M third phase voltage. Comparing 1 minute to 15 minutes time resolution it is seen that the voltage rise and drops appear smoothed and the time series is less volatile which is a natural

outcome of low pass filtering. As the peaks and drops are of concern for power systems operation, we instead aim to find a suitable model for 5 or 10 minute time resolution, which is compromise between data throughput (or computational burden) and data filtering. 10 minute filtering is initially chosen to have the possibility to incorporate environmental data, which has a time resolution of 10 minutes.

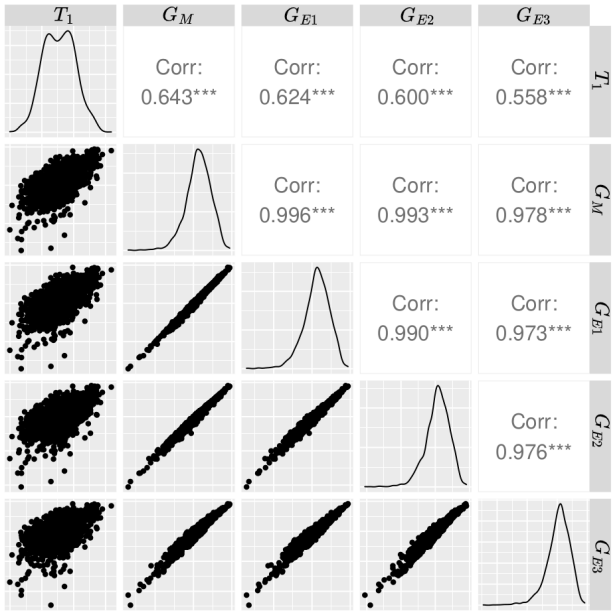


Figure 3: Scatter plots, data density and correlation for phase L3 voltages for all devices on green feeder, using a time resolution of 10 minutes.

3.3. Model selection

It was realized through both modeling approaches in Sections 3.4 and 3.5 that using only the transformer devices measurements was insufficient to model the edge node voltages. The following two sections present two modeling approaches, both following a forward selection process inspired from eqs. (1)–(3). The models are derived using the green feeder in the grid (Figure 1). There are two options for placing the second input device, either at G_M or G_{E1}, as the other two devices do not measure currents. Here, we chose to build models based on measurements from G_{E1}. In this way one end-node voltage will be known by the DSO with a higher accuracy (assuming very small observation errors). Furthermore, if voltage drops along one entire radial is known it will be supposedly be easier to derive models for other end-nodes in the network.

To estimate the voltage at G_M, using devices at T_1 and G_{E1}, two voltage drop drops can be

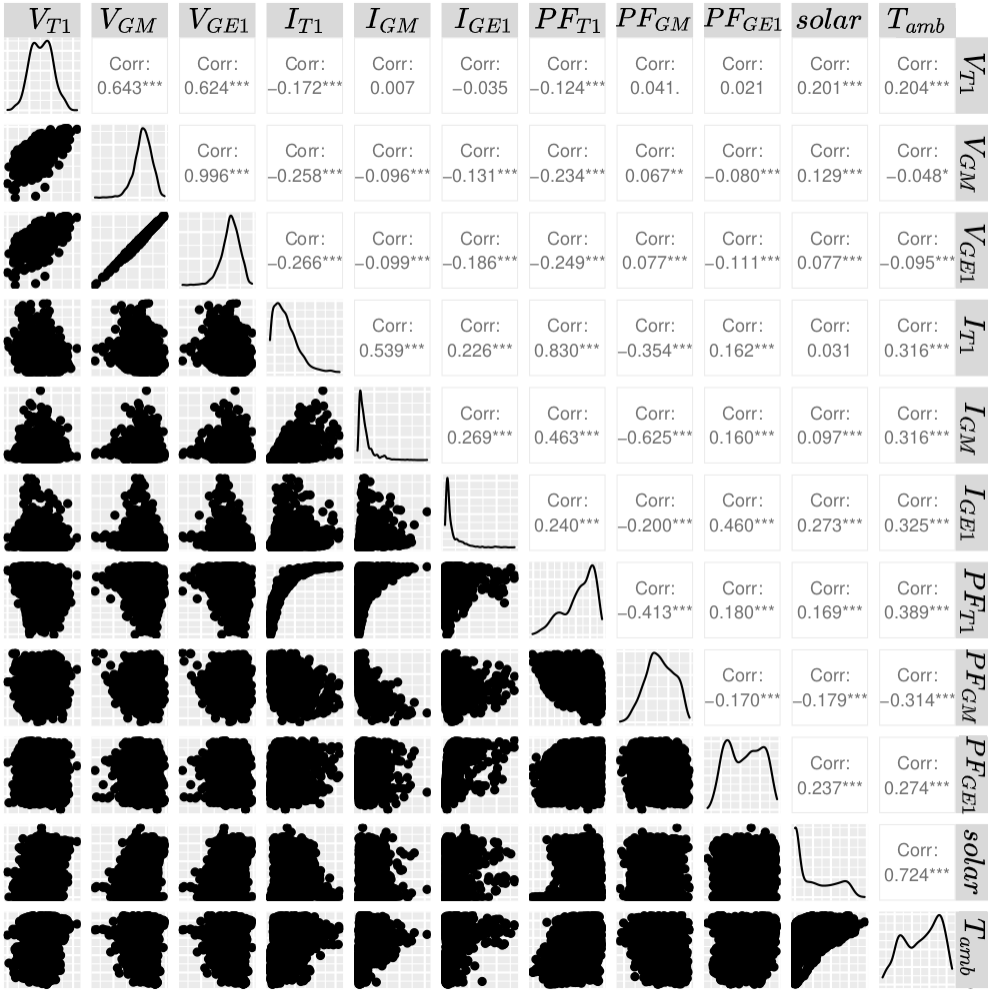


Figure 4: Scatter plots, data density and correlation for voltages and input variables, using a time resolution of 10 minutes.

expressed:

$$V_{GM} = V_{T1} - \Delta V_{T1-GM} \tag{4}$$

$$V_{GE1} = V_{GM} - \Delta V_{GM-GE1} \Leftrightarrow V_{GM} = V_{GE1} + \Delta V_{GM-GE1} \tag{5}$$

where V_{T1}, V_{GM} and V_{GE1} are phase-to-ground voltages at devices T_1, G_M and G_{E1}, respectively, ΔV_{T1-GM} is the voltage drop between devices T_1 and G_M and ΔV_{GM-GE1} is the

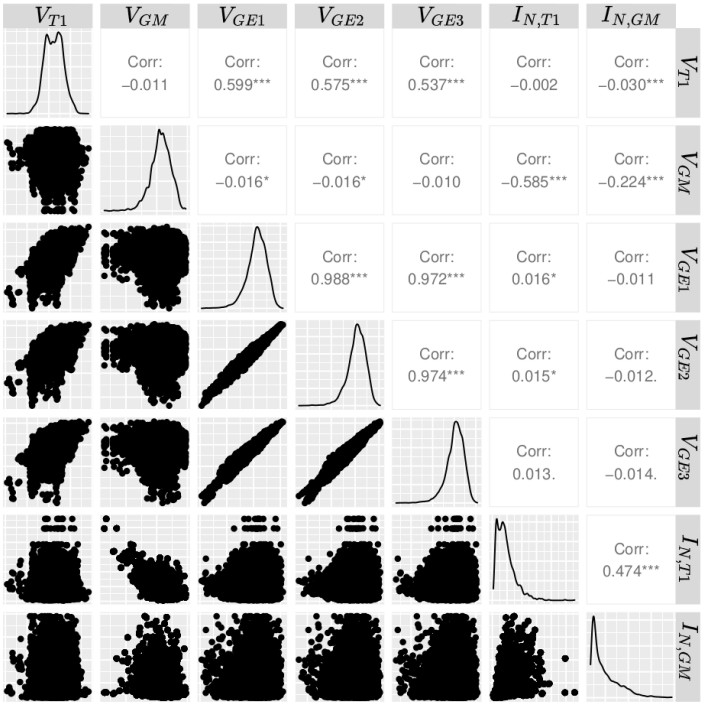


Figure 5: Scatter plots, data density and correlation for voltages and neutral current, using a time resolution of 10 minutes.

voltage drop between devices G_M and G_{E1}. Both GAMs and grey-box models structures are derived assuming that V_{GM} is partly described by Eq. (4) and partly by Eq. (5) such that:

$$\mathbb{E}(V_{GM}) = a(V_{T1} - \Delta V_{T1-GM}) + b(V_{GE1} + \Delta V_{GM-GE1}) + \epsilon \tag{6}$$

where a and b are coefficients to scale the contribution from eqs. (4) and (5), respectively, and ϵ are independent and identically distributed errors assumed to be Gaussian white noise, $\mathcal{N}(0, \sigma^2_{e_k})$. Both modeling approaches described below, start with this formula as an initial model structure.

To evaluate the models in the forward model selection process, we used a similar approach as described in [31]. For each tested model, the auto correlation function (acf) and cumulated periodogram for the residuals were evaluated to investigate if the model assumption of residual white noise had been achieved and if there are any patterns left in the data that are not captured by the model. Root mean squared errors (rmse) for both the training and test data set, along with visual inspection of model estimations on the training set and pre-

dictions on the test set were used to evaluate each model. We also evaluated the significance levels of the estimated parameters, and the model was reduced if higher p-values than 5% were observed. Log likelihood was also used in the grey-box modeling approach and Akaike Information Criterion (AIC) in the GAMs approach.

3.4. GAMs model

Generalized additive models (GAMs) were investigated as a possible model structure of finding the voltage estimation model. The general expression for GAMs is:

$$g(\mu_i) = \boldsymbol{A_i\theta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + ... \tag{7}$$

where μ_i is the expected value of a response variable Y_i, $\boldsymbol{A_i\theta}$ represent parametric part of the model with parameters θ and f_k represent smooth functions of variables x_j [32].

The initially derived GAMs model from Eq. (6) has the following structure:

$$g(V_{GM}) = V_{T1} + s_1(I_{T1}, PF_{T1}) + V_{GE1} + s_2(I_{GE1}, PF_{GE1}) \tag{8}$$

where the inputs are describe in Table 2, while $s_1()$ and $s_2()$ represent smooth functions using B-splines. The parameters of the model were optimized using *gam()* and *gamm()* functions in R package *mgcv version 1.8-40* [32–36]. Furthermore, a Gaussian distribution was used for $g(V_{GM})$.

The initial formula in Eq. (8) is derived using only terms related to resistance of the voltage drop equations. Following the data analysis in Section 3.2 and the voltage drop description in Section 3.1, various extensions of inputs were explored:

- adding a smooth term for the term in Eq. (1) relating to reactive current, by using line current and $\sin\theta = \sin(\arccos(PF_d))$ as inputs, where PF_d is the power factor at device d. This was done to investigate the impact of the reactance in the cable.
- adding voltage drop terms using $I_{N,T1}$ and $PF_{N,T1}$, to investigate the impact of the voltage drop in the neutral wire. Using data for the neutral current in G_{E1} was not possible, since it is not available in the measured data.
- adding temperature as an input by incorporating it in the smooth functions related to cable resistance ($s(I_d, PF_d)$), to investigate if temperature has an impact on the resistance.
- adding a smooth term for solar radiation to investigate potential impact from PV panels in the network.
- adding a seasonal term to investigate if there is an additional daily or hourly variation not explained by other data. This was done using cubic splines with and periodic incremental time step inputs (i.e. a vector $[1, ..., k]$, where k is the period length of a day or hour).

Additionally, variations to the smooth functions were explored.

3.5. Grey-box model

Grey-box models were also explored as another modeling approach as they have proven to be useful when developing data-driven models for physical systems (e.g. in [31, 37]). In grey-box models, known theory of the physical system is used to build a first structure for the model, while statistics and data is used to optimize the parameters in the model [38]. Thus, it is a mixture of deterministic modeling, relying purely on the known theory, and black-box models relying purely on statistics and data. Grey-box models, consisting of a set of first order stochastic differential equations, can be described in a continuous-discrete time state-space representation as follows:

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)dt + \mathbf{B}\mathbf{U}(t)dt + \boldsymbol{\sigma}d\mathbf{w}(t) \tag{9}$$

$$\mathbf{Y}(t) = \mathbf{C}\mathbf{X}(t) + \mathbf{D}\mathbf{U}(t) + \mathbf{e}(t) \tag{10}$$

where $\mathbf{X} \in \mathbb{R}^n$ is the state vector, $\mathbf{Y} \in \mathbb{R}^m$ is the vector of measured outputs, $\mathbf{U} \in \mathbb{R}^p$ is the input vector, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$, $\mathbf{C} \in \mathbb{R}^{m\times n}$ and $\mathbf{C} \in \mathbb{R}^{m\times p}$ are the state-space matrices, $\mathbf{w}$ are standard Wiener processes with incremental covariance matrix $\boldsymbol{\sigma} \in \mathbb{R}^{n\times n}$, and $\mathbf{e} \in \mathbb{R}^m$ are the measurement errors. We assume that Wiener processes are independent, resulting in a diagonal covariance matrix with corresponding variances, σ_i^2, to the ith Wiener process. Measurement errors are assumed to be Gaussian white noise $\mathcal{N}(0, \sigma_{e_k}^2)$ to the kth measured output and uncorrelated to other measurement errors. We further assume that the Wiener processes and the measurement errors are independent.

The initial grey-box model in the model selection process is described as follows:

$$d\Delta V_{T1-GM} = a\left(R_{T1-GM} * dI_{T1}/dt * PF_{T1}\right)dt + \sigma_1 dw_1 \tag{11}$$

$$d\Delta V_{GM-GE1} = b\left(R_{GM-GE1} * dI_{GE1}/dt * PF_{GE1}\right)dt + \sigma_2 dw_2 \tag{12}$$

$$V_{GM} = c(V_{T1} - \Delta V_{T1-GM}) + f(V_{GE1} + \Delta V_{GM-GE1}) + e \tag{13}$$

where a, b, c and f are parameters to be estimated, R_{T1-GM} and R_{GM-GE1} are cable resistance and other inputs are described in Table 2. The state equations are derived by taking the derivative of both sides in Eq. (1).

Again various extensions to the model structure were explored, for instance:

- adding voltage drop related to reactive current by adding $R_{d1-d2} * dI_d/dt * \sin\left(\arccos\left(PF_d\right)\right)$ to eqs. (11) and (12), where $d1$ is the device at the sending end and $d2$ is the device at the receiving end.
- adding a state for the voltage drop in the neutral wire using $I_{N,T1}$ and $PF_{N,T1}$.
- adding temperature as an input to eqs. (11) and (12).
- adding solar radiation as an input to eqs. (11) and (12).

The parameters were optimized using maximum likelihood through the R-package CTSM-R [38, 39].

4. Result and Analysis

Several models were developed and tested throughout the model selection process. While discussing a majority of the models, a special focus is given to the models with the best statistical results.

4.1. GAM model

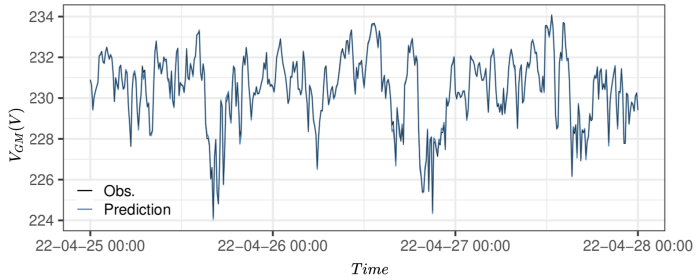
In the GAMs modeling approach it was discovered that neutral wire voltage drop, ambient temperature and reactance terms in the voltage drop equation, (1), were significant terms improving the results of the model while having statistically significant estimated parameters. Adding solar radiation and seasonal cubic splines resulted in insignificant parameters and/or worse predictive capabilities. The resulting GAMs model has the following formula:

$$g(V_{GM}) = s_1(V_{T1}) + s_2(I_{T1}, PF_{T1})T_{amb} + s_3(I_{T1}, \sin(\arccos(PF_{T1}))) + s_4(I_{N,T1}, PF_{N,T1}) + s_5(V_{GE1}) + s_6(I_{GE1}, PF_{GE1})T_{amb} + s_7(I_{GE1}, \sin(\arccos(PF_{GE1}))) \quad (14)$$

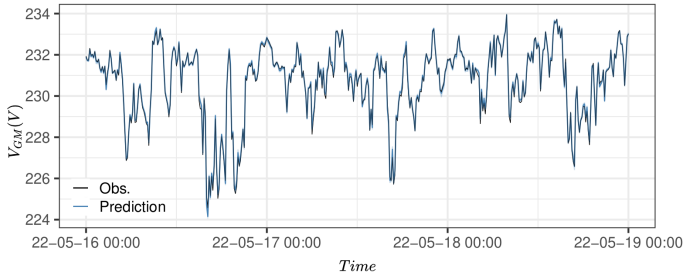
where $s()$ represent smooth functions as described in Section 3.4, where $s_2()$, $s_3()$, $s_6()$ and $s_7()$ are tensor product smooths, and input variables are described in Table 2. The parameter and function terms estimates are seen in Table 3 and log likelihood and rmse values for the training and test data set are presented in Table 4. Interestingly, the model maintains a similar rmse for the test data set, which is promising for estimations and predictions outside the training data set. This is further validated in Figure 6, showing model predictions for the training and test data set. Here, the predictions are very close to the observations and the 95% confidence intervals can barely be seen due to the small span. Although showing good estimations and low rmse, the acf (Figure 7a) and cumulative periodogram (Figure 7b) show that there are patterns in the data that the model does not capture. Additionally, the model has 146 parameters (due to the smooth functions) which limits the explainability and possibility to extend the model to explain other end-node behavior. For this we need a model where the voltage drop for a piece of the radial could be separated/extracted.

Parameter/term	**Estimated**	**P-value**
Intercept	230.6	$< 2e-16$
$s(L3_{T1})$	6.229	$< 2e-16$
$te(I_{T1}, PF_{T1})T_{amb}$	8.595	$1.03e-05$
$te(I_{T1}, \sin(\arccos(PF_{T1})))$	9.831	0.09692
$s(IN_{T1}, PF_{T1})$	10.546	0.00115
$s(L3_{GE1})$	4.733	$< 2e-16$
$te(I_{GE1}, PF_{GE1})T_{amb}$	15.979	$6.86e-06$
$te(I_{GE1}, \sin(\arccos(PF_{GE1})))$	.306	$< 2e-16$

Table 3: Estimated parameters and function terms in Eq. (14) as well as corresponding p-values.



