The Influence of Model Complexity on Demand Side Management with Distributed Energy Resources

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

M.Sc. Chang Li
geb. in Jingzhou, China

Tag der mündlichen Prüfung:	15.01.2026
Hauptreferent:	Prof. Dr. Veit Hagenmeyer
Korreferent:	Prof. Dr. rer. pol. Wolf Fichtner

Declaration

Karlsruhe, August 1, 2025

I hereby declare that I have undertaken this work independently, without any improper assistance, and have utilized only the resources explicitly cited.

Chang Li

Acknowledgements

As time swiftly passes, my past years of research are drawing to a close. At this juncture, I would like to extend my sincere thanks to the individuals listed below, without whose support I would not have been to complete this research.

Firstly, I would like to express my sincere gratitude to my supervisors Prof. Dr. Veit Hagenmeyer and Prof. Dr. rer. pol. Wolf Fichtner, as well as my mentors Dr. Kevin Förderer and apl. Prof. Dr. Jörg Matthes. Particularly, special thanks go to Prof. Dr. Veit Hagenmeyer for his insightful thoughts and great support in defining a meaningful topic that is worth researching at an early stage as well as in improving the quality of each scientific publication along this journey. I would like to extend my appreciation to my primary mentor Dr. Kevin Förderer for his direct guidance and valuable knowledge about energy system modeling, which constantly refine my topic. I am also grateful to apl. Prof. Dr. Jörg Matthes for his support, for providing constructive feedback both within and beyond the scope of the research. Furthermore, I sincerely thank Prof. Dr. rer. pol. Wolf Fichtner for the thoughtful comments and recommendations on this dissertation. In addition, I would like to thank Prof. Dr. Martin Cichon and the defense committee for their expertise.

I am deeply thankful to be a member of the Institute for Automation and Applied Informatics (IAI), which is an outstanding institute with talented employees and advanced research infrastructures. In particular, I would like to thank the IAI colleagues at the Energy Lab, including but not limited to Tobias Moser, Jan Wachter and Luigi Spatafora. With their support, I was able to utilize the experimental data to validate my research hypothesis in a first case study. I would also like to thank the Helmholtz Research Program Energy System Design (ESD) for providing a valuable platform which enables us to focus on the ideas, methods and technologies towards a more digitalized energy system in the future. I want to convey my appreciation to Dr. Christine Stürzer who is the program manager and has organized several meaningful exchange events with other participating Helmholtz centers such as Forschungszentrum Jülich (FZJ) and German Aerospace Center (DLR).

Furthermore, I would like to express my thanks to the numerous doctoral students and colleagues at the IAI who have accompanied me in this journey in different ways such as scientific exchange and cooperation in publications, namely Jovana, Moritz Weber, Felix, Tobias, Alex, Hao, K.

Çakmak, Jianlei, Gökhan, Jakob, Simon, Natascha, Frederik, Xinliang, Moritz Noskiewicz, Ellen, and many more. Specifically, I would want to take this opportunity to thank Jovana for being my trusted collaborator during this PhD journey. I am grateful for our fruitful discussions, joint supervision of a master's thesis and numerous joint publications. I would also like to thank Moritz Weber for giving me so many useful advice in various aspects, especially during our joyful coffee break. In addition, I would like to thank Daniel Boser for being a trustworthy friend and giving me many new thoughts from a german perspective.

Finally, I would like to thank my parents for their wholehearted support through my whole life, which, honestly, cannot be expressed in words. I also appreciate my cohort Miao Zhang for her support in various aspects.

Abstract

In the energy transition from fossil energy sources to renewable energy sources, residential buildings are increasingly integrating a range of components with high electricity consumption, including heat pump (HP) for space heating and domestic hot water production and electric vehicle (EV) charging points. These components could be collectively referred to as Distributed Energy Resources (DERs), which can either be individual physical assets or aggregated virtually. Furthermore, the corresponding technologies within DERs enable the adjustment of electricity consumption to align with power generation, a concept known as Demand Side Management (DSM). To investigate the effects of DERs in DSM, a wide variety of models and simulations of the relevant components are required. However, when modeling these components and their interactions for DSM purposes, the appropriate level of model complexity required for various DSM applications often remains ambiguous. This is primarily due to the inherent trade-off between the model's complexity and the utility it provides, as their interactions can significantly influence the overall performance and applicability of DSM strategies.

Inspired by Gossen's First Law in the economics, which describes the law of diminishing marginal utility, this thesis further investigates the hypothesis that Gossen's First Law also holds in the modeling for DSM. The proposed hypothesis states that in general, in the modeling of DERs for DSM, the relationship between the model complexity and the model utility could be represented by a diminishing marginal utility curve. To validate the proposed hypothesis, the quantification methods for model complexity and model utility are first reviewed, discussed and defined within the scope of this thesis. In the following, two heat pump case studies and a use case of EV charging stations are carried out for validation. Lastly, a new Shapley value-based quantification method for the contribution of submodels within a model class is introduced to further explore the DERs modeling.

The results in all use cases are basically in line with the proposed hypothesis and indicate that the hypothesis can act as a guideline that could help developing more simple and efficient models in the context of DSM.

Kurzfassung

Im Rahmen der Energiewende von fossilen zu erneuerbaren Energiequellen werden in Wohngebäuden zunehmend eine Reihe von Komponenten mit hohem Stromverbrauch integriert, darunter Wärmepumpen für die Raumheizung sowie die Warmwasserbereitung, und die Ladestationen für Elektrofahrzeuge. Diese Komponenten können gemeinsam als dezentrale Energiequellen (Distributed Energy Resources, DERs) bezeichnet werden. Die können entweder als einzelne Anlagen oder virtuell aggregierte Einheiten dargestellt werden. Darüber hinaus ermöglichen die entsprechenden Technologien innerhalb der DERs eine Anpassung des Stromverbrauchs an die Stromerzeugung, ein Konzept, das als Demand Side Management (DSM) bekannt ist. Um die Auswirkungen von DERs im Rahmen des DSM zu untersuchen, ist eine Vielzahl von Modellen und Simulationen der relevanten Komponenten erforderlich. Allerdings bleibt bei der Modellierung dieser Komponenten und ihrer Interaktionen für DSM-Zwecke jedoch oft unklar, wie komplex das Modell für verschiedene DSM-Anwendungen sein muss. Dies liegt in erster Linie an dem inhärenten Kompromiss zwischen der Komplexität des Modells und dem Nutzen, den es bietet, da ihre Wechselwirkungen die Gesamtleistung und Anwendbarkeit von DSM-Strategien erheblich beeinflussen können.