(a)



(b)

Figure 6: GAM model estimations on training, (a), and test, (b), data sets zooming in on three days, respectively. The black line represents the observations and the blue line the model predictions. There is also a 95 % confidence interval indicated by a blue area, but it is visually difficult to see in the graph due to the low standard deviation in the model.

	GAM	**Grey-box**
Log likelihood	1678	1685
rmse training data set	0.099	0.100
rmse test data set	0.109	0.107

Table 4: Comparison of log likelihood between the GAM and grey-box model as well as root mean squared errors (rmse) for the training and test data set.

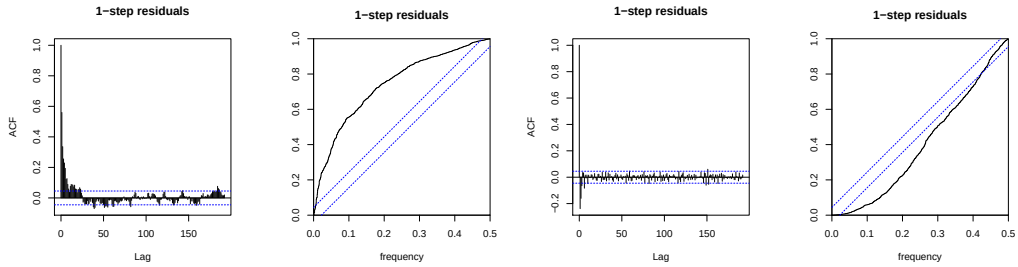
4.2. Grey-box model

In the grey-box model lower model order was achieved with the following model structure:

$$d\Delta V_{R,GM-GE1} = a \cdot R_{GM-GE1} \cdot \frac{dI_{GE1}}{dt} \cdot PF_{GE1} + \sigma_1 dw_1 \tag{15}$$

$$d\Delta V_{X,GM-GE1} = b \cdot X_{GM-GE1} \cdot \frac{dI_{GE1}}{dt} \cdot \sin\left(\arccos\left(PF_{GE1}\right)\right) + \sigma_2 dw_2 \tag{16}$$

$$V_{GM} = c \cdot V_{GE1} + d \cdot \left(V_{R,GM-GE1} + V_{X,GM-GE1}\right) \tag{17}$$



(a) **(b)** **(c)** **(d)**

Figure 7: Residual auto correlation functions and cumulative periodograms, for the GAM model in (a) and (b), and for the grey-box model in (c) and (d). Blue horizontal and diagonal lines indicate a 95 % confidence interval.

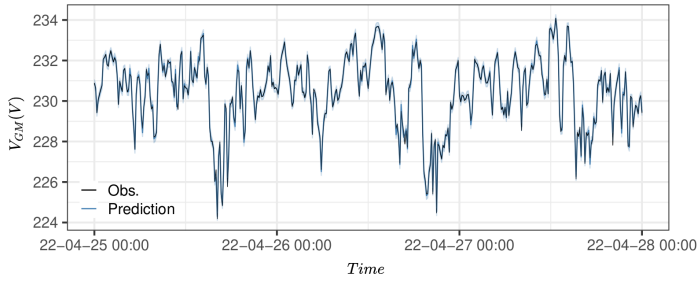
where $\Delta V_{R,GM-GE1}$ represents a state for the voltage drop related to resistance between G_M and G_{E1} (R_{GM-GE1}), whereas $\Delta V_{X,GM-GE1}$ represents a voltage drop state related to the reactance (X_{GM-GE1}) along the same cables. Although inputs from the transformer devices could be used in the model structure it had similar performance without those inputs. For instance, using the device at T_1 to model states for voltage drop along the radial from T_1 to G_M and as well as the neutral current voltage drop, gave a log likelihood of 1687, and rmse of 0.100 and 0.108, for the training and test data set, respectively. This is to be compared with the reported values in Table 4, keeping in mind that the model structure in the latter has 18 parameters as opposed to the model in eqs. (16)–(17) which has 9. Thus, the smaller model structure was preferred. All other extensions as described in Section 3.5 either proved to give insignificant parameters, or produced similar or even worse performance in the predictions.

The estimated model parameters and standard errors are reported in Table 5. Here it can also be seen that p-values are very low and are in general more significant than the estimated parameters in the GAMs model. Improvements compared to the GAM model are also seen in the residual acf (Figure 7c) and cumulative periodogram (Figure 7d), indicating that there is almost no auto correlation in the residuals and the model is better at capturing the behaviour in the data. According to the cumulative periodogram there are either higher frequencies left in the residuals or the model has a small tendency towards overfitting. It should however be noted that voltage tends to change quite rapidly and higher frequency patterns left in the residuals could be due to other load currents, for which we do not have the measurements. Looking at the model estimations on the training and test data set in Figure 8, it can be concluded that although there is a slight deviation in the cumulative periodogram, the model is preforming quite well even on the test data set.

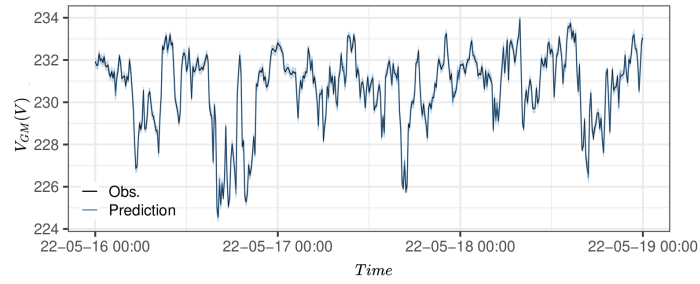
The advantage of the grey-box model is that it has a model structure that offers more explainability compared to the GAM model. Furthermore, the estimated states can be used to derive models for other end-nodes in the radial network.

Parameter	Estimated	Std. error	P-value
Initial state $V_{R,GM-GE1}$	$2.3693e+02$	$1.8476e+01$	$< 2.2e-16$
Initial state $V_{X,GM-GE1}$	$2.3716e+02$	$1.8498e+01$	$< 2.2e-16$
a	$1.5318e-01$	$3.7564e-03$	$< 2.2e-16$
b	$9.8561e-06$	$1.8247e-06$	$7.48e-08$
c	$9.5597e-01$	$2.0718e-03$	$< 2.2e-16$
d	$2.1358e-02$	$9.9780e-04$	$< 2.2e-16$
$ln(\sigma_1)$	$-1.7058e+00$	$5.0160e-02$	$< 2.2e-16$
$ln(\sigma_2)$	$2.0426e-01$	$< 2.2e-16$	
$ln(\sigma_e^2)$	$-4.5660e+01$	$7.3097e-02$	$< 2.2e-16$

Table 5: Estimated parameters for the final grey-box model in eqs. (16) – (17), with standard errors (Std. errors) and corresponding p-values.



(a)



(b)

Figure 8: Grey-box model estimations on the training, (a), and test, (b), data sets zooming in on three days, respectively. The black line represents the observations and the blue line the model predictions. There is also a 95 % confidence interval indicated by a blue area, but it is visually difficult to see in the graph due to the low standard deviation in the model.

4.3. End-node estimation

Since the GAM model indicated good estimation performance for the voltage at G_M, although having limited explainabilty, we return to this modeling approach when estimating the end-node voltages. Here less explainability is required, because the outputs will not

be used for further modeling. The best results were acquired using the following model structure:

$$V_{end} = s(V_{est,GM}) + s(\Delta V_{est,R,GM-GE1}, I_{T1}) + s(t_{day}) \tag{18}$$

where $s(t_{day})$ is a seasonal spline for the daily variation using B-splines of degree 3 and 144 knots and $\Delta V_{est,R,GM-GE1}$ and $V_{est,GM}$ are the estimated states from the grey-box model. It was tested to incorporate $\Delta V_{est,X,GM-GE1}$, however, insignificant parameters were observed and the model had to be reduced. Estimating the voltage at G_{E2}, V_{GE2}, gave rmse values of 0.22 V and 0.24 V for the training and test data sets, respectively, as well as the predictions in Figure 9. The voltage estimations, V_{GE3}, at G_{E3} are seen in Figure 10, for which the rmse was 0.39 V and 0.49 V, respectively. If instead using a GAM model with the measured voltage at G_M (such that $V_{end} = s(V_{GM}) + s(I_{GM}, PF_{GM}) + s(t_{day})$) rmse values were 0.19 V and 0.21 V for voltage at GE2, and 0.40 and 0.50 for voltage at G_{E3} for the training and test data set respectively.

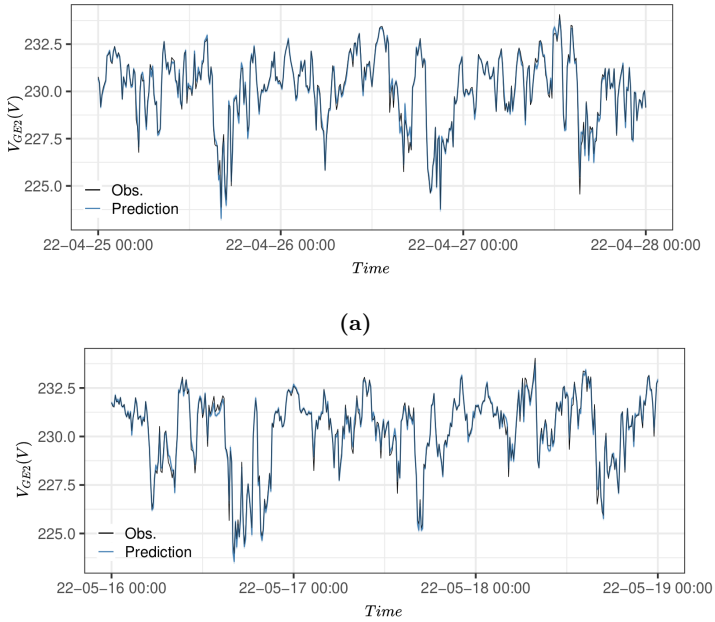
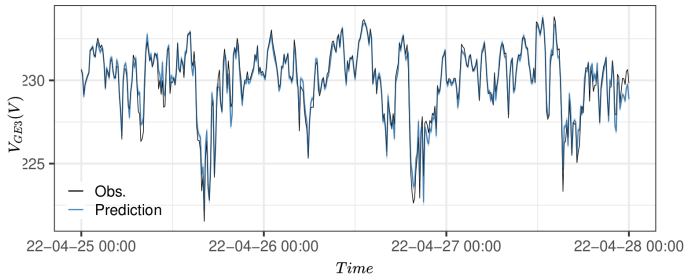
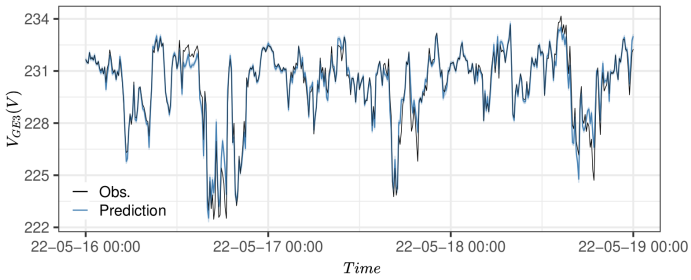


Figure 9: GAM model estimations on the training, (a), and test, (b), data sets for V_{GE2} zooming in on three days, respectively. The black line represents the observations and the blue line the model predictions. There is also a 95 % confidence interval indicated by a blue area.



(a)



(b)

Figure 10: GAM model estimations on the training, (a), and test, (b), data sets for V_{GE3} zooming in on three days, respectively. The black line represents the observations and the blue line the model predictions. There is also a 95 % confidence interval indicated by a blue area.

4.4. Application proposed method and future set up extension

Following the results from the end-node estimations using G_{E1} or G_M as model input produce quite similar model errors. Here, building the model using device input from G_{E1}, located at the end of the green feeder in the LV grid (Figure 1), has a clear advantage in that one end-node voltage is known with high certainty. Furthermore, the models can be used in an online estimation or forecasting algorithm due to fast computation times. The grey-box model parameters are optimized in 13.7 seconds, while the GAM models for voltages at G_{E2} and G_{E3} are optimized in 1.1 and 1.3 seconds, respectively, using an Intel core i7 @ 1.90 Ghz, 16 GB RAM and running on Linux Pop! OS version 21.10.

Although producing reasonable estimations the model and experimental setup can be improved. There are two main action paths to improve the experimental setup:

1. Improve measurements at G_{E1}, to measure all customers at the end-node as well as the neutral wire current.
2. Improve measurements at G_M, to measure all outgoing cables, including neutral wires.

In the installation setup the electrical data for two out of three customers at G_{E1} are measured. This could also explain that the potential presence of high frequency residuals for

the grey-box model in Figure 7c, meaning that there are high frequency patterns that the model might miss because the data is not available. Furthermore, the voltage drop along the neutral wire was a significant input to the GAM model, but was only possible to model between T_1 and G_M, since neutral phase data was only available at T_1. In the grey-box model using inputs from only G_{E1} was proven to provide the best results and therefore, the impact form neutral wire voltage drop could not be explored. Thus, in a future installation setup measurements for the neutral wire should be provided. This also highlights the importance of using real world data in radial LV grids as unpredicted harmonics and unbalance and could results in neutral wire currents and affect the model output. In setup improvement 2. it is suggested to measure all outgoing cables and neutral wire. The idea here is that the GAM model can be used directly to estimate end-node voltages. It should however, be noted that improvement 1. still has the advantage of knowing one end-node voltage with very high certainty and better explainabilty in extending the model structure to other radials. A good practice would be to install both setup improvements and evaluate which one that provides better model performance.

With the short computation time the models are also suitable for daily updates using smart meters. Using smart meters directly to monitor voltage might not be possible due to GDPR and costly communication infrastructure with large data flows. Instead smart meter data is quite often collected once a day by DSOs. If this is the case model parameters can be optimized using the data collected daily, to maintain the accuracy of the model, while the model provides online real-time estimates with low computational burden during the daily operation.

5. Conclusion

In this work we have developed a data-driven node voltage observability method for real-time monitoring in radial LV grids. The method uses input from only one device at the end of a radial and is designed to provide phase voltages in a 10 minutes time resolution. Such estimations are useful for distribution system operation as these grids are typically unbalanced.

The method compromises grey-box and GAM modeling techniques and has proven to give reasonable estimations on both the training data set (13 days) and the test data set (31 days), with rmse of 0.002 – 0.0004 p.u. for the studied nodes. The grey-box model demonstrated explainability in describing the voltage drop along parts of a radial, which could be used as input to the computationally lighter GAM model. The method also provides confidence intervals which gives the DSO the possibility to apply risk informed strategies.

The proposed method has a low computational burden, which makes it useful for online monitoring algorithms, as opposed to other techniques relying on heavy data flows and complex communication infrastructure.

Furthermore, the method is derived using data from a real-world radial LV grid. Working with real-world data offers considerable contribution towards application of observability models in DSO grids, as it reflects the real system with unavoidable disturbances, not captured through simulations in ideal conditions.

5.1. Future work

From analysing and using the data in the model development useful insights have been gained and a few improvements to the experimental setup can be suggested. The improvements involve more comprehensive measurements at the end-node used for model building as well as at the middle node (G_M). It is suggested that both improvements are installed and the resulting models evaluate to obtain the better model performance.

6. Acknowledgments

This work was supported by the Flexible Energy Denmark (FED) project funded by Innovation Fund Denmark under Grant No. 8090-00069B and Ebalanceplus project funded by the European Union's Horizon 2020 under the grant agreement number of 864283. The authors would also like to thank Claus Schack Urup for technical support and providing the data for this study.

References

[1] F. D'Ettorre, M. Banaei, R. Ebrahimy, S. A. Pourmousavi, E. Blomgren, J. Kowalski, Z. Bohdanowicz, B. Łopaciuk-Gonczaryk, C. Biele, H. Madsen, Exploiting demand-side flexibility: State-of-the-art, open issues and social perspective, Renewable and Sustainable Energy Reviews 165 (2022) 112605. doi:10.1016/j.rser.2022.112605.
URL https://linkinghub.elsevier.com/retrieve/pii/S1364032122005007

[2] I. Táczi, B. Sinkovics, I. Vokony, B. Hartmann, The challenges of low voltage distribution system state estimation—an application oriented review, Energies 14 (17) (2021). doi:10.3390/en14175363.

[3] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, F. Bu, A survey on state estimation techniques and challenges in smart distribution systems, IEEE Transactions on Smart Grid 10 (2) (2019) 2312–2322. doi:10.1109/TSG.2018.2870600.

[4] H. Wang, N. N. Schulz, A Revised Branch Current-Based Distribution System State Estimation Algorithm and Meter Placement Impact, IEEE Transactions on Power Systems 19 (1) (2004) 207–213. doi:10.1109/TPWRS.2003.821426.

[5] M. Pau, P. A. Pegoraro, S. Sulis, Efficient branch-current-based distribution system state estimation including synchronized measurements, IEEE Transactions on Instrumentation and Measurement 62 (9) (2013) 2419–2429. doi:10.1109/TIM.2013.2272397.

[6] M. E. Baran, J. Jung, T. E. McDermott, Including voltage measurements in branch current state estimation for distribution systems, in: 2009 IEEE Power and Energy Society General Meeting, PES '09, 2009. doi:10.1109/PES.2009.5275479.

[7] A. Monticelli, State Estimation in Electric Power Systems: A Generalized Approach, Vol. 7, Springer, New York, NY, USA, 1999.
URL https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civilwars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625

[8] W.-M. Lin, J.-H. Teng, State Estimation for Distribution Systems with Zero-Injection Constraints, Tech. Rep. 1 (1996).

[9] Q. Chen, D. Kaleshi, Z. Fan, S. Armour, Impact of Smart Metering Data Aggregation on Distribution System State Estimation, IEEE Transactions on Industrial Informatics 12 (4) (2016) 1426–1437. doi:10.1109/TII.2016.2573272.

[10] Y. Nie, C. Y. Chung, N. Z. Xu, System State Estimation Considering EV Penetration with Unknown Behavior Using Quasi-Newton Method, IEEE Transactions on Power Systems 31 (6) (2016) 4605–4615. doi:10.1109/TPWRS.2016.2516593.

[11] Y. Yao, X. Liu, D. Zhao, Z. Li, Distribution System State Estimation: A Semidefinite Programming Approach, IEEE Transactions on Smart Grid 10 (4) (2019) 4369–4378. doi:10.1109/TSG.2018.2858140.

[12] H. Zhu, G. B. Giannakis, Power system nonlinear state estimation using distributed semidefinite programming, IEEE Journal on Selected Topics in Signal Processing 8 (6) (2014) 1039–1050. doi:10.1109/JSTSP.2014.2331033.

[13] J. Wu, Y. He, N. Jenkins, A robust state estimator for medium voltage distribution networks, IEEE Transactions on Power Systems 28 (2) (2013) 1008–1016. doi:10.1109/TPWRS.2012.2215927.

[14] B. Liu, H. Wu, Y. Zhang, R. Yang, A. Bernstein, Robust Matrix Completion State Estimation in Distribution Systems, Tech. rep.

[15] C. Lin, W. Wu, Y. Guo, Decentralized Robust State Estimation of Active Distribution Grids Incorporating Microgrids Based on PMU Measurements, IEEE Transactions on Smart Grid 11 (1) (2020) 810–820. doi:10.1109/TSG.2019.2937162.

[16] S. Dahale, H. S. Karimi, K. Lai, B. Natarajan, Sparsity based approaches for distribution grid state estimation - A comparative study, IEEE Access 8 (2020) 198317–198327. doi:10.1109/ACCESS.2020.3035378.

[17] A. S. Zamzam, X. Fu, N. D. Sidiropoulos, Data-Driven Learning-Based Optimization for Distribution System State Estimation, IEEE Transactions on Power Systems 34 (6) (2019) 4796–4805. doi:10.1109/TPWRS.2019.2909150.

[18] K. Dehghanpour, Y. Yuan, Z. Wang, F. Bu, A Game-Theoretic Data-Driven Approach for Pseudo-Measurement Generation in Distribution System State Estimation, IEEE Transactions on Smart Grid 10 (6) (2019) 5942–5951. doi:10.1109/TSG.2019.2893818.

[19] D. Kim, J. M. Dolot, H. Song, Distribution System State Estimation Using Model-Optimized Neural Networks, Applied Sciences (Switzerland) 12 (4) (2 2022). doi:10.3390/app12042073.

[20] A. S. Zamzam, N. D. Sidiropoulos, Physics-Aware Neural Networks for Distribution System State Estimation (3 2019).
URL http://arxiv.org/abs/1903.09669

[21] J. H. Menke, N. Bornhorst, M. Braun, Distribution system monitoring for smart power grids with distributed generation using artificial neural networks, International Journal of Electrical Power and Energy Systems 113 (2019) 472–480. doi:10.1016/j.ijepes.2019.05.057.

[22] Y. Weng, R. Negi, C. Faloutsos, M. D. Ilic, Robust Data-Driven State Estimation for Smart Grid, IEEE Transactions on Smart Grid 8 (4) (2017) 1956–1967. doi:10.1109/TSG.2015.2512925.

[23] O. M. Anubi, C. Konstantinou, Enhanced resilient state estimation using data-driven auxiliary models, IEEE Transactions on Industrial Informatics 16 (1) (2020) 639–647. doi:10.1109/TII.2019.2924246.

[24] M. Pertl, P. J. Douglass, K. Heussen, K. Kok, Validation of a robust neural real-time voltage estimator for active distribution grids on field data, Electric Power Systems Research 154 (2018) 182–192. doi:10.1016/j.epsr.2017.08.016.

[25] A. T. Procopiou, L. F. Ochoa, Voltage Control in PV-Rich LV Networks Without Remote Monitoring, IEEE Transactions on Power Systems 32 (2) (2017) 1224–1236. doi:10.1109/TPWRS.2016.2591063.

[26] V. Rigoni, A. Soroudi, A. Keane, Use of fitted polynomials for the decentralised estimation of network variables in unbalanced radial LV feeders, IET Generation, Transmission and Distribution 14 (12) (2020) 2368–2377. doi:10.1049/iet-gtd.2019.1461.

[27] IEEE Standards Board, IEEE Std. 141-1993, Recommended Practice for Electric Power Distribution for Industrial Plants, Tech. rep. (1993). doi:10.1049/pe:19880018.

[28] S. S. Pandian, Various considerations for estimating steady-state voltage drop in low voltage AC power distribution systems, Conference Record - Industrial and Commercial Power Systems Technical Conference (2006). doi:10.1109/icps.2006.1677307.

[29] L. Degroote, B. Renders, B. Meersman, L. Vandevelde, Neutral-point shifting and voltage unbalance due to single-phase DG units in low voltage distribution networks, 2009 IEEE Bucharest PowerTech:

Innovative Ideas Toward the Electrical Grid of the Future (2009) 1–8doi:10.1109/PTC.2009.5281998.
[30] T. H. Jung, G. H. Gwon, C. H. Kim, J. Han, Y. S. Oh, C. H. Noh, Voltage Regulation Method for Voltage Drop Compensation and Unbalance Reduction in Bipolar Low-Voltage DC Distribution System, IEEE Transactions on Power Delivery 33 (1) (2018) 141–149. doi:10.1109/TPWRD.2017.2694836.
[31] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy and Buildings 43 (7) (2011) 1511–1522. doi:10.1016/j.enbuild.2011.02.005.
URL http://dx.doi.org/10.1016/j.enbuild.2011.02.005
[32] S. N. Wood, Generalized Additive Models : An Introduction with R, second edi Edition, CRC Press, 2017. doi:10.1201/9781315370279-4.
[33] S. N. Wood, Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models, Journal of the Royal Statistical Society (B) 73 (1) (2011) 3–36.
[34] S. Wood, N., Pya, B. S"afken, Smoothing parameter and model selection for general smooth models (with discussion), Journal of the American Statistical Association 111 (2016) 1548–1575.
[35] S. N. Wood, Stable and efficient multiple smoothing parameter estimation for generalized additive models, Journal of the American Statistical Association 99 (467) (2004) 673–686.
[36] S. N. Wood, Thin-plate regression splines, Journal of the Royal Statistical Society (B) 65 (1) (2003) 95–114.
[37] P. A. Stentoft, T. Munk-Nielsen, L. Vezzaro, P. A. Stentoft, H. Madsen, J. K. Møller, L. Vezzaro, P. S. Mikkelsen, Towards model predictive control: Online predictions of ammonium and nitrate removal by using a stochastic ASM, Water Science and Technology 79 (1) (2019) 51–62. doi:10.2166/wst.2018.527.
[38] R. Juhl, J. K. Møller, H. Madsen, ctsmr - Continuous Time Stochastic Modeling in R, arXiv (2016).
URL http://arxiv.org/abs/1606.00242
[39] R. Juhl, J. K. Møller, J. B. Jørgensen, H. Madsen, Modeling and Prediction Using Stochastic Differential Equations (2016) 183–209doi:10.1007/978-3-319-25913-0{_}10.

5 Paper D

Tilman Weckesser, Dominik Franjo Dominković , Emma M.V. Blomgren, Amos Schledorn, Henrik Madsen,
"Renewable Energy Communities: Optimal sizing and distribution grid impact of photo-voltaics and battery storage"
Applied Energy (2021)

Renewable Energy Communities: Optimal sizing and distribution grid impact of photo-voltaics and battery storage

Tilman Weckesser[a,*], Dominik Franjo Dominković[b], Emma M.V. Blomgren[b], Amos Schledorn[b], Henrik Madsen[b,c]

[a] *Danish Energy, Frederiksberg, Denmark*
[b] *Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark*
[c] *Norwegian University of Science and Technology (FME-ZEN project), Norway*

Abstract

In this paper, an extensive study of Renewable Energy Communities and their potential impact on the electric distribution grid has been carried out. For that purpose, a Linear Programming optimization model sizing the energy community's Photo-Voltaic and Battery Energy Storage System was developed. The linear programming model was soft coupled with power flow analysis to investigate the impact of different energy community configurations on distribution grids. Different distribution grids (city, suburban, village), different energy community configuration, different operating strategies and different battery placements were investigated. The results showed that when the battery is located at the beginning of the feeder, then the energy community does not impact the observed minimum and maximum voltage. Moreover, it was found that depending on the energy community's operating strategy the low-voltage grid loading can be reduced by up to 58 %. The energy community's sizing showed that optimal capacities of photo-voltaics and communal batteries were up to three times larger for the case of city grid, following the operating strategy of maximizing the energy community's own economic benefit than in other operating strategies and grid types.

Keywords: Energy Communities, energy system sizing, electric battery, distribution grid, power flow analysis, grid impact

*Corresponding author
Email address: TWE@danskenergi.dk (Tilman Weckesser)

Sets	
$\mathcal{T}$	Set of time periods t
Parameters	
T	Technology lifetime
C^{b}	annualized battery investment costs [EUR/MWh]
C^{p}	annualized PV investment costs [EUR/MW]
P_t^{buy}	purchase price for power in time step $t \in \mathcal{T}$ [EUR/MWh]
P_t^{sell}	sales price for power in time step $t \in \mathcal{T}$ [EUR/MWh]
D_t	electricity demand in the EC in time step $t \in \mathcal{T}$ [EUR/MWh]
G_t^{p}	Normalised electricity production from PV in time step $t \in \mathcal{T}$
γ_t^{ch}	battery charging efficiency in time step $t \in \mathcal{T}$
γ_t^{dc}	battery discharging efficiency in time step $t \in \mathcal{T}$
$\underline{\lambda}$	minimum storage level in percent
Variables	
Z	total costs [EUR]
c^{b}	installed battery capacity [MWh]
c^{p}	installed PV capacity [MW]
q_t^{g2c}	power imported from the central grid to the EC in time step $t \in \mathcal{T}$ [MWh]
q_t^{c2g}	power exported from the EC to the central grid in time step $t \in \mathcal{T}$ [MWh]
q_t^{g2d}	power imported from the central grid used for demand satisfaction in time step $t \in \mathcal{T}$ [MWh]
q_t^{g2b}	power imported from the central grid used for charging the battery in time step $t \in \mathcal{T}$ [MWh]
q_t^{p}	power produced by PV and used in time step $t \in \mathcal{T}$ [MWh]
q_t^{p2d}	power production from PV used for demand satisfaction in time step $t \in \mathcal{T}$ [MWh]
q_t^{p2g}	power production from PV exported to the central grid in time step $t \in \mathcal{T}$ [MWh]
q_t^{p2b}	power inflow to the battery from PV in time step $t \in \mathcal{T}$ [MWh]
q_t^{b2d}	power outflow from the battery used for demand satisfaction in time step $t \in \mathcal{T}$ [MWh]
q_t^{b2g}	power exported from the battery to the central grid in time step $t \in \mathcal{T}$ [MWh]
q_t^{dc}	power outflow from the battery in time step $t \in \mathcal{T}$ [MWh]
q_t^{ch}	power inflow from the battery in time step $t \in \mathcal{T}$ [MWh]
l_t	battery storage level in time step $t \in \mathcal{T}$ [MWh]
Additional nomenclature	
C_{norm}	norm investment costs [EUR/MWh]/[EUR/MW]
r	discount rate
f	generic objective function
$\mathcal{M}$	generic set of objectives
$\mathbf{x}$	generic vector of decision variables
$\mathcal{S}$	generic feasible space
ϵ_m	ϵ-bound on objective $m \in \mathcal{M}$
α	peak power exchange
κ	iteration counter in the ϵ-constraint method

Table 1: Nomenclature

1. Introduction

To mitigate climate change the European Union (EU) aims at climate neutrality in 2050 [1]. Realizing the central role of the energy sector in the climate challenge EU launched the Clean Energy package, stating that the share of Renewable Energy (RE) in the energy mix should be at 32 % in 2030 and further that the energy market should be modernized to account for flexibility [2]. To accommodate for increasing RE, the power system is undergoing a paradigm shift, leading to more Decentralized Energy Resources (DERs) in the grid. In consequence, Energy Communities (ECs) have been developed as an innovative and cooperative strategy to share RE DERs, minimize their own consumption and/or to ease the loading of the power grid through utilizing energy flexibility of the active consumers.

ECs are recognized by the EU and the Clean Energy package as "collective energy actions that foster citizens' participation across the energy system" [3]. There are approximately $3,500$ ECs in the EU (in 2019) and in Denmark specifically there are around 700 (in 2019) [3]. Moreover, there are several initiatives and research projects on energy communities among which, IEletrix [4] is a recent addition to the development. ECs are often formed based on social and/or environmental objectives and the participants share or exchange energy resources in a non-commercial manner [3]. In this work, we focus on Renewable Energy Communities (RECs), utilizing locally produced RE and Battery Energy Storage System (BESS). Given the already large number of ECs, the decentralization of the grid and the goals on increasing RE, it is a natural conclusion that such RECs will have an increasing impact on the energy system.

As generation is moved to the edge of the power system, it might stress the Low Voltage (LV) grid. The increasing share of Photo-Voltaic (PV) electricity generation may cause voltage violations in LV grids and lead to changes in load profiles potentially causing reversed power flows [5]. In the context of smart communities with peer-to-peer trading and high adaptation of PVs, [6] finds that as the number of active members rises the risk of overloading increases. These consequences could limit further adaptation of PV panels [5] and calls for research on grid impact and mitigation strategies. This is highly relevant for RECs, generally having a high share of PV installations as well as exchanging the generated RE locally, leading to different power flows as reported in [5].

To mitigate grid related issues associated with RE generation in LV networks, such as violation of voltage and thermal limits, BESS can be utilized. Di Clerico et al. [7] investigate BESS providing ancillary services, owned and operated by TSO or DSO and finds that the solution can improve grid management and potentially delay grid investments if used for power peak reduction. In [8] the authors find that a multi-purpose and multi-tasking BESS supporting grid relief, reduces local demand and power peak stress. Furthermore, Manbachi et al. [9] find that their designed communal BESS operating in all 4 power quadrants optimally supports Volt-VAR optimization by discharging during peak hours. However, if the BESS is owned by a REC the strategy to operate the BESS can differ quite widely from grid supportive objectives. Resch [10] conclude that the most profitable operation strategies, from a customer perspective, for large scale batteries in Germany 2017 were self-consumption optimization and primary frequency control. Hence, communal BESS can greatly contribute to reliving the grid, but it is of importance to investigate the impact of REC BESSs with various operational strategies, realistically applied by consumers such as RECs.

Communal BESS can also have social benefits and has the potential to be more profitable than individual BESS [11]. Barbour et al. [12] states that communal BESSs decrease the total amount of deployed storage capacity, increase the total rate of return and self-sufficiency compared to individually owned BESSs. Asimakopoulou & Hatziargyriou [13] compare centralized dispatch of DER with DER dispatched by local ECs/aggregator and finds that the later is more beneficial for the consumer as well as giving lower system marginal prices and average load serving costs. As the market of behind-the-meter BESS is increasing, Liu et al. [14] point out the importance of research on behind-the-meter energy storage systems to ensure optimal sizing and operation as well as their role in improving reliability in local electricity grids. However, these type of studies often overlook the impact on the power system. Consequently, there is a lack of both comparative research for operational strategies in economical terms, but also to investigate their respective impact on distribution grids. In addition, [10] states that research is needed on combining profitable and grid supportive operation strategies. If this is not done there is a risk of further stressing the grid instead of relieving it. Hatta et al. [15] further stress the importance of interoperation between the smart community and the rest of the power system to not compromise on power quality.

Through our research, we aim to provide a grid impact study comparing different optimized operational strategies of RECs with shared BESS. The REC configuration in the distribution grid is varied to account for various power flow scenarios. Additionally, we perform case studies for three different grid types typical for city, suburban and village environments. The objective is to fill the gap of comparative research for various REC and the respective impact on the power system as well as to spur communication between RECs and Distribution System Operators (DSOs). Ideally, the comparative case study can support policy and regulatory frameworks, especially for smart cities and municipalities.

It should be mentioned that smart energy systems and smart energy communities have also been a matter of discussion in recent years [16]. However, in this work we focus on the energy communities' impact on power systems and thus, the electrical side of energy communities only.

Similar work has been conducted, however, not as comprehensive on the comparative side as this work. Faessler et al. [17] develop an autonomous optimization routine for communal and individual BESS to support the grid through load shifting and optimizing self-consumption. However, the grid impact of the developed system is not investigated. The authors in [18] claim to investigate grid impact, in terms of loading, current and voltage for an LV grid in Townsville, Australia, in three different cases: no battery present, with individual battery or with communal battery installed in the system. Yet, only voltage at transformer is presented, whereas the voltage violations could occur elsewhere in the LV grid. Additionally, full insight in behind-the-meter devices is assumed, which might not always be practically applicable.

The authors in [19] design a MILP to optimize a smart community energy system (in the UK) including buildings, energy storage and energy conversion devices and optimize on energy usage, cost and electricity self-sufficiency. Both residential and commercial communities are included in the analysis as well as different price signals. Contrary to our contribution, insight within the buildings is assumed which limits practical implementation. Reis et al. [20] investigates the impact of individual end-users' goals on the energy community's self-sufficiency, using distributed artificial intelligence and optimization. The authors find that individual economic goals can have a negative impact on the community's self-sufficiency. Furthermore, Fina et al. [21] estimate economic potential of shared rooftop

PV systems through developing an optimization model to determine profitability. The study investigates city area, town area (extended city area), mixed area (city outskirts) and rural area. Different house type combinations and different direction of PVs, direct use, storage and fed into grid is also accounted for. Focusing more on the market set up Moura et al. [22] designs a transactive energy market for community microgrids involving public and commercial buildings with BESSs and EVs. The authors find that their method leads to improved matching between demand and local PV production, while also lowering costs. Although these studies investigate ECs, there is no sufficient analysis of the grid impact with relevant measures such as voltage and line loading.

Resch et al. [11] develop an operation strategy model for a community energy storage both for maximizing self-consumption and performing primary frequency control, respectively. Yet again, the authors do not investigate grid types and different REC configurations and indeed not the impact on the LV grid.