Inspiriert von Gossens erstem Gesetz in den Wirtschaftswissenschaften, das das Gesetz des abnehmenden Grenznutzens beschreibt, stellt diese Arbeit daher die Hypothese auf, dass Gossens erstes Gesetz auch in der Modellierung für DSM gilt. Die aufgestellte Hypothese besagt, dass im Allgemeinen bei der DER-Modellierung im Rahmen des DSM die Beziehung zwischen Komplexität und Nutzen durch eine abnehmende Grenznutzungskurve dargestellt werden kann. Um die aufgestellte Hypothese zu validieren, werden die Quantifizierungsmethoden für die Modellkomplexität und die Nützlichkeit von Modellen zunächst im Rahmen dieser Arbeit überprüft, diskutiert und definiert. Im Anschluss werden zwei Use Cases zu Wärmepumpen sowie ein Use Case zu Ladestationen für Elektrofahrzeuge durchgeführt, um die Validierung sicherzustellen. Abschließend wird eine neue, auf dem Shapley Value basierende Quantifizierungsmethode zur Bewertung des Beitrags von Teilmodellen innerhalb einer Modellklasse im Kontext von DSM vorgestellt, um die Modellierung von DERs weiter zu erforschen.

Die Ergebnisse in allen Anwendungsfällen stimmen im Wesentlichen mit der aufgestellten Hypothese überein und deuten darauf hin, dass die Hypothese als Leitlinie dienen kann, die bei der Entwicklung einfacherer und effizienterer Modelle im Kontext von DSM hilfreich sein könnte.

Preface

"Essentially, all models are wrong, but some are useful."

"Remember that all models are wrong;
the practical question is how wrong do they have to be to not be useful."

George E. P. Box and Norman R. Draper
in *Empirical Model-Building and Response Surfaces*

Publications

Jovana Kovačević, Chang Li, Kevin Förderer, Hüseyin K Çakmak, and Veit Hagenmeyer. Improving reduced-order building modeling: Integration of occupant patterns for reducing energy consumption. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems*, pages 569–579. ACM, 2024.
DOI: https://doi.org/10.1145/3632775.3662163.

Jovana Kovačević*, Chang Li*, Gina Brecher, Kevin Förderer, Hüseyin K Çakmak, and Veit Hagenmeyer. The impact of model complexity reduction on system dynamics: A modelica case study of a residential heating system. In *Proceedings of the 12th IEEE Conference on Technologies for Sustainability (SusTech 2025)*, pages 1–8. IEEE, 2025. *co-first authorship
DOI: https://doi.org/10.1109/SusTech63138.2025.11025706.

Chang Li, Kevin Förderer, Tobias Moser, Luigi Spatafora, and Veit Hagenmeyer. Gossen's first law in the modeling for demand side management: A first heat pump case study. In *Energy Informatics Academy Conference*, pages 111–125. Springer Nature, 2023.
DOI: https://doi.org/10.1007/978-3-031-48652-4_8.

Chang Li, Miao Zhang, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. A comparative analysis of machine learning algorithms for aggregated electric chargepoint load forecasting. In *Proceedings of the 9th International Conference on Sustainable and Renewable Energy Engineering*, volume 545. EDP Sciences, 2024.
DOI: https://doi.org/10.1051/e3sconf/202454501004.

Chang Li, Felix Langner, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. A survey on the current state of demand response implementations: Models and approaches. *Energy Informatics Review*, 5(1), ACM, 2025.
DOI: https://doi.org/10.1145/3757888.3757890.

Chang Li, Gina Brecher, Jovana Kovačević, Hüseyin K Çakmak, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. Gossen's first law in the modeling for demand side management: a thorough heat pump case study with deep learning based partial time series data generation. *Energy Informatics*, 7(47), Springer Nature, 2024.
DOI: https://doi.org/10.1186/s42162-024-00353-z.

Contents

1 Introduction

1.1 Motivation

Advancing towards a sustainable energy system, enhancing energy flexibility has become increasingly important to accommodate the growing integration of renewable energy sources. According to the federal statistical office of Germany (Statistisches Bundesamt 2023), the total energy consumption of private households in Germany has only slightly increased by 7.8% from 2015 to 2021. During the same period, however, the share of renewable energy experienced a substantial growth of 23.3%, eventually comprising over 14% of the total energy mix. Energy flexibility refers not only to the ability to adjust power generation, which is challenging with renewables due to their intermittent nature, but also to adapting electricity load on the demand side, which is also known as Demand Side Management (DSM) (Impram et al. 2020). As a strategy to control electricity demand and to manage demand over different terms, DSM is a very broad term which is usually further defined based on specific use cases (Li et al. 2025). To control electricity load on the demand side, the load-changing ability is required and this ability is usually provided by Distributed Energy Resources (DERs). DERs are small-scale electricity generation, storage and consumption systems that are able to be controlled. For instance, in the residential sector, assets such as heat pumps for heating and supplying domestic hot water, rooftop solar panels for self-generating electricity, home charging stations for e-mobility and even fuel cell systems as home power plants (Thomas et al. 2020) can be classified as DERs. Although the DERs have the ability to enhance energy flexibility on the demand side, substantial efforts are still required to address ongoing challenges: in 2024, the federal environmental agency in Germany has reported that despite of the biggest decline in terms of climate emissions in 2023 since 1990, the building sector still exceeded the annual emissions permitted under the Federal Climate Protection Act (KSG) by around 1.2 million tonnes of CO_2 equivalents (Umwelt Bundesamt 2024). To improve this critical situation, one suggestion in (Umwelt Bundesamt 2024) is increased utilization of heat pumps. This again underlines the importance of utilizing DERs on the demand side, especially in the residential sector, and provides motivation to pursue research in this direction.

To investigate the effects of DERs in DSM applications, a wide variety of models and simulations of the relevant energy systems and components are required. However, when modeling such

components and their interactions for DSM applications, determining the appropriate level of model's complexity remains challenging. This is due to the inherent trade-off between model utility and complexity. A complex model can usually provide more meaningful results than a simple model, but the effort required for modeling increases accordingly (Chiasson 1999). Besides, the main problem with complex models is that the mathematical optimizations could no longer be solved numerically in a reasonable time which is usually essential in DSM. Therefore, simple models are also needed for DSM and it is necessary to investigate the relationship between the utility and the complexity of a model in order to provide a quantified reference for different DSM applications.

The first inspiration to solve this core question addressed by this thesis comes from daily life in different domains. In the sports medicine domain, (Morton et al. 2018) investigate the relationship between the change in fat-free mass and the relative total protein intake. The results show that fat-free mass stagnates after the so-called breakpoint, which means fat-free mass increases no more with further increasing protein intake after this point. In the consumption domain, (Horowitz et al. 2007) conduct three types of experimental treatments. The first experiment cover typical goods-money tradeoffs to investigate, how the compensation demand, namely money, of subject groups changes with respect to different amount of goods. The results show that the compensation demand per unit decreases when the amount of goods increases.

The primary question of this thesis is, whether a similar result manifests for DSM applications, with the horizontal axis being complexity and the vertical axis being utility, or not. Another inspiration is derived from Gossen's First Law in economics. Gossen's laws are named after Hermann Heinrich Gossen (1810-1858) and Gossen's First Law is also known as the law of diminishing marginal utility, which states that as consumption increases, marginal utility gained from the consumption decreases (Gossen 1854). To quantitatively solve this main question of the thesis, a novel approach and hypothesis is first proposed for investigating the relationship between the model utility and the model complexity within the domain of DSM. The proposed hypothesis states that in general the complexity-utility relationship in the field of DSM modeling could be represented by a diminishing marginal utility curve.

For testing the proposed hypothesis, it is imperative to undertake a validation process through diverse use cases, in which different sub-questions are addressed. More details on this are provided in Section 1.4.

1.2 Related Work

To establish the broader context of this thesis, this section reviews related work across relevant domains with modeling and model complexity in DSM with DERs in focus. By analyzing different contributions of existing research, this section also aims to identify existing gaps in the literature and therefore underlines the importance and novelty of investigating the proposed hypothesis in the present thesis. Furthermore, it should be pointed out that the subsequent chapters will present further related work specific to their respective topics.

In recent years, the literature dedicated to the modeling of Demand Response (DR) or DSM has grown substantially. For example, the part highlighted in orange in Fig. 1.1 shows the publication trend from 2010 to 2024 based on the keyword combination "('Demand Response' or 'Demand Side Management') and 'modeling'" in Scopus. This reflects the increasing academic and practical

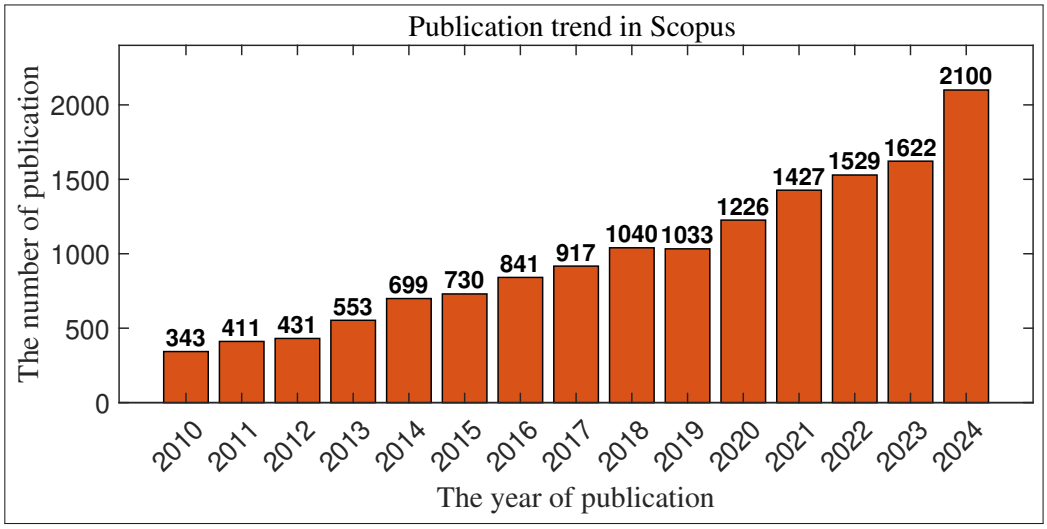


Figure 1.1: Publication trend from 2010 to 2024 based on the keyword combination "('Demand Response' or 'Demand Side Management') and 'modeling'" in Scopus

interest in these areas. For instance, (Turitsyn et al. 2011) propose a modeling framework for 4 types of individual devices, that is, optional loads, deferrable loads, control loads and storage loads, which are expected to participate in future DR markets on distribution grids. The proposed framework is based on the standard Markov Decision Process (MDP) approach to focus only on the qualitative features of the loads. Therefore, the differentiated 4 types of models are optimal and generic. The modeling of specific systems and synergies between different systems is also not investigated. In 2013, (Petersen et al. 2013) propose a more generic taxonomy to model flexibility

on the demand side in Smart Grids (SG), categorizing systems into three distinct groups to facilitate the optimization and resolution of challenges related to flexibility within SGs. This modeling approach offers the advantages of simplifying the modeling process and enhancing optimization efficiency. However, it falls short in addressing the complexities of real-world energy systems, particularly the influence of factors such as temperature, due to the high level of abstraction in the models. Consequently, these models face significant limitations in their direct applicability to energy systems on the demand side. Similarly, the authors in (Barth et al. 2018) develop a comprehensive modeling framework to mathematically describe demand side flexibility based on different existing models. The modeling framework is also very abstract by integrating up to 14 different features, which represent constraints or parameters, into mathematical models.

In contrast, to obtain more details and to address the complexities of real-world energy systems, detailed and complex models, i.e., white box models that are based on detailed physical entities of a system, the laws of physics such as the thermodynamics and complex numerical techniques, are up to the task. For instance, (Śliwa and Gonet 2005) develop a detailed theoretical model of borehole heat exchangers, which is an important part of geothermal heat pump systems, utilizing very complex heat flow equations, such as a three-dimensional equation for heat flow in the Cartesian coordinates system. While the proposed model can lead to accurate simulation results, the high complexity of the model and its corresponding solving method viz. the so-called Finite Difference Method (FDM) limits the efficiency of solving the model. Similarly in another work (Peralta et al. 2021), a detailed mathematical model of a ground source heat pump (GSHP) system with vertical U-pipe heat exchangers is proposed to calculate the thermal load and then to participate in day-ahead and real-time electricity markets. Although several assumptions are made by the authors to reduce the complexity of the proposed model, the overall heat-flow equations remain very detailed and require numerical techniques such as Lax-Wendroff finite difference approximations to solve the model. Yet, the time resolution of the model is set to one hour in the simulation. It's evident that while detailed white box models can describe the interactions within an energy system very well, the effort for modeling is enormous and the expansibility as well as the practicability is limited, since the optimization tasks can no longer be solved in a reasonable time. Other than detailed white box models, there is another common modeling approach which pre-defines the model structure based on physical knowledge and identifies system parameters based on measurement data, that is, gray box modeling. For example, (Frahm et al. 2023) utilize the lumped resistance and capacitance approach to build a RC gray box model for the multi-zone building structure as basis to evaluate the DR performance when using flexible electrical loads. By building a gray box model with fewer necessary parameters, the complexity of the whole thermal building model is reduced. This also helps to improve the efficiency when combining the models with advanced control strategies, such as Model Predictive Control (MPC), and smaller time step, such as 15 min in (Frahm et al. 2023). However, due to the reduced complexity of

the models, gray box models can exhibit lower utility compared to white box models when the identification data sets are equally accurate (Picard et al. 2016).