Recently, Rocha et al. [23] performed a benefit assessment for RECs and local energy markets, in which they also considered grid impacts. They find issues with voltage and power losses in their case study and state that utilizing flexibility could solve these issues. However, they do not investigate different community strategies, but only optimize on cost. Additionally, they do not include line loading in their grid assessment.

Similar to our work, Hatta et al. [15] conduct a simulation study to evaluate the grid impact of a smart community optimizing their local objectives (cost minimization). The study considers residential consumers in LV grids, PV panels and communal batteries as well as commercial customer areas in high voltage grid, including PV panels, BESS and combined heat and power systems. The authors find that maximum voltage and maximum current increases when including BESS for both residential and commercial areas. Additionally, the authors state that the smart community operation result in worse power quality. However, the study only considers cost optimization and do not compare and investigate other strategies such as optimizing self-consumption which is reported as one of the most profitable strategies for ECs that is expected to grow in number of business cases [10].

In [24] the authors optimally design a communal battery in a system where prosumers own PVs and EVs. They aim to maximize net present value, improve the voltage profile, reduce power loss and mitigate transformer aging and they notice improvements on all points. Additionally, they optimize the battery size and account for maintenance, investment cost and battery bank replacement. However, it is assumed that the battery is designed to meet the needs of DSOs and does not account for the perspective of an EC and their strategies. Furthermore, there is no comparison of configurations and different grid types.

To the best of our knowledge no research has provided a comprehensive comparison for RECs including typical distribution grids, which consist of a Medium Voltage (MV) and LV grids, representing different settlement areas, REC configuration and operational strategies. The aim is to provide a grid impact analysis, that could support bridging the communication gap between RECs and DSOs. As our research is only focused on REC, we hereafter refer to REC simply as Energy Community (EC).

Our contribution are case studies using data from representative LV/MV grids in Germany. The load profiles are based on real measurement data from $300,000$ customers in Denmark. We design a communal PV-BESS system and develop four battery operational strategies through Linear Programming (LP). The grid impact of these are then assessed using power flow analysis in 45 distribution grid scenarios including:

- 5 different REC composition configurations

- 3 different grid types (city, village, suburban)
- 3 different BESS placements

The main research questions are:

- What are cost-optimal installed capacities of PVs and community BESS for ECs based on a current regulatory framework?
- How do different operational strategies of an EC and the EC's configuration impact the distribution grid?
- How do different placements of a communal BESS impact the distribution grid?

The novelty of this paper stems from the application of different methods on the same problem in order to represent the potential scenarios if different perspectives are taken into account, for example a grid perspective and the EC's perspective. Different methods and different assumptions resulted in a much more comprehensive paper than the current state-of-the-art. Thus, this paper primarily focuses on distribution system operators, as well as on investors in ECs and policy makers. Many developed scenarios makes it possible to investigate different potential outcomes, in line with the current significant development of energy communities.

The rest of the paper is organized as follows. Section 2 describes the formulation of the BESS operational strategies, using LP programming, through which the optimal capacities of installed PV and BESS are designed. The section further describes the investigated grids, EC configurations and the structure of the grid impact assessment. Section 3 presents the optimal design of the installed BESS and PV in the ECs, followed by the results from the grid impact assessment for the investigated scenarios. In Section 4 the main findings are discussed and finally, Section 5 concludes the paper.

2. Methods

The overall problem of assessing the impact of different ECs on the grid consisted of two soft-linked sub-problems. First, the LP optimization model was run in order to detect the optimal capacities of PVs and a communal battery that could be installed based on economic goals, taking into account different grid types, EC configurations and different control strategies. Apart from capacities, the LP model also determined the operation profile of the battery. Second, the obtained capacities and battery profiles were loaded into the power flow model that assessed the impact of the EC on the power grid. By solving the capacity extension problem for the EC first, realistic capacities of the distributed PV and the communal battery could be obtained, creating a more realistic EC size than defining the energy capacities based solely on generic assumptions.

The following four control strategies were applied. The first strategy (Strategy 1), maximization of economic benefit, modelled the EC's goal of achieving the lowest system cost for electricity for the EC in the given market and regulation framework. The second strategy (Strategy 2) of the EC, peak shaving, featured a multi-objective optimization that resulted in a Pareto front of the EC's peak demand versus total system costs, as explained in section 2.1.3. The third strategy (Strategy 3), maximization of self-sufficiency has the goal of maximizing the consumption of the electricity generated by the EC's own PV. To this end, the optimization enforced that the

power generated by the EC's PVs was either locally consumed or stored in the communal battery. Strategy 3 had two alternations depending on whether the battery was located behind the meter (Strategy 3*a*) or in front of the meter (Strategy 3*b*) - as this distinction has an important consequence regarding the taxation. The two alternations are presented in sections 2.1.4 and 2.1.5.

Furthermore, the impact on three different distribution grid types (city, suburban and village grid) was evaluated. Five different connection schemes of the community members to the low and medium voltage feeders consisted of different combinations of households and a large commercial customer connected to the distribution grid via different set-ups of low and medium voltage feeders. Both of those variations are described in section 2.2.3.

2.1. Optimizing sizing and operation of the PV & BESS

In this section, model equations of the LP for the communal energy system are given. The model formulation is flow-based, where a flow refers to the virtual, market-level, flow of electricity. This formulation allows to formulate the problem in a general way, and later introduce additional constraints for the four strategies. If not mentioned otherwise, the range of values for all variables is set of positive real numbers $\mathbb{R}^+$.

2.1.1. General Model

Objective function. The objective comprises the minimization of operational costs and investment costs:

$$\begin{aligned} Z = \quad & c^{\mathrm{b}} C^{\mathrm{b}} + c^{\mathrm{p}} C^{\mathrm{p}} \\ & + \sum_{t \in \mathcal{T}} P_t^{\mathrm{buy}} q_t^{\mathrm{g2c}} - P_t^{\mathrm{sell}} q_t^{\mathrm{c2g}} \end{aligned} \tag{1}$$

Here, c^{b} and c^{p} denote the installed capacities for battery storage and PV respectively at the investment costs C^{b} and C^{p}, which are annualized over the technology lifetime T at a discount rate r and the operation period of one year is set in relation to the technology lifetime, such that $C = rC_{\mathrm{norm}}(1+r)^T((1+r)^T - 1)^{-1}$, where C_{norm} denotes the norm investment costs. The sum of the operational costs over the set of time steps $\mathcal{T}$ included the costs associated with the flow from the grid to the community q_t^{g2c} at a price P_t^{buy}, and with the flow from the community to the grid q_t^{c2g} at a price P_t^{sell}.

Demand constraint. Electricity demand D_t needs to be satisfied by the flow from the grid to demand q_t^{g2d} as well as the flow from the PV units (battery) q_t^{p2d} (q_t^{b2d}) to demand:

$$q_t^{\mathrm{g2d}} + q_t^{\mathrm{p2d}} + q_t^{\mathrm{b2d}} = D_t \quad \forall t \in \mathcal{T} \tag{2}$$

PV generation. The product of the normalized electricity production from PV G_t^{p} and the installed capacity c^{p} sets an upper bound for the electricity generated from PV q_t^{p}, which is the sum of the flows to demand satisfaction q_t^{p2d}, to the grid q_t^{p2g} and to the battery storage q_t^{p2b}:

$$q_t^{\mathrm{p}} \leq c^{\mathrm{p}} G_t^{\mathrm{p}} \quad \forall t \in \mathcal{T} \tag{3a}$$

$$q_t^{\mathrm{p}} = q_t^{\mathrm{p2d}} + q_t^{\mathrm{p2g}} + q_t^{\mathrm{p2b}} \quad \forall t \in \mathcal{T} \tag{3b}$$

Battery storage. Battery storage is modelled by the following constraints. First, the storage level l_t in time step t is equal to the storage level in the preceding period adjusted by the charging (discharging) quantity q_t^{ch} (q_t^{dc}) at the respective efficiencies γ^{ch} (γ^{dc}):

$$l_t = l_{t-1} + \gamma^{\mathrm{ch}} q_t^{\mathrm{ch}} - \frac{1}{\gamma^{\mathrm{dc}}} q_t^{\mathrm{dc}} \quad \forall t \in \mathcal{T} \setminus \{0\} \tag{4a}$$

$$l_t = \gamma^{\mathrm{ch}} q_t^{\mathrm{ch}} - \frac{1}{\gamma^{\mathrm{dc}}} q_t^{\mathrm{dc}} \quad , t = 0 \tag{4b}$$

Also, battery outflow is the power sold to the grid (q_t^{b2g}) and used to supply demand q_t^{b2d}; battery inflow can originate from the grid (q_t^{g2b}) and PV units (q_t^{p2b}):

$$q_t^{\mathrm{ch}} = q_t^{\mathrm{p2b}} + q_t^{\mathrm{g2b}} \quad \forall t \in \mathcal{T} \tag{5a}$$

$$q_t^{\mathrm{dc}} = q_t^{\mathrm{b2d}} + q_t^{\mathrm{b2g}} \quad \forall t \in \mathcal{T} \tag{5b}$$

Besides that, the storage level is bound by the storage capacity and a minimum storage level $\underline{\lambda} c^{\mathrm{B}}$ (6a, 6b), whereas we assume $\underline{\lambda} = 0.1$. Besides, the battery cannot be discharged by more than half its capacity within one hour:

$$l_t \leq c^{\mathrm{b}} \quad \forall t \in \mathcal{T} \tag{6a}$$

$$l_t \geq \underline{\lambda} c^{\mathrm{b}} \quad \forall t \in \mathcal{T} \tag{6b}$$

$$q_t^{\mathrm{ch}} \leq \frac{1}{2} c^{\mathrm{b}} \quad \forall t \in \mathcal{T} \tag{6c}$$

Finally, the storage level in the initial and final period is set to be at least as high as the initial storage level:

$$l_0 \geq l_{|\mathcal{T}|} \tag{7}$$

Interaction with the central grid.. The EC can purchase and sell power from the central grid. Purchased power can fill the storage or contribute to supplying demand:

$$q_t^{\mathrm{g2c}} = q_t^{\mathrm{g2d}} + q_t^{\mathrm{g2b}} \tag{8}$$

Sales quantities consist of flows from PV and the battery storage:

$$q_t^{\mathrm{c2g}} = q_t^{\mathrm{p2g}} + q_t^{\mathrm{b2g}} \tag{9}$$

The general problem can then be written as:

$$\min \quad Z \tag{10a}$$

$$s.t. \quad \text{eqs. (2) to (9)}, \tag{10b}$$

whereas specific constraints are added for the respective strategies.

2.1.2. Strategy 1: Maximization of Economic Benefit

In the case of maximization of the economic benefit of the EC with the battery behind the meter, no additional constraints were needed. Power was purchased at a price that included taxes, fees and tariffs and was sold at day-ahead prices. The overview of price formation is given in Appendix A.1.

2.1.3. Strategy 2: Reducing peak power exchange with the central grid

This strategy aims at reducing the maximum power that the EC is exchanging with the grid. With that the aim of the EC would be to reduce its cost. E.g. the costs for the required grid connection or the grid tariff charge, which is associated with a customer's capacity usage. In certain countries, e.g. Norway, it is already today practice that some customers are charged based on their capacity usage and, in [25], it is stated that there is a push towards capacity-based (or power-based) tariffs.

Multi-objective optimization. Multi-objective optimization means solving an optimization problem comprising more than one objective. Only basic definitions of the field are given here, following [26] and [27], to which the reader is referred for an introduction to the topic.

In multi-objective optimization, the notion of optimality becomes the one of Pareto-optimality and since more than one Pareto-optimal solution can exist, there can be multiple "optimal" solutions. In a problem with the objective space $\mathcal{M}$, the aim is to find a solution $\mathbf{x}$ from the feasible region $\mathcal{S}$ that minimizes the vector of objective functions $\mathbf{f} \in \mathcal{R}^{|\mathcal{M}|}$ [27].

$$\min \quad \mathbf{f}(\mathbf{x}) \tag{11a}$$

$$s.t. \quad \mathbf{x} \in \mathcal{S} \tag{11b}$$

A solution $\mathbf{x}^*$ is said to be Pareto-optimal, if there does not exist another solution $\mathbf{x} \in \mathcal{S}$ that $\mathbf{x}^*$ is dominated by, in the sense that this solution is no worse than $\mathbf{x}^*$ with respect to all objectives and strictly superior to $\mathbf{x}^*$ with respect to at least one objective [27].

ϵ-constraint method. The ϵ-constraint method [28] is a method for exploring the search space of a multi-objective optimization problem. In each iteration, a problem of the following form is solved [27]: Minimize for one objective k, while constraining all other objectives $m \in \mathcal{M} \setminus \{k\}$ by upper bounds $\epsilon_1, \epsilon_2, ..., \epsilon_{|\mathcal{M}|}$:

$$\min \quad f_k(\mathbf{x}) \tag{12a}$$

$$s.t. \quad f_m(\mathbf{x}) \leq \epsilon_m \quad \forall m \in \mathcal{M} \setminus \{k\} \tag{12b}$$

$$\mathbf{x} \in \mathcal{S} \tag{12c}$$

Any solution to this problem is Pareto-optimal [26] and a variation of the bounds allows exploring the search space of the problem. Here, the objectives are the minimization of costs and the minimization of peak consumption from the central grid. Peak consumption α can be written as the maximum inflow from the grid to the community:

$$\alpha \geq q_t^{\mathrm{g2c}} \quad \forall t \in \mathcal{T} \tag{13}$$

The bi-objective problem of minimising peak power exchange from the central grid α and costs Z can be written as (14). Hence, we refer to peak power exchange as the maximum of the flow both to and from the grid in the remainder of this paper.

$$\min \quad [Z, \alpha] \tag{14a}$$
$$s.t. \quad \text{eqs. (2) to (9) and (13)}, \tag{14b}$$

Here, the ϵ-constraint method is applied to solve that problem (14). That means that the set of objectives $\mathcal{M}$, introduced in (12) becomes $\mathcal{M} = \{Z, \alpha\}$. The linear program (eqs. (2) to (9)) is solved minimizing for costs Z only, with costs corresponding to objective k with objective function f_k in (12), whereas peak consumption α is successively constraint, corresponding to $\{f_m | m \in \mathcal{M} \setminus \{k\}\}$.

20 iterations κ are run, where an upper bound on α is set, that is linearly decreasing with respect to the peak power exchange α_0, when maximising economic benefit only. In other words, α_0 is obtained from the solution to the single-objective problem of cost-minimization and then, the model is solved 20 times, while forcing $\alpha \leq \alpha_0, \alpha \leq 0.95\alpha_0, \alpha \leq 0.9\alpha_0, ..., \alpha \leq 0.05\alpha_0$.

That corresponds to writing the problem in iteration κ as:

$$\min \quad Z \tag{15a}$$
$$s.t. \quad \alpha \leq (1 - \frac{\kappa - 1}{20})\alpha_0 \tag{15b}$$
$$\text{eqs. (2) to (9) and (13)}, \tag{15c}$$

Please note that in general, in order to obtain a sensible range for the ϵ-constraints, it is recommended to first optimize for all objective values (see e.g. [29]). In our case, we already know that the optimization for peak power exchange would lead to $\alpha^* = 0$, hence the range for α is known a-priori.

2.1.4. Strategy 3a: Maximization of Self-Sufficiency - Battery Behind the Meter

When the energy community's self-consumption is maximized with the battery behind the meter, all flows from the community to the central grid were set to zero:

$$q_t^{\mathrm{c2g}} = 0 \quad \forall t \in \mathcal{T} \tag{16}$$

2.1.5. Strategy 3b: Maximization of Self-Sufficiency - Battery in Front of the Meter

If self-consumption is maximized with the battery placed in front of the meter, additionally to (16), grid tariffs, taxes and fees P^{fix} needed to be paid when charging and discharging the battery, so the objective function changed:

$$\begin{aligned} Z = \quad & c^{\text{b}} Q^{\text{b}} + c^{\text{p}} Q^{\text{p}} \\ & + \sum_{t \in \mathcal{T}} P_t^{\text{buy}} q_t^{\text{g2c}} - P_t^{\text{sell}} q_t^{\text{c2g}} \\ & + \sum_{t \in \mathcal{T}} (q^{\text{p2b}} + q^{\text{b2d}}) P^{\text{fix}} \end{aligned} \tag{17a}$$

$$q_t^{\text{c2g}} = 0 \qquad \forall t \in \mathcal{T} \tag{17b}$$

2.2. Grid impact analysis

The impact of ECs on the distribution grid was investigated based on power flow simulations in three test grids. The open-source python tool pandapower [30] was used to carry out the grid impact analysis.

2.2.1. Test distribution grids

Depending on the supplied customers and the area, distribution grid design criteria differ [31]. Three distribution grids are analysed in the paper, namely a city, a suburban and a village grid. In order to investigate EC configurations, where the EC members are connected to two different LV grids, it is necessary to include the MV grid in the power flow simulations. For simplicity, in this work the same MV grid was assumed for each of the above mentioned areas, namely the CIGRE MV distribution grid described in [32]. In order to investigate the impact of ECs, the following modifications were made to the MV grid.

- At bus 5, an MV/LV transformer was added supplying a commercial customer through LV cables.
- At bus 6 and 7, the MV loads were each replaced by an LV distribution grid (city, suburban and village, respectively).
- The LV grid connected to bus 7 was extended with an additional LV feeder, which is supplying a commercial customer

An example of the MV grid extended with two suburban LV grids can be seen in Fig. 1. The LV grids used in this work can be found in [33] and are representative for LV grids in Germany. The different grid types are briefly described in the following.

Village. This LV grid is representative for a grid in a village in Germany (called "Dorfnetz" in [33]). The MV/LV transformer is a 400 kVA 20/0.4 kV transformer with six short LV feeders and 57 loads, which each corresponds to one consumer considered to be detached houses.

Suburban. This LV grid is representative for grids in German suburban areas (called "Vorstadtnetz" in [33]), which are often located around cities or urban areas. They are characterized by a high density of buildings/consumers and the consumer type are households in detached, semi-detached or terraced houses. The MV/LV transformer is a 630 kVA 20/0.4 kV transformer with nine long LV feeders and 144 loads, which each corresponds to one consumer.