The above mentioned literature dedicated to the modeling of DERs in the context of DSM and DR have been either focused on a new modeling framework or the performance of a new specific approach. As mentioned in Section 1.1, no straightforward and comprehensive investigation of the effect of DER model complexity on model utility in DSM has been done yet. Nevertheless, there are some intriguing research results regarding model complexity in other domains. The authors (Robinson 2023) explore the relationship between model complexity and model accuracy in simulation by considering 3 ideal test models, which are the manufacturing model, the service model and the panorama televisions model. The results show that diverse accuracy to model complexity trajectories have been observed in different test models. For Building Information Modeling (BIM), (McArthur 2015) consider the marginal utility of each geometric component to develop a BIM model and investigate the relationship between the increasing benefit to the team and the increasing extra effort required for a higher level of completeness. In the grid modeling domain, (Weber et al. 2023) present a novel static network reduction method to investigate its impact on a model of the German transmission grid. The reduction of the power grid model is mainly topology-driven and thus focuses on the preservation of the overall network topology. According to (Weber et al. 2023), the proposed method produces interpretable models that offer a better reusability while achieving a comparable performance in the reduction grade.

The literature survey reveals an underexplored research gap in DERs modeling in the field of DSM. Therefore, it's necessary and worthwhile to investigate the relationship between the utility and complexity of a model quantitatively in order to provide a better reference for diverse practical DSM applications.

1.3 Research Questions and Contributions of the present Dissertation

As delineated in the previous sections, the primary goal of this thesis is to investigate the relationship between the model complexity and the model utility in the context of DSM quantitatively. Mainly inspired by Gossen's First Law in economics and relevant literature in other domains, a hypothesis is first formulated to address this core issue. To validate this hypothesis, several open research questions regarding the validation process, that have not been resolved satisfactorily in the literature, are formulated in this section.

Research Question 1: How to quantify the complexity of DER models in the context of DSM?

Based on the hypothesis, complexity and utility of a model represents the two axes in a coordinate system respectively. To quantitatively investigate the relationship between them, quantifying the model complexity is an essential preliminary step. In the literature, there are many existing measures to quantify the complexity of models or functions in different scenarios such as the famous Big O notation for classifying the time complexity of algorithms. However, not all of these are applicable to the scenarios discussed in this thesis. Therefore, this thesis discusses different possibilities to quantify the model complexity and then identifies the concrete quantification methods for different scenarios with DERs in DSM.

Research Question 2: How to quantify the utility of DER models in the context of DSM?

Similar to the first research question, quantifying utility of a DER model is another necessary preliminary step. However, differing from complexity, utility is a more general concept and its definition depends strongly on the specific applications. Therefore, there is a need for an accurate specification of its scope.

Research Question 3: What is the relationship between the model complexity and the model utility of DERs models in the context of DSM?

A complex model can usually provide more meaningful results, i.e., model utility than a simple model, but the effort required to modeling it increases accordingly. The question of how to achieve a reasonable balance when modeling for DSM considering the interactions between model complexity and model utility is essential in practice. This question could be answered sensibly after the relationship between model complexity and model utility is quantitatively investigated. As mentioned in Section 1.1, this research question forms the core question of this thesis.

Research Question 4: How to quantify the contributions of a DER's submodel to model utility in DSM?

The DER models in practical application can be a simple model or a combined model with submodels. For DER models composed of several submodels, assessing the individual contribution of each submodel to the overall accuracy is of significant value for a further informed decision-making. This research question can help to take another step forward in expanding and applying the hypothesis mentioned above to practical applications.

Contributions:

- Different quantification methods for model complexity are analyzed and summarized based on a literature review. After the discussion, the number of parameters as the model complexity quantification approach is selected for white-box DER models and input size is selected for black-box DER models.
- This thesis restricts the definition of model utility to accuracy. And according to the definition of "accuracy" in ISO 5725-1, different useful metrics in descriptive statistics are discussed and utilized for quantifying model utility on a per-use-case basis.
- A hypothesis that in general the complexity-utility relationship in DER models for DSM can be represented by a diminishing marginal utility curve is proposed and tested with different case studies.
- A new approach for quantifying the contribution of a DER's submodel for DSM based on the Shapley value is introduced and tested with a case study.

1.4 Thesis Outline

Based on the research questions, this thesis is structured as follows: Chapter 2 conducts a comprehensive survey on the current state of DSM implementations, focusing on the utilized models and the corresponding approaches. The survey offers an extensive and relatively complete background and discussion of how models have been applied in practical DSM. The review results suggest that simple DER models, such as linear models, can deliver relatively strong performance in several practical DSM applications. This underlines the relevance and the importance of the subsequent investigations. Chapter 3 responds to the Research Question (RQ) 1 and RQ2 by reviewing and discussing specific quantification methods for model complexity and model utility. Specific quantification methods are then selected and explained as the foundation for the subsequent validation. Chapter 4 follows as a further step to answer RQ3 by validating the diminishing marginal utility hypothesis with three use cases, including two use cases based on white-box models and one use case based on black-box models. Chapter 5 addresses RQ4 by suggesting a method for quantifying the contribution of DER's submodels in the context of DSM based on the Shapley value. In the same chapter, one of the use cases from Chapter 4 is demonstrated to indicate the effectiveness of the proposed method. Finally, Chapter 6 concludes the present thesis with a summary and an outlook on possible future research work. To provide a more intuitive representation of the outline, Fig. 1.2 illustrates the detailed workflow of the present thesis.

In addition, it's worth noting that some previously published contents in the current thesis are block quoted from author's own publications. To clearly identify them, the direct block quotations are marked with a gray bar on the left and the sources are always provided in a box at the beginning of the corresponding chapter or section.

Major parts of this chapter are adapted or reproduced from the author's own previous work:

This is an example box, intended to illustrate the definition of this type of block quotation.

> This example sentence serves to illustrate the use of a block quotation with a gray bar on the left.

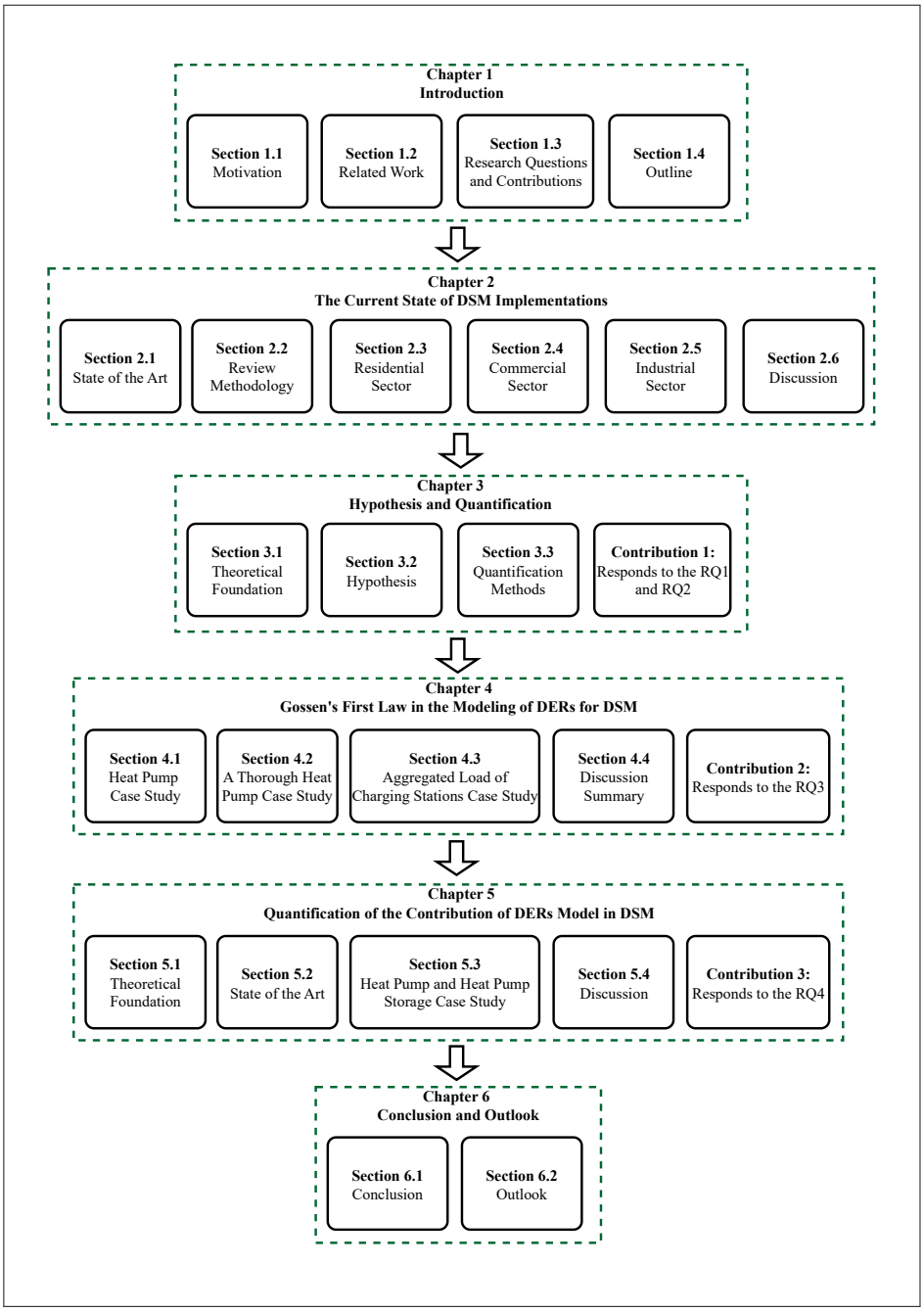


Figure 1.2: Illustration of the outline of the present thesis

2 The Current State of Demand Side Management Implementations

Major parts of this chapter are adapted or reproduced from the author's own previous work:

Chang Li, Felix Langner, Kevin Förderer, Jörg Matthes and Veit Hagenmeyer.
"A survey on the current state of demand response implementations: models and approaches". In: *ACM Energy Informatics Review*, Volume 5, Issue 1.
doi: 10.1145/3757888.3757890.

In this chapter, the current state of DSM implementations, i.e., practical experiments and applications, are reviewed and discussed in detail. First and foremost, it should be emphasized that DSM as a strategy to control electricity demand is a very broad term and seeks to manage demand over different terms. For instance, in a paper from 2011, (Palensky and Dietrich 2011) categorize DSM into four types of measures, namely Energy Efficiency, Time of Use (ToU), Demand Response (DR) and Spinning Reserve. In the more recent definitions from the IEA (IEA 2023), which are used in this thesis, ToU is part of DR and DR is subdivided into implicit DR based on prices and explicit DR based on monetary incentives. Within the scope of this thesis, DSM is primarily centered on DR.

The structure of this chapter is as follows: the first section describes the state of the art regarding existing review articles about DR researches and outlines their findings as well as limitations. Section 2.2 presents the utilized review methodology in this work, which determines the corresponding literature review results. The DR implementations in different sectors, namely residential, commercial and industrial, are discussed in Section 2.3, Section 2.4 and Section 2.5 respectively. Section 2.6 summarizes the findings and the contributions of this chapter and discusses potential directions for further development in the context of DR implementations.

2.1 State of the Art

For short-term actions, DR as a type of DSM (Dornellas et al. 2016, Impram et al. 2020), has proven to be an operational tool in maintaining reliability (Dornellas et al. 2016). In general, the goals of DR can be divided into three categories: electrical demand increase, decrease and re-planning (Impram et al. 2020). More specifically, electrical demand increase contains valley filling (Deng et al. 2015, Lund et al. 2015) and load growth (Lund et al. 2015). Demand decrease includes peak shaving (Deng et al. 2015, Lund et al. 2015) and load conservation (Lund et al. 2015). Re-planning is load shifting over a specific time horizon (Deng et al. 2015, Lund et al. 2015). Fig. 2.1 illustrates these goals in the context of DR based on the literature mentioned above.