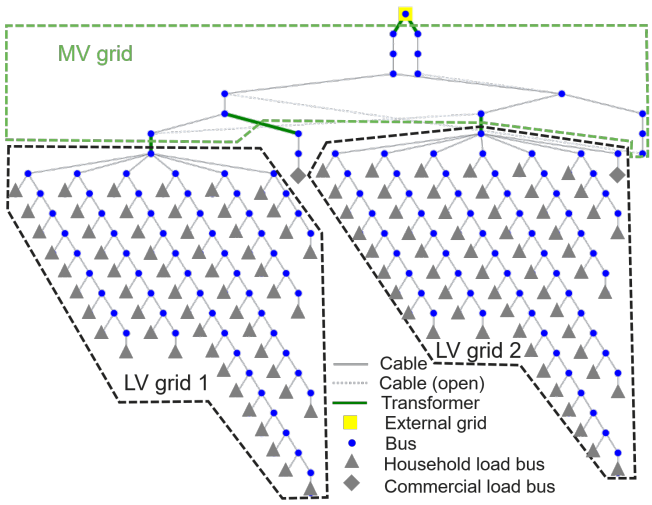


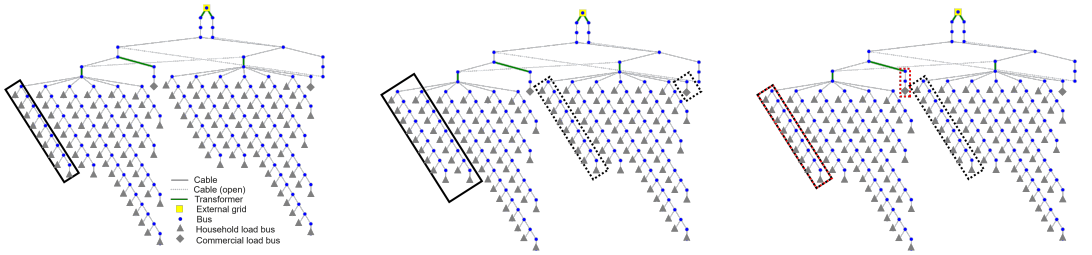
Figure 1: Single-line diagram of the CIGRE medium voltage distribution grid [32] extended with two suburban LV grids.

City. In city centers, the customer concentration is high and the LV feeders are short. The grid topology is similar to the village grid, but the supplied customers are generally households, mainly flats in apartment buildings and few semi-/detached houses. There is no urban or city grid available in [33]. Since the grid topology is similar to the village grid, that grid topology was used, but the consumption patterns correspond to typical consumers in city centers. The MV/LV transformer is a 400 kVA 20/0.4 kV transformer with six LV feeders and 57 loads. In a residential area of an urban grid, the consumers are a mix of flats and detached houses. In order to determine how many consumers are located in the city grid, the following assumptions were made. An average number of flats per apartment building were calculated based on data for the capital region of Denmark, which is dominantly urban area. Based on data from Statistics Denmark[1] for the capital region, there were ca. $31,000$ apartment buildings with in total approx. $552,000$ flats in 2019. Consequently, there were in average ca. 18 flats in an apartment building. Moreover, housing statistic for the region show that approx. 95 % of the homes are flats and the remaining 5 % are houses. Based on these assumption, there are a total of 555 consumers in the city grid, which consist of 527 flats and 28 houses.

2.2.2. Consumption profiles

The consumption profiles have a resolution of one hour and were extracted from real measurement data. The database is an anonymized data set comprising one year of consumption data for approx. $300,000$ customers in Denmark, which were categorized into consumer categories. The categories allow identifying different household consumer types, e.g. detached house without/with electric heating, detached house with heat pump, apartment with/without electric heating etc., and different industrial/commercial consumers, e.g. supermarkets. The resulting consumption profiles differentiate between season (winter/summer) and day (weekday/weekend). Moreover, for all

[1] www.dst.dk



(a) $EC1$

(b) $EC2a$ (solid rectangle) and $EC2b$ (dashed rectangle)

(c) $EC3a$ (black dashed rectangle) and $EC3b$ (red dashed rectangle)

Figure 2: The five different energy community configurations, which are investigated in the paper.

profiles a 95 % quantile was used and the profiles take into account simultaneity, when a feeder with multiple consumers is simulated.

Profiles for the following consumer types where used in the simulations.

- Household loads: Detached houses without electric heating, detached houses with heat pump and apartments without electric heating
- Commercial load: Supermarket

The used PV and battery profiles are the output from the optimization described in Section 2.1.

In the analysis, two summer weeks followed by two winter weeks will be simulated.

2.2.3. Energy community configurations

From a grid perspective, a variety of EC configurations are possible, where configuration refers to how the members are connected to the distribution grid. In the following, three configurations, where two configurations have two subcategories each, are considered. The subcategories are added, due to the fact that the EC can be composed of different customer types, e.g. households and commercial customers. In Fig. 2 the different configurations are indicated in the one-line diagram. The rectangles indicate which buses and, hence, consumers are part of the EC.

Below the characteristics of the five EC configurations are described.

1. All members of the energy community are connected to the same LV feeder
   (a) $EC1$: Only households (see Fig. 2a)
2. All members are connected to the same MV/LV transformer, but to different LV feeders (see Fig. 2b)
   (a) $EC2a$: Only households (solid rectangle)
   (b) $EC2b$: Households and one large commercial customer (dashed rectangle)
3. The members are connected to two different MV/LV transformers (see Fig. 2c)
   (a) $EC3a$: Only households (black dashed rectangle)
   (b) $EC3b$: Households and one large commercial customer (red dashed rectangle)

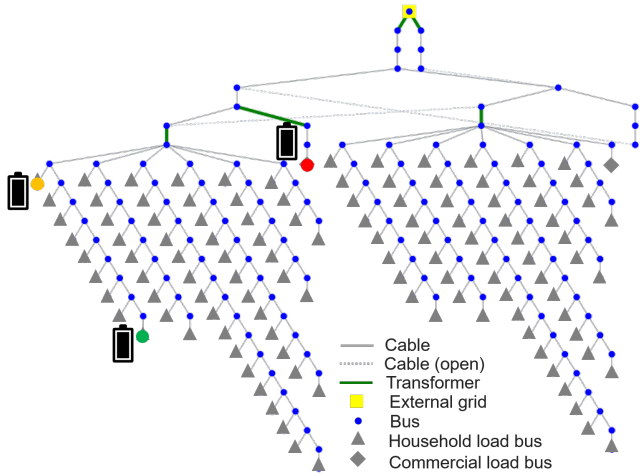


Figure 3: Example for the investigated battery locations.

2.2.4. Approach for grid impact assessment

The grid impact is investigated in three steps. In each step, power flow computations are performed and their results are analysed. Each power flow calculation simulates two summer weeks and two winter weeks, in total four weeks. In order to evaluate the grid impact, minimum and maximum voltage magnitudes observed within the EC in the simulated four weeks were extracted, as well as maximum LV line, MV line and MV/LV transformer loading.

Step 1 - Impact of battery placement. In order to investigate the impact of battery placement, the profiles from Strategy 1 are used and the communal battery is connected at the following three locations (see also Fig. 3):

- Beginning of the feeder (yellow): It is expected that the impact of the battery is limited, since it is located close to the MV/LV transformer and, hence, the electrical distance to the remaining grid is relatively short.
- End of the feeder (green): In this case, the battery presumably has a larger impact on the grid, since the electrical distance to the remaining grid is long in comparison and power flowing from the MV grid to the battery results in a voltage drop along the entire feeder.
- Commercial customer (red): The commercial customer has its own feeder or even its own MV/LV transformer, which may limit the impact on the grid.

The impact is evaluated by investigating minimum/maximum voltage magnitude at the buses within the energy community, maximum loading of LV cables, which interconnect the members of the EC and connect them to the grid, and MV/LV transformer loading. Based on these results, from a grid perspective a preferred location for the communal battery is selected. This location is used for all simulations in the second and third step.

Step 2 - Impact of peak-shaving. The methodology described in Section 2.1.3 was used to calculate the needed PV and battery capacity as well as the corresponding charging profile to gradually reduce the maximum power

consumption P_{max} of the energy community in the base case (Strategy 1, see Section 2.1.2). The maximum power consumption of the EC P_{max} is gradually reduced by enforcing a power cap, which is expressed relative to the maximum power consumption of the EC $P_{max,0}$, when no cap was applied. The cap is reduced in steps of 5%. Hence, in the first peak-shaving step P_{max} is limited to $95\%P_{max,0}$. In total 20 steps are simulated, which corresponds to a cap equal to $5\%P_{max,0}$ in the last step. Based on an economical and grid impact analysis, a feasible level for peak-shaving is selected.

Step 3 - Comparison of the impact of different battery operation strategies. In this step, the different strategies derived in the section 2.1.2 to section 2.1.5 are compared based on their grid impact and economical factors. In all strategies it is assumed that the battery is connected at the preferred location determined in *Step* 1. For Strategy 2 the peak-shaving level, which was chosen based on the analysis carried out in *Step* 2, will be used.

3. Results

To assess the impact of the different ECs on the distribution grids, first the LP optimization model was used to determine the optimal PV and communal battery capacities. These results are presented in Sec. 3.1. Subsequently, the impact of the ECs on the different grid types is investigated and discussed in Sec. 3.2.

3.1. Optimal capacity and operational results

The sizing of the EC's PVs and battery are not only dependent on the configuration of the EC but also on the operational strategy. In this section, first the capacities for the operation strategies 1 and 3*a*&*b* are discussed and then for Strategy 2, where the amount of peak-shaving was gradually increased.

3.1.1. Strategy 1 and Strategy 3a&b

The average electricity price for households was 0.282 EUR/kWh, while the capacity factor of PV was 12 %. More details about the taxes and the exact price formation is presented in Appendix A.1. Calculated yearly annuities for investments were 124.9 EUR/kW of PV capacity and 41.76 EUR/kWh of the communal battery capacity. The optimal capacities for different grid set-ups of EC can be seen in Fig. 4. Due to different consumption profiles, optimal installed capacities in the city grid were generally larger than in the other two types of grid. Moreover, the set-ups that included a large consumption point (a supermarket) had larger optimal capacities of PV and BESS, which can be seen in cases EC2*b* and EC3*b*. For Strategy 3*a* for the case of the city grid, the optimal PV capacities were lower between 25 % and 31 %, and the optimal communal battery capacities were lower between 24 % and 28 % than in Strategy 1. The reason for this is that Strategy 1 maximizes the economic benefit, which includes also market arbitrage by using the battery. A larger battery then also provides larger value for the PVs, as energy generated by PV can be stored for longer, until the more beneficial prices are achieved. On the other hand, Strategy 3*a* optimizes the self sufficiency and does not allow selling the energy community's electricity to the main grid. This reduces the value of PV generated and the batteries, as they cannot utilize high market prices in the most economic way. Furthermore, for Strategy 3b (battery located in front of the meter) for the case of city grid, a communal battery was not installed for any of the configurations and the optimal PV capacity was between

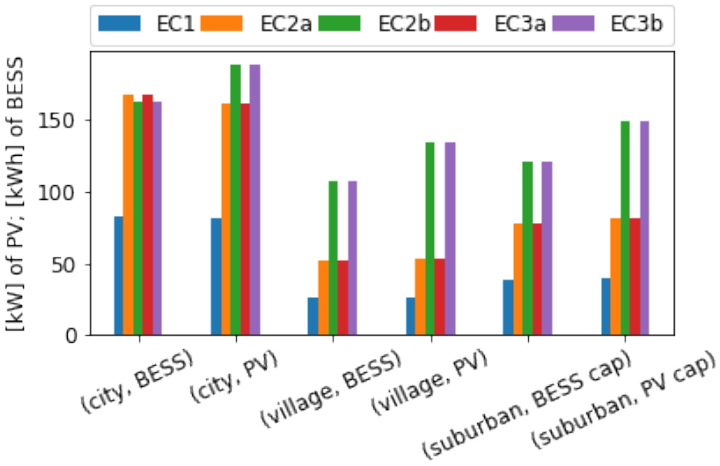


Figure 4: Strategy 1: Installed capacities of PV and BESS for the different grid set-ups investigated in this work

52 % and 57 % lower than in Strategy 1. This shows that the regulatory framework regarding the taxation of the communal battery operation is very important for the feasibility of the investment. Optimal capacities of PV and BESS, that were obtained by focusing on different possible strategies of EC, were used for the analysis of EC's impact on the power grid.

However, cases EC1-EC3*b* all had different total loads. Thus, in order to evaluate the average cost of acquiring electricity, the levelized cost of acquired electricity (LCOE) expressed in EUR/kWh of electricity is a useful measure. In Table 2, average electricity prices are presented for 45 different combinations of strategies and grid set-ups. The lowest LCOE occurred for the village grid for Strategy 1, for the EC3*b* setup. Contrary to that, the highest LCOE occurred in the case of Strategy 3, in village and suburban grids. The difference between the highest and the lowest LCOE for the EC was 6.4 %.

Moreover, the operation of one EC that follows Strategy 1 for the city grid can be seen in Fig. 5. The demand during the three weekdays followed a relatively uniform pattern and the peak occurred between $17:00-19:00$ in all three days. PV generation in the summer was especially high on the 6^{th} of July. The peak PV generation on that day was three times larger than the EC consumption. Hence, one part of the energy was stored in the battery and the larger share was exported to the central grid.

In general, the battery was utilized more in the winter time than in the summer time, as it could arbitrage better during the winter week when the price oscillations on the market were larger. In all three winter days, the peak demand hour was completely satisfied by the stored energy in the batteries, as the peaks in demand and day-ahead electricity prices coincided. During the summer week, the batteries were usually charged during the periods of high PV generation and they were usually discharged during the early evening. In the winter week, batteries were charged during the night (starting after midnight), using the power from the central grid. This analysis shows

Table 2: Total levelized cost of acquired electricity (LCOE) [EUR/kWh] in different Strategies and different grid set-ups. The total levelized cost included both the annualized investment costs, as well as costs of electricity taken from the grid

		EC1	EC2*a*	EC2*b*	EC3*a*	EC3*b*
City	Strategy 1	0.258	0.258	0.251	0.258	0.251
	Strategy 3*a*	0.262	0.262	0.256	0.262	0.256
	Strategy 3*b*	0.263	0.263	0.258	0.263	0.258
Village	Strategy 1	0.260	0.260	0.250	0.260	0.250
	Strategy 3*a*	0.264	0.264	0.254	0.264	0.254
	Strategy 3*b*	0.267	0.267	0.256	0.267	0.256
Suburban	Strategy 1	0.260	0.260	0.251	0.260	0.251
	Strategy 3*a*	0.267	0.267	0.256	0.267	0.256
	Strategy 3*b*	0.267	0.267	0.258	0.267	0.258

that the day-ahead prices in the summer were negatively correlated with the peak PV consumption, which made it optimal for the batteries to charge it during those low-cost periods. However, in the winter week the day ahead prices were lowest during the night and thus, the batteries were mostly charged over the night time. The battery profiles were obtained for all the different combinations and they were further utilized as an input for the grid analysis.

3.1.2. Strategy 2

Results of the optimization for Strategy 2 in terms of LCOE and investment capacities, the latter relative to Strategy 1 results, are plotted against the enforced reduction in peak power exchange with the main grid in Figure 6. In order to generate a one-dimensional KPI for the change in investment behaviour with respect to the baseline case of no constraints on peak demand, investment capacities for PV and BESS were normalised with respect to the baseline case - i.e. 0 % corresponding to the same value as in the baseline. Then, an average value across normalized investment capacities for the two capacities was calculated.

Following the curves from right to left with increasingly strict constraints on the optimization, one can note that until the elbow points, reductions in the peak power exchange had only a minor effect on resulting system costs and investment behaviour. The reason for this behaviour was that stricter additional constraints led to a shift in the EC's system operation with respect to the timing of electricity sales and purchases. This behaviour was a result of finding strictly optimal solutions meaning that due to power prices varying slightly across hours, the LP solution shifted as much purchase/sales volume as possible to the most favourable hours, even though a flatter curve would only yield a minor deterioration in the objective value.

The values left to the elbow points indicate that stricter constraints on the peak power exchange forced investments in additional BESS and PV capacities. At these points, mere operational changes were not sufficient to fulfill the constraints on peak power exchange. As a result, a sharp increase in system costs was observable.

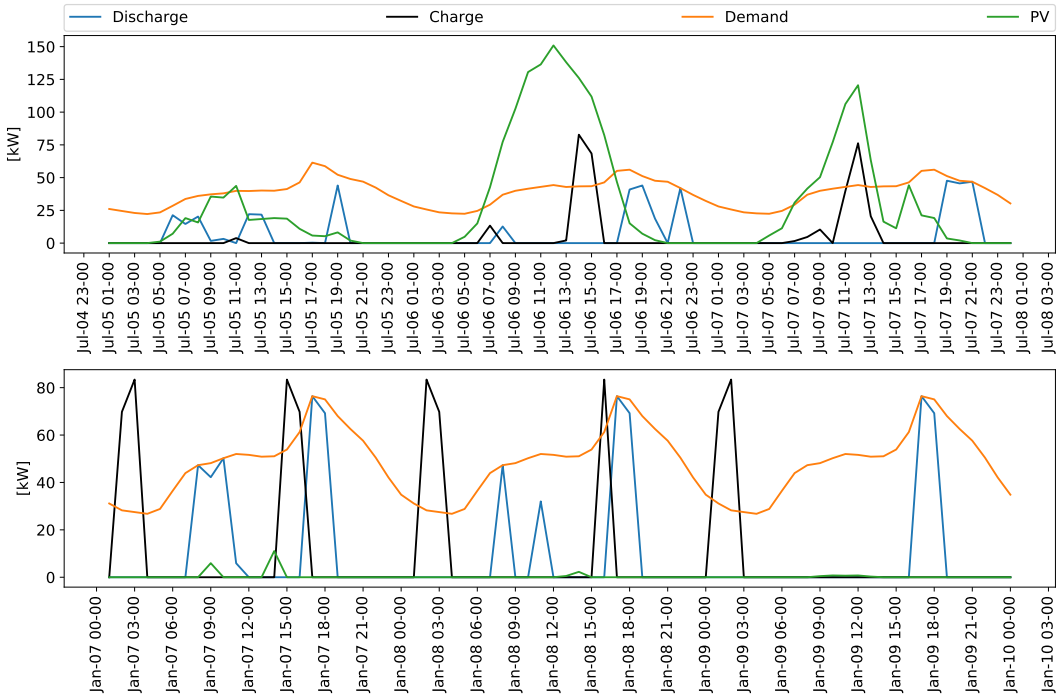


Figure 5: Operation of the EC during three days for a winter (bottom) and a summer (top) week. The figure includes BESS operation, PV generation and the total electricity demand of the community. The presented operation is for the city grid, Strategy 1 (maximizing of EC's own economic benefit) for the EC3a configuration. Presented profiles in the figure correspond with the summer and the winter week that was used for the power grid analysis.

3.2. Grid impact analysis results

The analysis is split into three parts. First, the impact of the battery location was investigated for the three grid types and for the five EC configurations. This allowed to identify favourable locations for the communal battery. Second, the results from Strategy 2 were investigated in the City grid and for the different levels of peak-shaving. Finally, the three strategies were compared to identify their differences with respect to grid impact.

3.2.1. Impact of battery location

The location, where the battery is connected, can have a significant impact on the grid.

Figure 7 shows the observed minimum and maximum voltages observed for each grid and EC configuration. The bars stretch from the observed minimum to the maximum value in the power flow simulation and, consequently, all voltages lay within the depicted range. In the graph, the blue bars show the voltage range, when the simulation was carried out without a battery. The yellow bars correspond to the case, when the battery was located at the beginning of the feeder, the green bars to the case, when the battery was connected at the end of the feeder, and

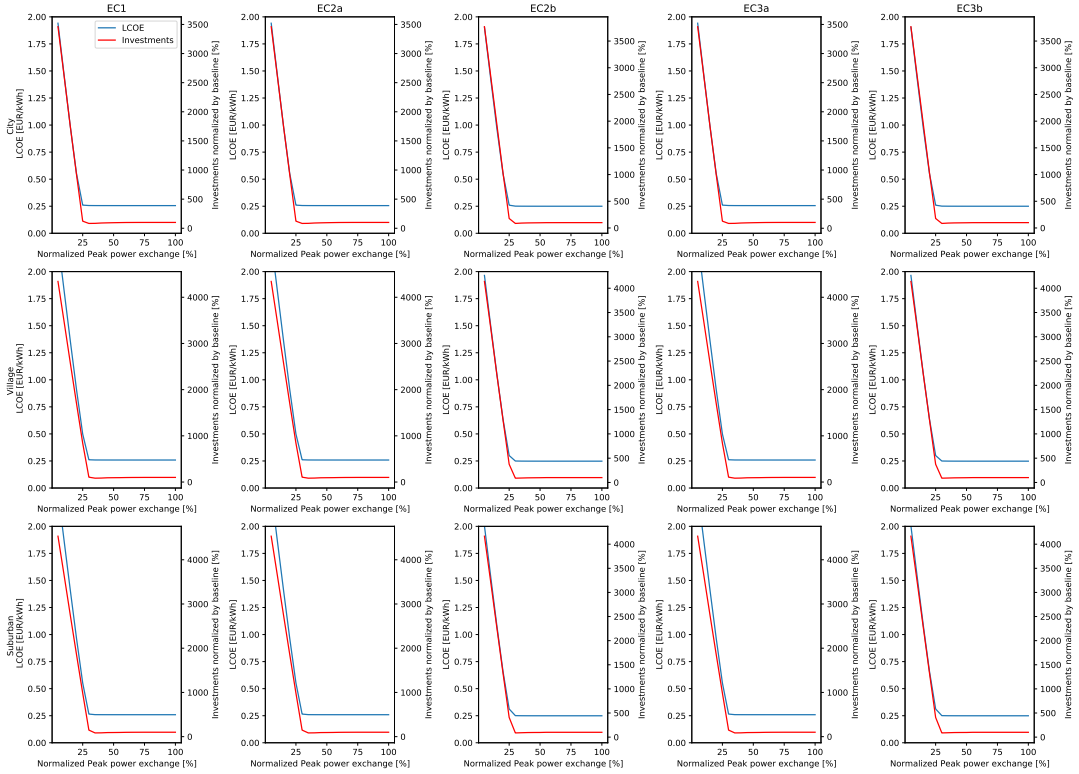


Figure 6: LCOE and investments for Strategy 2 against peak power exchange with the main grid. Peak power exchange is given as a percentage of Strategy 1 values and investments as a change with respect to Strategy 1 values as a baseline.

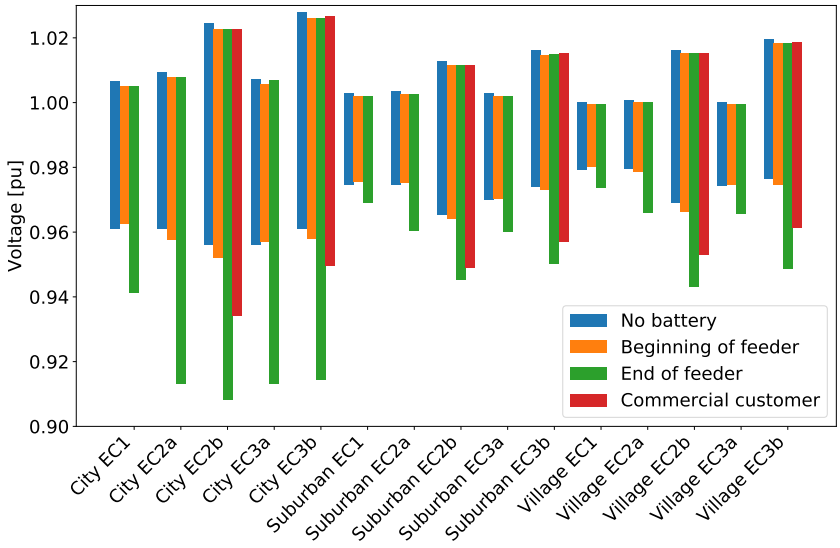


Figure 7: Comparison of the impact of different investigated battery locations on the minimum and maximum voltage observed in the EC.

the red bars, when the battery was connected at the location of the commercial customer (only applicable for EC2*b* and EC3*b*).

The grid operator aims at keeping the voltages within a narrow band. According to the European Standard EN50160 [34], the voltage should be maintained within a band ±10 % of its nominal value. It could be observed that the battery location greatly impacts the voltage. In general, it could be observed that, when the battery was located at the end of the feeder or the commercial customer, the impact on the voltage was detrimental and the voltage range increased. Connecting the battery at the end of the feeder led to significantly lower voltages, e.g. in the case City EC2*b* the lowest voltage was reached with 0.91 pu, while without battery the minimum voltage was 0.96 pu. This was due to the fact that the additional power flow to charge the battery increased the voltage drop along the feeder. If the battery was located at the beginning of the feeder (yellow), then the impact was limited. In some cases the battery improved the voltages by decreasing the observed range of voltages, mainly by increasing the minimum voltage. For example in City EC1, City EC3*a*, Suburban EC1 and Suburban EC2*a* the minimum voltage with battery connected to the beginning of the feeder was slightly improved. With respect to EC configurations, it could be observed that across all grid types in the cases where the EC configurations included a commercial customers (namely EC2*b* and EC3*b*), then the impact on voltage was more pronounced. This may be explained by the fact that the consumption of the commercial customer was large in comparison to residential consumers and, hence, the power exchanged with the battery was larger. A comparison of the different grid types revealed that the voltage in the City grid was generally more impacted. The consumer concentration was higher in that grid type (see Section 2.2.1) and, consequently, also its loading. This can also be observed in Fig. 8, where the maximum loading of the most loaded LV line is shown in each grid type, for each EC configuration and the different

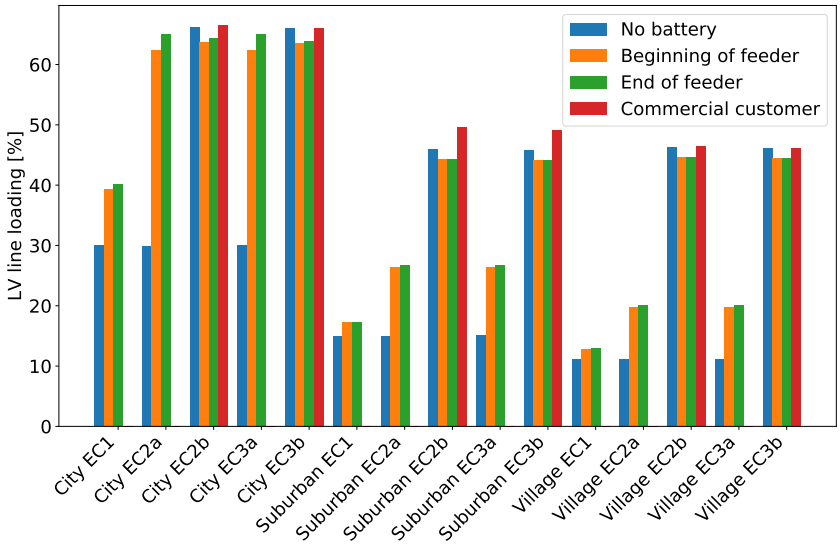


Figure 8: Comparison of the impact of different battery locations on the loading of the LV lines in the EC.

battery locations. The figure shows that the impact of the battery location on the LV line loading was limited. Across grid types and with the configuration EC1, EC2*a* and EC3*a*, the introduction of a battery led to increase of LV line loading. Due to the economical optimization, the battery was charged during hours where the electricity price was low and discharged during hours where the price was high. This created a new consumption peak during the night, where electricity prices were low (see Fig. 5). The new peak correlated with the number of consumers in the EC. When the EC only consisted of consumers connected to a single feeder (EC1), then the impact was considerably lower than in the case when the members spread over two feeders (EC2*a* and EC3*a*). When the EC included a commercial customer (EC2*b* and EC3*b*), then the installed PV capacity was larger (see Fig. 4). This resulted in a higher loading of the grid, when no battery was included (blue bars in Fig. 8) and a slight reduction of the maximum LV line loading, when the battery was connected at the beginning or end of the feeder.

Figure 9 shows the impact of the battery and its location on the transformer loading for the different grid types and EC configurations. It can be observed that the transformer loading was less impacted in comparison to the LV lines. In a few cases (e.g. City/Suburban/Village EC3*a*) the battery had virtually no impact on the highest loaded MV/LV transformer. In cases City/Suburban/Village EC1 as well as Village EC2*a* the battery appeared to slightly decrease the loading. When a commercial customers was part of the EC (EC2*b* and EC3*b*), the battery generally slightly increased the transformer loading.

Lastly, the impact on the loading of the MV grid, in particular, MV lines was investigated. Figure 10 displays the highest loading of the highest loaded MV line in the different grid types and EC configurations. The graph reveals that the battery and its location did not have a significant impact on the MV line loading. This may be explained by the fact that the power consumption of the ECs were small compared to the total power delivered

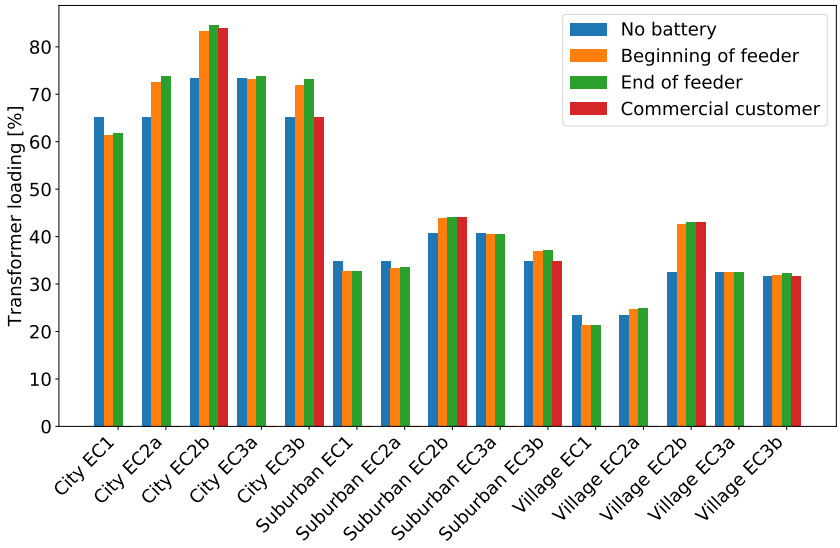


Figure 9: Comparison of the impact of different battery locations on the loading of the MV/LV transformers in the EC.

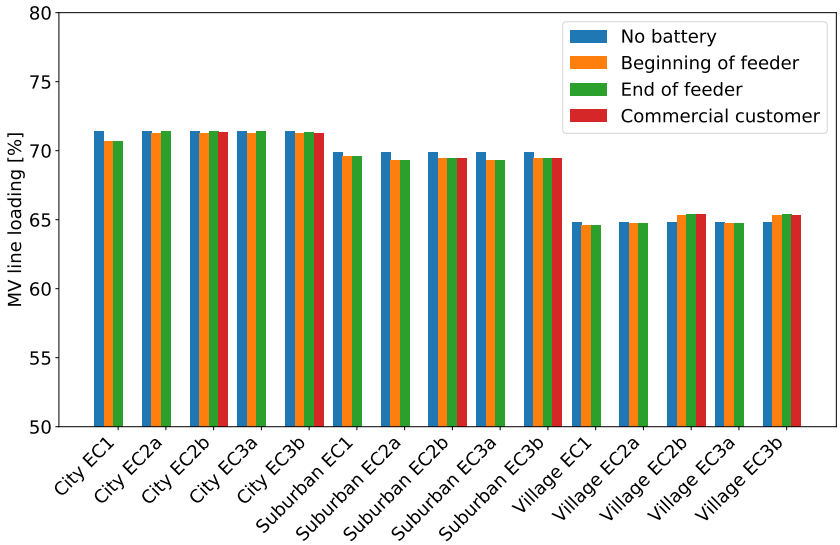


Figure 10: Comparison of the impact of different battery locations on the loading of the MV lines in the EC.

through the MV grid. If multiple ECs would be considered, they could potentially impact the MV grid. However, this was out of the scope of the paper.

The results presented in this section revealed that the battery location greatly affects the minimum and maximum

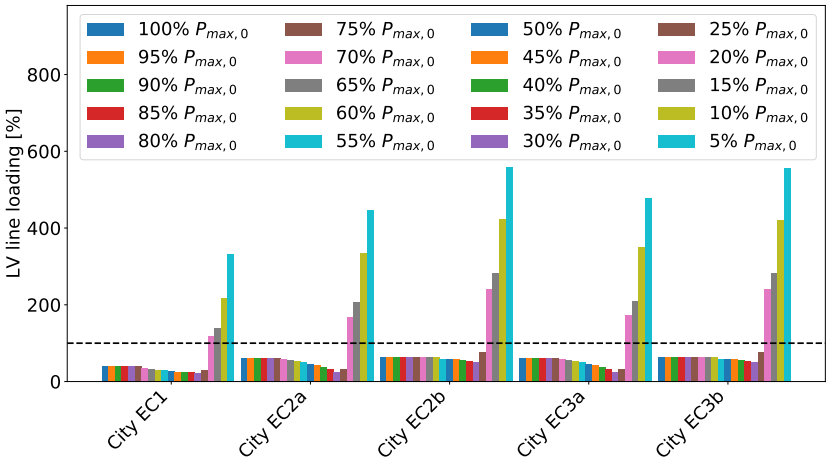


Figure 11: Comparison of the impact of peak-shaving on the loading of the LV lines in the EC. 100% loading is indicated by the black dashed line.

voltages in the LV grids. When the battery was located at the end of the feeder or at the commercial customer, the impact on the voltage was detrimental, while the battery did not affect the voltage, when it was located at the beginning of the feeder. Loading of LV lines, MV lines and transformers was not significantly impacted by the battery location. Based on these results, the preferred location for connecting the communal battery seems to be at the beginning of the feeder.

3.2.2. Impact of consumption peak-shaving

Due to the results in Section 3.2.1 and aiming at reducing the impact of the battery on the grid, in this analysis it was assumed that the battery was connected at the beginning of the feeder (see Fig. 3). Moreover, only results for simulation of the City grid will be discussed, since in the previous section, it was identified as the grid type most impacted by the EC.

Figure 11 shows the impact on LV line loading, when the amount of peak-shaving was gradually increased. It can be observed that initially peak-shaving was reducing the loading of the LV lines until a cap of 30%$P_{max,0}$. Further reducing the cap resulted in a massive increase of installed PV capacity and battery capacity (see Section 3.1.2), which led to overloading of the LV lines. This shows that a peak shaving of 70% is both economic and technical maximum, and that further peak shaving has only deleterious effect on the system. Future research applied on different case studies should show whether a similar break-off point exists for energy communities in general. Subsequently, only the peak-shaving cases up to 30%$P_{max,0}$ were further analysed, since the aim with this strategy was to relieve the grid.

Figure 12 shows the impact of peak-shaving up to 70 % on the line loading in the LV grid. It can be observed that in all EC configurations the peak-shaving strategy reduced the loading of the LV lines to some extent. For EC1 the loading of the highest loaded LV line was reduced from ca. 39 % to ca. 23 %. In configuration EC2*b* and

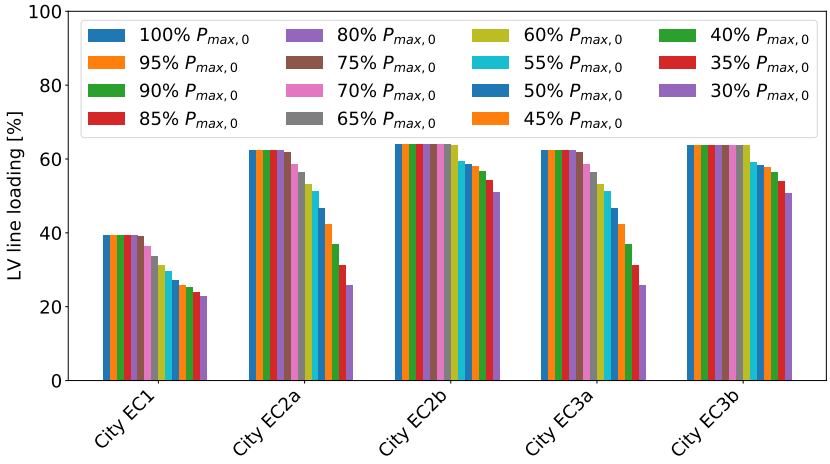


Figure 12: Loading of the LV lines in the EC with up to 70 % peak-shaving.