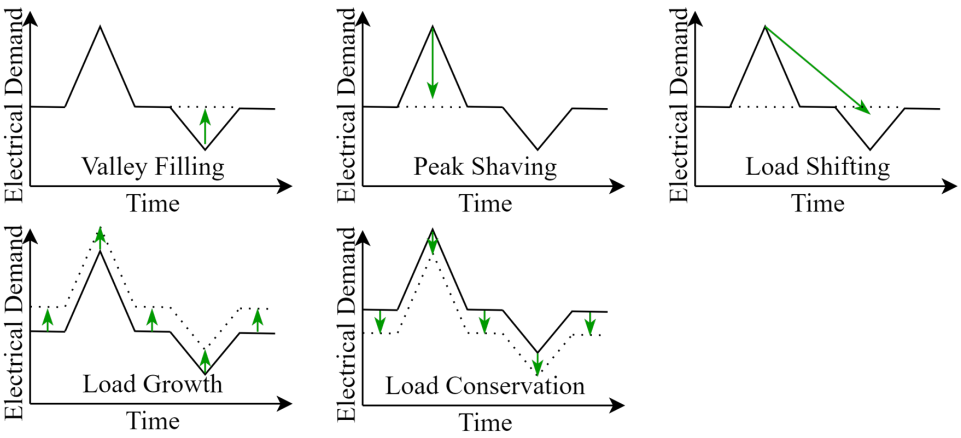


Figure 2.1: Illustration of the goals in the context of DR based on (Deng et al. 2015, Lund et al. 2015)

Over the course of DR development, modeling and optimization of different objectives are two of the major aspects to handle. Other aspects such as Information and Communication Technology (ICT) advancement and new energy storage technology (Panda et al. 2022) are also essential for the practical implementation of DR. In recent years, there has been a growing number of reviews about DR and DSM researches with various focuses in different sectors worldwide: (Kohlhepp et al. 2019) evaluate control and information technology as well as the economic and regulatory frameworks for the DR flexibility potential on the demand side based on several field studies. The focus lies on projects

that include a large share of domestic HVAC/P2H (Power-to-Heat) and TES (Thermal Energy Storage)-based DR in the residential sector. In another review of the broad scope of DSM in the residential sector (Panda et al. 2022), the authors focus on the modeling, optimization methods, major objectives and related topics. Regarding related literature, they do not distinguish between simulation and implementation works. (Darwazeh et al. 2022) survey peak load management strategies in commercial buildings, in which the load forecasting models and the impact of peak load reductions on indoor environment have been discussed in detail. In the industrial sector, (Golmohamadi 2022) review DSM in heavy industries such as cement production, aluminum smelting and oil refinery. Employed optimization algorithms and software tools are discussed at length in the review. However, the highlighted results are still based on simulation. In a more recent review, the authors (Ranaboldo et al. 2024) focus on the advances in European markets and regulations regarding industrial demand response (IDR) applications. In terms of different technologies or methods, they emphasize the importance of digitalisation to provide better DR services, but only to a relatively small extent.

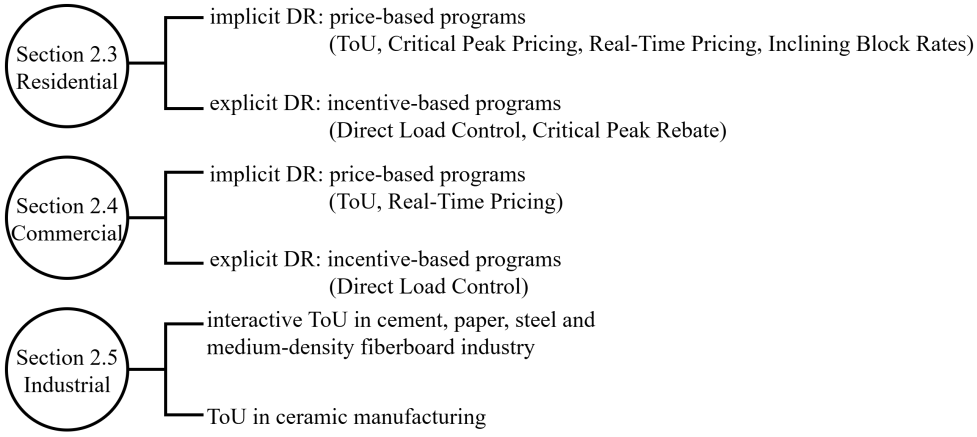


Figure 2.2: A graph overview of the discussed DR concepts in the chapter

Based on the related work and to the best of our knowledge, there is a lack of focus on the survey of new practical DR applications in the existing literature. In the present chapter, we focus on the current state of practical DR implementations in residential, commercial and industrial sectors respectively, especially on their corresponding modeling

and optimization approaches. Fig. 2.2 presents a graph overview of the discussed DR concepts in this chapter and the main contributions of this chapter are:

- The review provides a deeper insight of the utilized models and optimization/control approaches of DR implementations
- Different findings, such as best practices and limitations of specific DR concepts, in residential, commercial and industrial sectors are highlighted in each section separately
- Based on the review results, several improvement suggestions are proposed in the discussion

2.2 Review Methodology

"The quality of literature reviews is particularly determined by the literature search process." as stated in (Brocke et al. 2009). For that reason, the literature search process used in this chapter is described at first to improve the objectivity and transparency. Based on the framework proposed in (Brocke et al. 2009), a review methodology is developed in the following. Firstly, the review scope is to explore the practical experiments and implementations in the field of DSM during the past five years with a focus on different state-of-the-art modeling and optimization approaches. At the second step, different key thematic terms such as 'Demand Side Management', 'Commercial/Residential/Industrial', 'Demand Response', 'Modeling' and 'Optimization' are interchangeably combined as topics to identify the highly relevant open access articles, which usually have broader visibility and citation rates compared to closed access articles. The third step is selecting a literature database and conducting the search. In our work, Scopus is chosen as a main database while it is described as the world's largest abstract database of indexed, peer reviewed, scientific literature (Schotten et al. 2017). By using the backward reference searching, other databases are also utilized. However, it is worth noting that the overview based on the mentioned review methodology is not able to cover every aspect and literature comprehensively and is focused on identifying research gaps from different categorical perspectives, that are, different sectors. The analysis of the searched literature is then carried out based on three different sectors in the fourth step. In the last step, a research agenda with possible improvement suggestions is provided for extending this survey. Fig. 2.3 illustrates the employed review methodology.