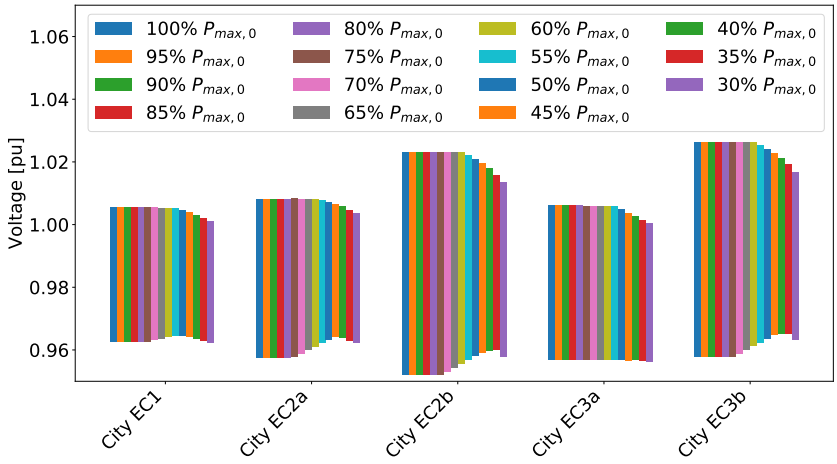


Figure 13: Impact on voltage in the EC with up to 70% peak-shaving.

EC3*b* a very similar reduction could be observed from ca. 64 % to 51 %. For EC2*a* and EC3*a* the reduction were the same and were the largest compared to the other EC configurations. The loading dropped from 62 % to 26 %.

Figure 13 shows the observed minimum and maximum voltage in the City grid, when applying peak-shaving up to 70 %. The graph shows that the impact on the voltage was very limited. In general more than 30 % peak-shaving was required to notice any change. For further increase in peak-shaving in EC2*a*, EC2*b* and EC3*b* a reduction of the voltage range could be observed. In comparison to the impact of battery location on the voltage, the impact of the peak-shaving strategy was negligible.

The influence of the peak-shaving strategy on transformer loading is depicted in Fig. 14. For EC3*a* no impact

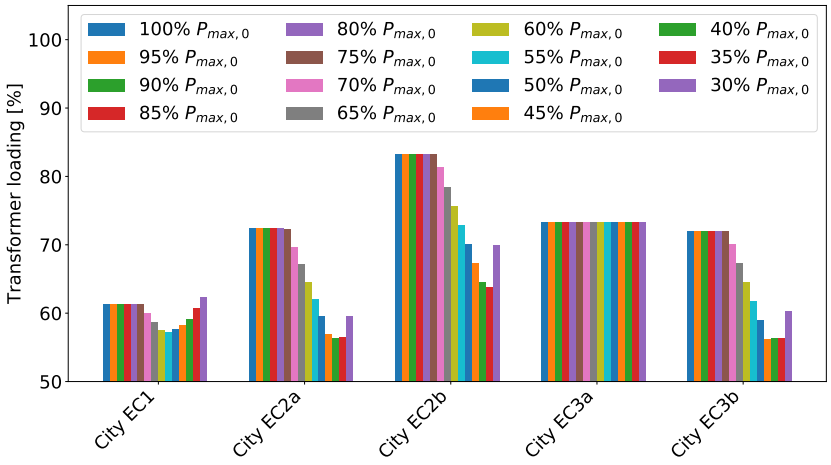


Figure 14: Transformer loading in the EC with up to 70% peak-shaving.

could be observed. The reason was that the highest loaded MV/LV transformer was the one supplying LV grid 2 (see Fig. 1 and Fig. 2c). Additionally to the household loads in the LV grid, this transformer was providing one commercial customer, which resulted in the higher loading. In this case, the battery was connected in LV grid 1 and, hence, the battery could not reduce the loading of that transformer located in LV grid 2.

For EC1, it could be observed that the loading remained constant around 61 % until 75%$P_{max,0}$. Afterwards the loading gradually decreased to 57 % (55%$P_{max,0}$). This reduction may be due to the EC covering more of their demand with power provided by the communal battery and PVs. This, consequently, reduced the demand of the EC from the MV grid. Beyond 55%$P_{max,0}$, the EC was again importing more power from the MV grid, which was visible in an increasing loading of the transformer. Similar behaviour could be observed in the configurations EC2a, EC2b and EC3b. The largest reduction in loading could be observed for EC2b, where the loading dropped from 83 % to 64 %.

Figure 15 shows for the different EC configurations the LCOE for the EC and maximum LV line loading for all peak-shaving levels, where the LV line loading was below 120 %. On the x-axis the load-shaving levels are shown. The blue curve indicates the costs for the EC on the left y-axis, while the red curve shows the maximum LV line loading on the right y-axis. The graph reveals that the impact on the total costs for the EC community was generally relatively small up to the peak-shaving level, where the LV line loading again began to increase. For example, in EC2a, where the largest reduction in loading could be observed, the total cost increased with ca. 0.3 % at the peak-shaving level 30% $P_{max,0}$, where the LV line loading reached it's minimum of ca. 26 %. This raised the question of finding a point which is optimal from the perspective of a DSO (the optimum from the EC perspective being the base case 100% $P_{max,0}$). Table 3 shows the maximum reduction in line loading and the corresponding increase in costs (both in percentage). While the reductions in maximum line loading were between 20.51 % and and 58.51 %, additional costs stayed below 0.4 % for these points.

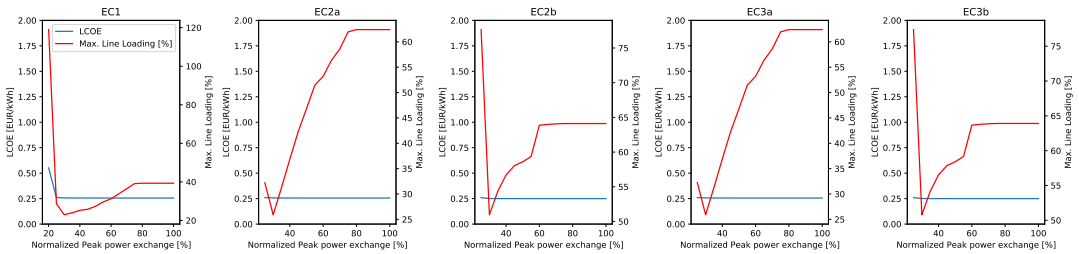


Figure 15: Comparison of impact of peak-shaving on LV line loading and the EC's LCOE in the City grid. Peak power exchange is given as a percentage of Strategy 1 values.

Table 3: Highest reduction in maximum line loading and corresponding cost increase expressed in percentage relative to Strategy 1 for the City grid.

	EC1	EC2*a*	EC2*b*	EC3*a*	EC3*b*
Line loading reduction	41.68	58.51	20.51	58.51	20.53
Cost increase	0.31	0.31	0.36	0.31	0.36

3.2.3. Impact comparison of the different strategies

The impact on the grid depends on the battery operation strategy. In Strategy 3*b* where the battery was located in front of the meter, the optimization resulted in no battery being installed. Since the impact of the battery was in this section the main interest, Strategy 3*b* was not investigated. Strategy 3*a* will in the following only be referred to as Strategy 3. A comparison of the three remaining strategies (described in Section 2.1) was presented.

For all strategies, the battery was located at the beginning of the feeder based on the results in Section 3.2.1 and aiming at reducing the impact on the grid.

For Strategy 2, the costs for EC were generally only minor impacted by increasing the peak-shaving level until the LV line loading reached its minimum. In Section 3.2.2 this was shown for the City grid and similar behaviour could be observed in the other two grid types (see Fig. A.19 in Appendix A.2). Based on these observation, in the comparison of the strategies in this section the peak-shaving levels were chosen, which achieved the lowest maximum loading of the LV lines (see Table A.4 in Appendix A.2).

In order to assess the different strategies, their impact on the minimum and maximum voltage, LV/MV line loading and transformer loading were investigated.

Figure 16 shows the minimum and maximum voltages observed within the EC. The graph shows that a smaller voltage range could be observed for Strategy 2 (peak-shaving) and Strategy 3 (self-sufficiency) compared to Strategy 1 (economical optimisation). Although Strategy 3 generally resulted in a narrower voltage range, it should be noted that peak-shaving in several cases improved the minimum voltage slightly more than the other two strategies. Raising the minimum voltage decreases losses and hence, makes grid operation more efficient. Consequently, raising the minimum voltage may be of greater importance compared to having a narrower voltage band. However, it should be noted that battery location seemed to have the greatest impact on voltage (see Section 3.2.1).

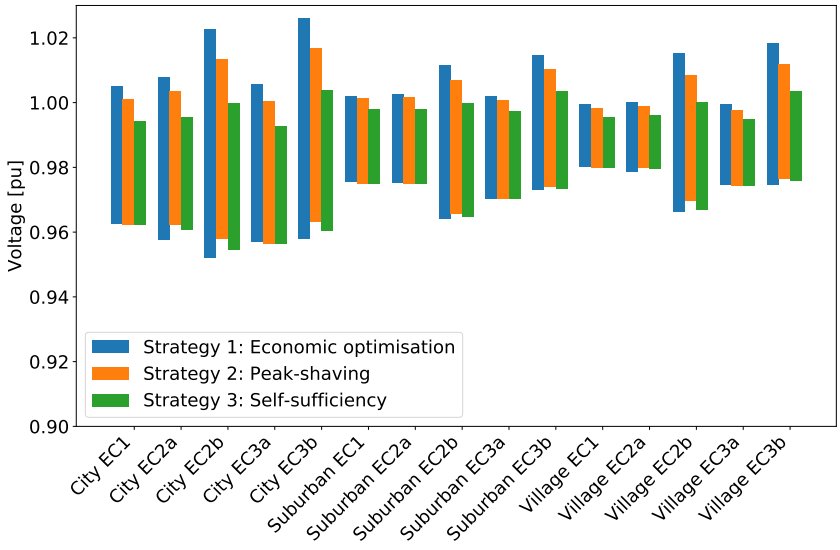


Figure 16: Comparison of minimum and maximum voltage in the EC for the three strategies.

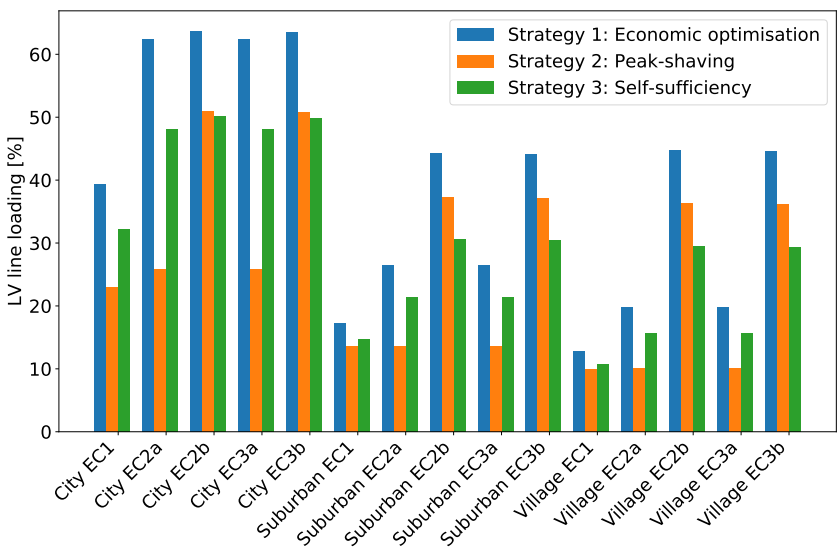


Figure 17: Comparison of line loading in the EC for the three strategies.

In Fig. 17 the impact on LV line loading of the different strategies is shown. It could be observed that generally Strategy 1 resulted in the highest LV line loading, which was due to the fact that this strategy generated new demand peaks during the night, where prices were low (see Fig. 5). Depending on the EC configuration it was

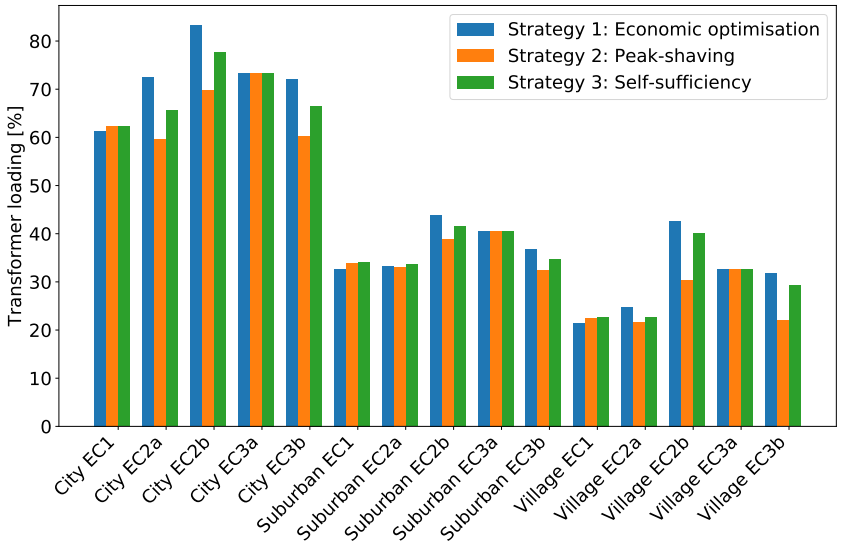


Figure 18: Comparison of transformer loading in the EC for the three strategies.

either Strategy 2 or Strategy 3 where the maximum LV line loading was lowest. For EC1, EC2*a* and EC3*a*, it was Strategy 2 where the maximum loading was lowest and for EC2*b* and EC3*b* the loading was lower in Strategy 3.

When considering transformer loading (see Fig. 18), little difference could be seen when comparing the three strategies. One explanation could be the fact that the EC corresponds to only a fraction of all the customers that the transformer was supplying. For EC2*a* and EC2*b*, it could be observed that Strategy 2 reduced slightly the maximum transformer loading. In this strategy, the EC aimed at reducing its peak-demand. This reduction is visible in the component/point, which is connecting the members of the EC to the overlying grid (similar to the point of common coupling). For EC2*a* and EC2*b* this point was the MV/LV transformer and, consequently, a reduction of its loading could be observed.

4. Discussion

This study focused on studying the impact of a communal battery owned by an EC and presented results for four possible EC strategies (Strategy 1 - Maximization of EC's economic benefit, Strategy 2 - Reducing the peak demand and Strategy 3 maximizing the self-sufficiency of the EC with batteries located in front of or behind the meter), three different grid types (city, suburban and village), five different EC configurations (based on the ECs connection to the distribution grid) and three different battery placement locations, totalling in 180 different configurations.

One can note that the amount of possible configurations explodes quickly, as there are many different potential setups, both in terms of EC configurations and distribution grids. That makes it complicated to encompass all the possible community setups when evaluating possible impacts of ECs on the distribution grid. However, this paper expanded the scientific literature with a significant amount of possible configurations, which made it possible

to investigate which factors (e.g. battery location, battery operation strategy and EC configuration) contribute significantly to the impact on the distribution grid. With respect to voltage, the results showed that battery location has the most significant impact. If the battery is located at an unfavourable location (e.g. 'End of feeder'), the range of observed voltages is significantly increased and thus, the risk of voltage violations increases. However, if the battery is connected to a favourable point (e.g. 'Beginning of feeder'), the battery does not contribute to increasing the range of voltages. This demonstrates the need for communication between the EC and the DSO when an EC is established with a communal BESS. How much the ECs impact the loading of the LV grid was investigated based on the observed maximum LV line loading and voltage, respectively. The results showed that the battery operation strategy has the highest impact on LV line loading. Strategy 2 and 3 were generally decreasing the LV line loading in comparison to Strategy 1. Due to the higher concentration of consumers, the highest LV line and transformer loading could be observed in the City grid. With respect to EC configuration, the impact grows with the total consumption of the EC. In the investigated configurations and grids, it appears that EC1, which consists of only households on one LV feeder, generally shows the least impact on the distribution grid. The configurations with the largest impact appear to be EC2*b* and EC3*b*, which consists of households and a commercial customer.

EC sizing LP problem revealed that the capacities of BESS and PVs were significantly larger in the City grid for Strategy 1 than in the other grid types. As the EC sizing problem encompassed all the costs imposed on the EC today, the obtained results present realistic outcomes of the potential EC sizing based on the current regulatory framework. Currently, the potential ECs do not need to take into account a potential distribution system impact, and thus, they do not need incentives to alternate their operation based on the grid conditions. However, Strategy 2 revealed that changing the operation strategy in line with the needs of the DSO (reducing peak load), the EC could significantly help to mitigate the adverse distribution grid impact, for an only minor increase in costs. In a green transition, the latter could potentially help integrating more electric vehicles (EVs), heat pumps and PVs into the distribution grid without the need for grid reinforcements. The ECs would provide flexibility by adjusting their behaviour based on the grid conditions. The potential regulatory framework for incentivizing this behaviour was out of the scope of this paper, and it presents an interesting topic for future research.

Some aspects of this study can be compared to the available literature. El-Batawy et al. [24] perform their study from a DSO perspective and with optimal parameters for their designed BESS they find service voltage improvements of 2.6 % and 4.5 % for two of their scenarios. This study was carried out from the EC's perspective, which caused slight voltage deterioration on the main grid. This shows that the grid impact can be significantly different based on the operating strategy and goals of the ECs. Furthermore, Hatta et al. [15] have reported voltage violations in the upper end of the allowed voltage range, whereas this study found risks for voltage violations in the lower end. This could be the consequence of different capacities of PVs and BESS installed in the ECs. Moreover, we found that the voltage violation risk highly depends on the location of the BESS and to some extent, the operational strategy. As [15] has reported only one scenario, the conclusions based on [15] are encumbered with considerable uncertainty.

There are also limitations in the performed case study. First, as the number of possible configurations explodes quickly, it is possible that some interesting configurations have been left out in this study. However, this study covered significantly more configurations than the presented literature. The City grid resulted in the largest optimal

capacities of both PV and BESS. As cities tend to be densely populated, it could be possible that the assumed capacities violated the real-life spatial constraints. However, investigating and modeling of space and rooftop area availability for installing those technologies was out of the scope of this study.