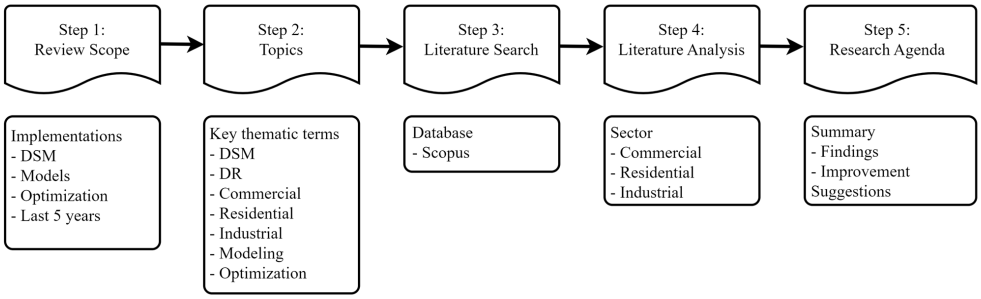


Figure 2.3: Illustration of the review methodology

2.3 Residential Demand Side Management

In the residential sector, the term DR is more often used compared to DSM (Deng et al. 2015, Parrish et al. 2020, Panda et al. 2022). The International Energy Agency (IEA) has defined DR as "... balancing the demand on power grids by encouraging customers to shift electricity demand to times when electricity is more plentiful or other demand is lower, typically through prices or monetary incentives. ..." (IEA 2023). This definition reveals two main mechanisms for DR applications, which are price-based programs as implicit DR and incentive-based programs as explicit DR. These two mechanisms have a direct impact on the modeling and optimization approaches in practical applications which are discussed separately as follows.

2.3.1 Price-based programs as implicit Demand Response

In residential implicit DR applications, different price signals and tariffs are utilized to motivate consumers to shift electricity consumption. Based on the features of the changing price signals, four common tariffs are further subdivided, namely Time-of-Use Pricing (ToU), Critical Peak Pricing (CPP), Real-Time Pricing (RTP) and Inclining Block Rates (IBR) (Deng et al. 2015, Yan et al. 2018). Each of them has practical applications in the residential sector. Fig. 2.4 presents the illustration of these four tariffs. Table 2.1 summarizes the identified real-world implementations of implicit DR under different price-based programs in residential sectors, which will be discussed in more detail below.

Table 2.1: Identified implicit DR applications in the residential sector with a focus on modeling and optimization approaches

Research	Place	Objective Function	Tariff	Modeling	Optimization	Comments
Fraija et al., 2022	Canada	Residential agents cost minimization	ToU	Linear + Quadratic models	Convex	RL for an online implementation
Rahman et al., 2021	Bangladesh	Total financial benefit maximization	ToU	Linear model	Concave	Optimized by GA
Jang et al., 2024	South Korea	Marginal cost minimization	ToU	Linear + Third-order polynomial models	Convex	Conditional logit estimation and instrumental regression
Pelletier et al., 2022	Canada	/	CPP	/	/	Questionnaire based with deployment results
Wang et al., 2023	China	Peaking cost minimization	CPP	Linear-segment model	Convex	Solved by an improved two-layer PSO algorithm
Burkhardt et al., 2023	USA	Generation cost minimization	CPP	Linear estimation model	/	Field experiment
Finck et al., 2020	Netherlands	Operational cost minimization	RTP	Grey-box + Black-box models	Convex	Combination with EMPC framework
Thorsteinsson et al., 2023	Denmark	Cost minimization	RTP	Linear + Second-order polynomial models	Convex	Combination with hierarchical MPC
Bandyopadhyay et al., 2020	USA	Overall peak load minimization	RTP	Linear models	Convex	Identification of the adverse impact of RTP-based DR
Elazab et al., 2022	Egypt	Electricity cost minimization	IBR	MILP model	Convex	With linear constraints
Shaban et al., 2024	Egypt	Net daily cost minimization	IBR	MIQP model	Convex	Under a 7-block IBR tariff

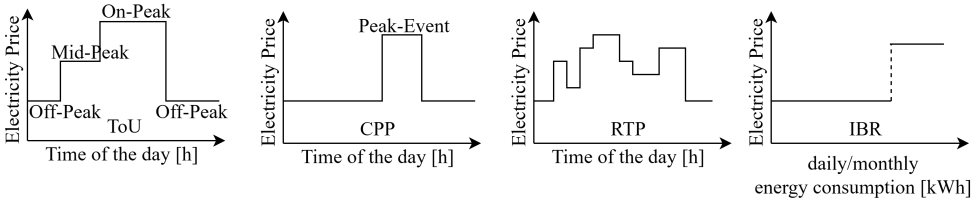


Figure 2.4: Illustration of different price-based programs based on (Deng et al. 2015, Yan et al. 2018)

ToU is a strategy that adjusts tariff rates in different blocks such as a common three-block strategy in Fig. 2.4 within a day. The main purpose of this price strategy is to motivate consumers to shift their loads to off-peak periods for balancing the overall energy consumption and reducing costs. In (Fraija et al. 2022), a data-driven based Demand Response Aggregator (DRA) for generating near-to-optimal Discount-Based ToU (DB-ToU) tariffs has been developed. On the aggregator side, a Reinforcement Learning (RL) technique is proposed for generating DB-ToU tariffs. On the customer side, a typical convex optimization problem is formulated by combining a linear model for indoor temperature and quadratic models for thermal comfort and penalization. In the case study, a multi-agent environment of 20 residential houses, located in Quebec, Canada, has been utilized. The results imply that the generated DB-ToU tariff has the capability to learn the DR of the residential agents and thus is able to mitigate energy consumption peaks and improve load factor. In another application (Rahman et al. 2021), the authors provide a ToU pricing scheme by supplying different consumer groups via two measuring meters, one of which is optimized by Genetic Algorithm (GA) for peak and off-peak energy rates. The objective function w.r.t. the total financial benefit is established with a simple linear model. The developed method is applied to the residential consumers of a practical distribution system in Bangladesh and the results suggest that significant financial savings can be achieved simultaneously by consumers and utilities. More recently, the researchers in (Jang et al. 2024) analyzed the effects of an applied three-block ToU strategies on the load patterns of residential customers and on producer surplus in South Korea from an economic perspective. On the demand side, a linear model with the conditional logit estimation and instrumental regression is implemented for analyzing consumption and load changes at the household level. On the supply side, a general form of a third-order polynomial for calculating system marginal price and the producer surplus is specified. The results according to the real-world application data show that ToU can be an effective measure for shifting electricity consumption from periods with higher marginal cost of

generation (peak) to those with lower marginal cost (off-peak), while simultaneously increasing producer surplus.