This study is highly relevant for the current development of ECs, primarily in European countries. Recently, there has been an increased amount of initiatives on energy communities but there have not been many studies systematically assessing the impact of energy communities from different perspectives. This study significantly enriched this debate by assessing different impacts if different strategies are to be applied. The results of this paper can be used by both policy makers, as well as DSO operators. It has been shown that by slightly alternating the incentives to the ECs, a significantly different outcomes can be found. Thus, DSOs and policy makers can use the study to create an inclusive framework for energy communities that already anticipates the development and impact on the distribution grid.

5. Conclusion

In this paper, the sizing of energy communities with photo-voltaics and a communal battery was investigated as well as its impact on the distribution grid. In order to determine what are cost-optimal installed capacities of photo-voltaics and of a community battery of an energy community based on a current regulatory framework, sizing the energy communities was carried out using linear-programming optimization based on four different strategies, maximizing of the energy community's own economic benefit, reducing the peak demand of the energy community and maximizing self-sufficiency with the BESS located behind and in front of the meter. Since the design criteria for distribution grids depend on the supplied customers, three different grid types were considered, namely a city, a suburban and a village grid. In all types the distribution grid consisted of two low-voltage and one medium-voltage grid. Different energy community configurations based on type of customers were also assessed. Taking into account all the mentioned configurations resulted in 180 different scenarios that were assessed in this study.

The levelized cost of electricity was the lowest in the village grid for the energy community that included a large consumer (a supermarket, configuration EC3*b*), with a value of 0.25 EUR/kWh. However, the levelized cost of electricity was similar in all the cases and was in the most expensive setup (village and suburban grid, configurations EC2*a* and EC3*a*) only 5.9 % higher. The optimal installed capacities of photo-voltaics and a communal battery were the largest in the city grid, which included a supermarket (EC3*b*), with capacities of 189 kW and 163 kWh, respectively. These values were 52 % and 35 % larger than in the same energy community configuration for a village grid and 41 % and 27 % larger than in the same energy community configuration for a suburban grid.

To assess the impact of the location of a communal battery on the distribution grid, grid analysis were carried out, which showed that the battery location mainly impacts the voltage. The largest voltage deviations could generally be observed when the battery was located at the end of the feeder. In the City grid for configuration EC2*b* and EC3*b* the observed voltages where between 0.91 pu to 1.02 pu. When the battery was located at the beginning of the feeder this range could be reduced to $0.96 - 1.02$ pu. To assess the impact of different energy community configurations and operation strategies, additional grid analysis were carried out in the three distribution grid types. The analysis showed that the largest impact on low-voltage line loading was observed in the City grid and it was shown that the chosen operation strategy greatly impacts the loading. By switching from Strategy 1 (maximizing

economic benefits) to Strategy 2 (peak-shaving) the maximum low-voltage line loading could be reduced by up to 58.5 % (City grid EC2a and EC3a), while the costs only increased by 0.3 %. It was further shown that peak shaving of 70% of the load was both technical and economic maximum, after which the system costs, as well as the line loading dramatically increased.

Since in this study the focus was solely on electricity, future work could extend this study to include communal heating and interactions with charging of electric vehicles.

Acknowledgments

This project was funded by CITIES project nr. DSF1305-00027B and by Flexible Energy Denmark (FED), funded by the Danish Innovationsfonden, as well as the Research Centre on Zero Emission Neighbourhoods in Smart Cities (FME ZEN) funded by the ZEN partners and the Research Council of Norway. Their support is greatly appreciated.

References

[1] European Commission, A Clean Planet for all. A European long-term strategic vision for a prosperous, modern, competitive and climate neutral economy (2018).

[2] W. Lutsch, Clean energy for all Europeans (2019). doi:10.2833/9937.

[3] A. Caramizaru, A. Uihlein, Energy communities : an overview of energy and social innovation, 2019. doi:10.2760/180576.
URL https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/energy-communities-overview-energy-and-social-innovation

[4] IElectrix.
URL https://ielectrix-h2020.eu/

[5] O. Gandhi, D. S. Kumar, C. D. Rodríguez-Gallegos, D. Srinivasan, Review of power system impacts at high PV penetration Part I: Factors limiting PV penetration, Solar Energy 210 (June) (2020) 181–201. doi:10.1016/j.solener.2020.06.097.
URL https://doi.org/10.1016/j.solener.2020.06.097

[6] E. Elliott, N. Shanklin, S. Zehtabian, Q. Zhou, D. Turgut, Peer-to-Peer Energy Trading and Grid Impact Studies in Smart Communities, 2020 International Conference on Computing, Networking and Communications, ICNC 2020 (2020) 674–678doi:10.1109/ICNC47757.2020.9049665.

[7] M. Di Clerico, L. Cocchi, C. Noce, F. Cazzato, C. Pezzato, G. Caneponi, V. Bufano, Prosumers' Battery Electrical Storage Systems: new ancillary services, impact on network planning and operation, in: CIGRE Session 46, 2016. doi:10.1109/PESGM.2012.6345361.
URL http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5211198

[8] M. Müller, L. Viernstein, C. N. Truong, A. Eiting, H. C. Hesse, R. Witzmann, A. Jossen, Evaluation of grid-level adaptability for stationary battery energy storage system applications in Europe, Journal of Energy Storage 9 (2017) 1–11. doi:10.1016/j.est.2016.11.005.

[9] M. Manbachi, H. Farhangi, A. Palizban, S. Arzanpour, Community Energy Storage impacts on smart grid adaptive Volt-VAR Optimization of distribution networks, 2016 IEEE 7th International Symposium on Power Electronics for Distributed Generation Systems, PEDG 2016 (2016) 3–10doi:10.1109/PEDG.2016.7527005.

[10] M. Resch, Impact of operation strategies of large scale battery systems on distribution grid planning in Germany, Renewable and Sustainable Energy Reviews 74 (December 2016) (2017) 1042–1063. doi:10.1016/j.rser.2017.02.075.
URL http://dx.doi.org/10.1016/j.rser.2017.02.075

[11] M. Resch, J. Bühler, B. Schachler, R. Kunert, A. Meier, A. Sumper, Technical and economic comparison of grid supportive vanadium redox flow batteries for primary control reserve and community electricity storage in Germany, International Journal of Energy Research 43 (1) (2019) 337–357. doi:10.1002/er.4269.

[12] E. Barbour, D. Parra, Z. Awwad, M. C. González, Community energy storage: A smart choice for the smart grid?, Applied Energy 212 (2018) 489–497. doi:10.1016/j.apenergy.2017.12.056.

[13] G. Asimakopoulou, N. Hatziargyriou, Models for the Integration of Local Energy Communities in Energy Markets (0123) (2018) 7–8.

[14] J. Liu, C. Hu, A. Kimber, Z. Wang, Uses, Cost-Benefit Analysis, and Markets of Energy Storage Systems for Electric Grid Applications, Journal of Energy Storage 32 (February) (2020) 101731. doi:10.1016/j.est.2020.101731.
URL https://doi.org/10.1016/j.est.2020.101731

[15] H. Hatta, E. Omine, N. Takahashi, Proposal of impact assessment method for autonomous operation of smart community using battery energy storage systems, IEEE PES Innovative Smart Grid Technologies Conference Europe (2016) 384–388doi:10.1109/ISGT-Asia.2016.7796416.

[16] F. Ceglia, P. Esposito, E. Marrasso, M. Sasso, From smart energy community to smart energy municipalities: Literature review, agendas and pathways, Journal of Cleaner Production 254 (2020) 120118. doi:10.1016/j.jclepro.2020.120118.
URL https://doi.org/10.1016/j.jclepro.2020.120118

[17] B. Faessler, M. Schuler, M. Preißinger, P. Kepplinger, Battery storage systems as grid-balancing measure in low-voltage distribution grids with distributed generation, Energies 10 (12) (2017). doi:10.3390/en10122161.

[18] F. Boulaire, A. Narimani, J. Bell, R. Drogemuller, D. Vine, L. Buys, G. Walker, Benefit assessment of battery plus solar for customers and the grid, Energy Strategy Reviews 26 (June) (2019) 100372. doi:10.1016/j.esr.2019.100372.
URL https://doi.org/10.1016/j.esr.2019.100372

[19] H. Wang, N. Good, P. Mancarella, Economic analysis of multi-service provision from PV and battery based community energy systems, 2017 IEEE Innovative Smart Grid Technologies - Asia: Smart Grid for Smart Community, ISGT-Asia 2017 (2018) 1–6doi:10.1109/ISGT-Asia.2017.8378390.

[20] I. F. G. Reis, I. Gonçalves, M. A. R. Lopes, C. H. Antunes, Assessing the Influence of Different Goals in Energy Communities' Self-Sufficiency—An Optimized Multiagent Approach, Energies 14 (4) (2021) 989. doi:10.3390/en14040989.

[21] B. Fina, H. Auer, W. Friedl, Cost-optimal economic potential of shared rooftop PV in energy communities: Evidence from Austria, Renewable Energy 152 (2020) 217–228. doi:10.1016/j.renene.2020.01.031.
URL https://doi.org/10.1016/j.renene.2020.01.031

[22] P. Moura, U. Sriram, J. Mohammadi, Sharing Mobile and Stationary Energy Storage Resources in Transactive Energy Communities (2020).
URL http://arxiv.org/abs/2008.08971

[23] R. Rocha, J. V. Collado, T. Soares, F. Retorta, Local Energy Markets for Energy Communities with Grid Constraints (2020) 1–6doi:10.1109/eem49802.2020.9221878.

[24] S. A. El-Batawy, W. G. Morsi, Optimal design of community battery energy storage systems with prosumers owning electric vehicles, IEEE Transactions on Industrial Informatics 14 (5) (2018) 1920–1931. doi:10.1109/TII.2017.2752464.

[25] NordREG, Electricity distribution tariffs - report by the nordreg network regulation wg (2 2021).
URL http://www.reguleringsmyndigheten.no/

[26] K. Deb, Multi-Objective Optimization using Evolutionary Algorithms, 2nd Edition, John Wiley & Sons, Ltd, Chichester, 2002.

[27] K. Miettinen, Some Methods for Nonlinear Multi-objective Optimization, in: E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, D. Corne (Eds.), Evolutionary Multi-Criterion Optimization, Springer, 2001.

[28] Y. Haimes, Yacob, L. S. Lasdon, D. A. Wismer, One a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization, IEEE Transactions on Systems, Man and Cybernetics July 1971 (1971) 296–297.

[29] G. Mavrotas, Effective implementation of the ϵ-constraint method in Multi-Objective Mathematical Programming problems, Applied Mathematics and Computation 213 (2) (2009) 455–465. doi:10.1016/j.amc.2009.03.037.
URL http://dx.doi.org/10.1016/j.amc.2009.03.037

[30] L. Thurner, A. Scheidler, F. Schäfer, J. Menke, J. Dollichon, F. Meier, S. Meinecke, M. Braun, pandapower — An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems, IEEE Transactions on Power Systems 33 (6) (2018) 6510–6521. doi:10.1109/TPWRS.2018.2829021.

[31] J. Dickert, M. Domagk, P. Schegner, Benchmark low voltage distribution networks based on cluster analysis of actual grid properties, 2013 IEEE Grenoble Conference (2013) 1–6doi:10.1109/PTC.2013.6652250.
URL http://ieeexplore.ieee.org/document/6652250/

[32] K. Strunz, E. Abbasi, R. Fletcher, N. Hatziargyriou, R. Iravani, G. Joos, TF C6.04.02 : TB 575 – Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources, 2014.

[33] G. Kerber, Aufnahmefähigkeit von Niederspannungsverteilnetzen für die Einspeisung aus Photovoltaikkleinanlagen, Ph.D. thesis, Technische Universität München (2011).

[34] CENELEC 2019, EN 50160:2010+A3:2019 Voltage characteristics of electricity supplied by public electricity networks, Tech. rep., CENELEC (2019).

Appendix A.

Appendix A.1. Input data for the LP

This appendix presents more detailed input data that is needed to recreate the case study. In the main part of the manuscript, the most important figures were presented, while in the appendix the remaining figures are reported.

PV hourly generation profile was taken from from [ref. https://re.jrc.ec.europa.eu/pvg_tools/en/#PVP] With the following parameters used: Slope angle 40^o

Azimuth angle 0^o

Summary of the profile (the profile downloaded from the referenced link consists of a time series with 8760 hourly values):

	kWh generated / kW installed
mean (capacity factor)	0.12
min	0
max	1
count	8760

Formation of the electricity price for households (2019): Data has been taken from Energinet and Radius, Energinet (energinet.dk) being a transmission system operator and Radius (https://radiuselnet.dk/) a distribution system operator.

In 2019, the end-user total price, used for our case study of Denmark, consisted of:

- Electricity tax: 0.119 EUR/kWh
- Public Service Obligation (PSO) charge: 1^{st} quarter: 0.0083EUR/kWh, 2^{nd} quarter 0.021 EUR/kWh, 3^{rd} quarter 0, 4 th quarter 0.01 EUR/kWh
- Transmission tariff: 0.013 EUR/kWh

- Distribution tariff: 0.027 EUR/kWh; except from October - March from 17:00 - 20:00: 0.075 EUR/kWh
- Electricity price on Nordpool day-ahead market DK2 sector (changes hourly): mean price: 0.0399 EUR/kWh, max 0.1097 EUR/kWh, min -0.061 EUR/kWh
- 15% of retailer margin on the Nordpool price
- 25% VAT on the total price

Summing up all the costs and the day ahead electricty price, an hourly changing price time series has been created. A summary of these used electricity price for households can be seen in the following table (electricity from the grid):

	EUR/kWh
mean	0.282
min	0.123
max	0.463
count	8760

Investments, fixed operating and maintenance costs and other economic parameters can be seen in the following table. These parameters were fed to the objective function of the LP problem.

	PV	Li-ion Community Battery
Investment cost I EUR/kW for PV; EUR/kWh for a battery	1130	337.4
Project lifetime n	15	12
Discount rate d	0.07	0.07
Fixed O&M EUR/(kW * year)	12.8	-
Efficiency	1*	charging: 0.98; discharging 0.97

* Because the rated power already took into account the efficiency of the panel.

Equated yearly installment of the investment calculation (PV: EUR/(kW * year); batteries: EUR/(kWh * year)):

$$EYI = \frac{I \cdot d \cdot (1+d)^n}{(1+d)^{(n-1)}}$$

Appendix A.2. Strategy 2 - Grid impact results

In this section the Strategy 2 results for all grid types are shown. Figure A.19 shows the LCOE and maximum LV line loading as a function of peak demand given as a percentage of the demand when no peak-shaving is performed. Table A.4 shows relative difference in line loading for the Strategy 2 for each of the set-ups.

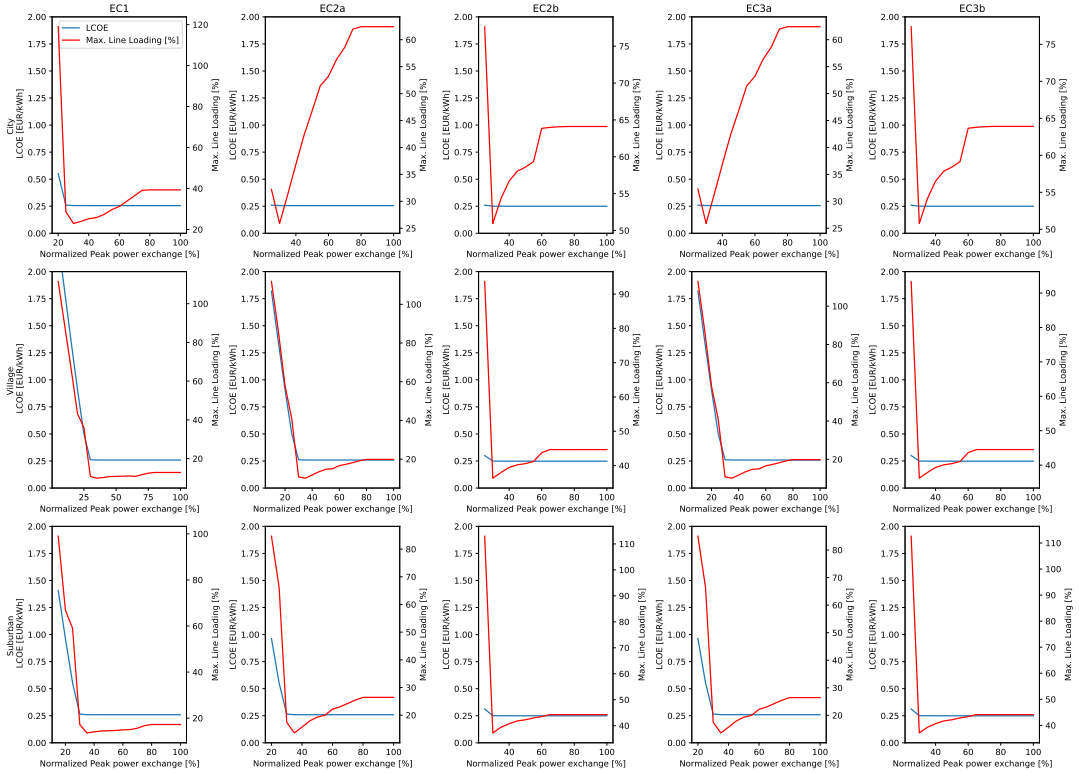


Figure A.19: LCOE and maximum line loading for Strategy 2 against peak demand. Peak demand is given as a percentage of Strategy 1 values as a baseline.

Table A.4: Highest reduction in maximum line loading (cost increase at that iteration with respect to Strategy 1 in percent) for Strategy 2 [%]

	City	Village	Suburban
EC1	41.68 (0.31)	23.02 (0.18)	21.26 (0.22)
EC2*a*	58.51 (0.31)	48.98 (0.18)	48.63 (0.22)
EC2*b*	20.51 (0.36)	18.67 (0.56)	15.94 (0.52)
EC3*a*	58.51 (0.31)	48.98 (0.18)	48.51 (0.22)
EC3*b*	20.53 (0.36)	18.69 (0.56)	15.95 (0.52)

{2.7182818284}ϑφερτυθιοπσδφγηξκλ