Compared to ToU, CPP is another simpler option, which charges a higher rate during peak events of the day, for helping reduce demand on the electric grid during peak hours. In a question-and-answer format article (Pelletier and Faruqui 2022), the authors discuss the implementation experience of a winter peaking utility in Canada. In the focus groups of residential customers, CPP is one of the two preferred options. It's confirmed in practice that a high price signal, CPP in this case, with the potential for greater savings rather than a low-price signal, is preferred. Another practical implementation of CPP is introduced in (Wang et al. 2023). The authors propose a fair residential CPP pricing scheme based on a so-called penalty mechanism of electricity price for high-electricity consumers (PMEP-H). In this scheme, the demand response load model is a simple linear-segment function and three objective functions from different perspectives constitute a convex optimization function. To solve this optimization problem, an improved two-layer Particle Swarm Optimization (PSO) algorithm is utilized for a real residential area in a populated area in northern China. The results show a promising effect of peak cut and increasing renewable energy consumption in that area. Similar load-shifting effects when utilizing CPP are also observed in a two-year field experiment involving 280 Texas households in (Burkhardt et al. 2023). The modeling is carried out at the household level with a linear estimating equation. With the use of appliance-level data, the authors (Burkhardt et al. 2023) also provide the first evidence on the large contribution of air conditioning as DR to the CPP and the potential of electric vehicle charging during off-peak times for managing electricity load with minimal or no discomfort costs.

As one of the dynamic pricing strategies, RTP can provide electricity prices that vary frequently, for instance hourly (Wang et al. 2020). This dynamic feature makes RTP more suitable for a combination with advanced control such as model predictive control (MPC). In a real-life demonstration (Finck et al. 2020), an economic model predictive controller (EMPC) framework under RTP is tested in a residential building for regulating demand flexibility. In this framework, two modeling techniques are utilized for the six main system components. For the heating storage, two grey-box models are used for the modeling. For demand and weather forecasting, different black-box models such as ANN and Markov chains are utilized. The EMPCs are then applied to a convex optimization problem for the minimization of operational costs. The results show that together with the introduced dynamic RTP, EMPC can achieve the lowest operational costs while providing more demand flexibility among the three implemented control strategies. In another four-month long experimental study in Denmark (Thorsteinsson et al. 2023), the authors

present a hierarchical MPC approach for DR under an hourly price model, namely a RTP tariff. Fig 2.5 presents the overview of the control implementation.

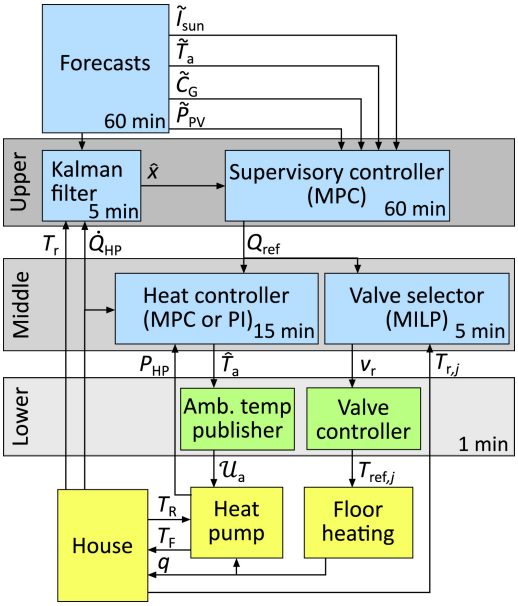


Figure 2.5: Overview of the control implementation, where blue represents computation process, green represents onsite units and yellow represents the heating system and the experimental house (Thorsteinsson et al. 2023)

To reduce the complexity of the control scheme, a mixed integer convex cost function, which contains linear models and second-order polynomial functions, is formulated. The results show that load shifting is feasible and cost-effective based on the proposed approach. Regarding the Danish price scheme, it is further established that the evening peak is the decisive cost factor, and heat pump operations should be blocked during this evening peak time. Although RTP is the most dynamic tariff, it cannot conclude that RTP would be "cure-all" solutions to high peak demand issues. In addition to the advantages of RTP, the adverse impact is also identified. According to the analysis results on a community of 100 single-family houses in Austin (Bandyopadhyay et al. 2020), RTP-based DR programs can effectively shift the residential peak away from the time of overall peak load but increase the magnitude of the residential peak load by incentivizing customers to concentrate usage within the low-price hours. To support this conclusion, a convex optimization model is developed. In this optimization model,

different submodels as constraints are contained, such as a linear one-parameter thermal model for modeling the HVAC system and the electric water heater. These submodels involve energy conservation around the home and the bounds for the room as well as the water temperature, which have a decisive impact on the optimization and the following analysis along with the conclusion.
Other than the three tariffs mentioned above, which are based on the time of the day (x-axis in Fig. 2.4), IBR provides different, usually two, price blocks depending on the customers' consumption. For instance, a power utility in Canada, BC Hydro, offers a residential IBR with two blocks. As of April 1, 2024, a higher block price is almost 40% more than the lower block price when a bi-monthly threshold of 1,350 kWh is exceeded (BC Hydro 2024). This simple structure has resulted in a widespread adoption of IBR by many power utilities since the 1980s (Deng et al. 2015), but this may also be one of the reasons that it has ceased to be a focus of DSM applications research in the recent years. IBR is usually treated as a known and unchanged input in the context of DSM. In (Elazab et al. 2022), a three-time-frame energy management scheme under IBR is proposed and applied to a house in Egypt. To reschedule the load and minimize the electricity cost, a mixed-integer linear programming (MILP) model is developed. The results show that reducing AC (air conditioner) power affects significantly the energy cost in a positive way under the Egyptian IBR tariff. Also for the load scheduling problem in Egyptian households, a more detailed application under a 7-block IBR with residential customer-based data is investigated in (Shaban et al. 2024). The authors first design a questionnaire to survey the characteristics of the Egyptian household appliances' usage in both cities and villages. To address the household load scheduling problem within a time horizon of 24h, a mixed-integer quadratic programming (MIQP) model is developed that aims to minimize the net daily cost. The findings underscore that the presence of PV systems together with the proposed MIQP model can significantly reduce electricity bills under IBR.

Focused on the modeling and optimization approaches, we derive the following findings:

- RTP is the most dynamic tariff and thus suitable for a combination with complex optimization and control methods such as MPC
- RTP is not the "cure-all" solution to high peak demand issues in the residential sector
- Linear models can already achieve a lot of improvement despite their simplicity
- Most of the identified applications make use of convex optimization regarding cost or energy consumption minimization

- Advanced techniques such as RL, GA, PSO can help improving the optimization performance
- IBR has not been the focus of research-oriented implementations in recent years

2.3.2 Incentive-based programs as explicit Demand Response

In incentive-based programs or incentive-based demand response (IBDR) (Chen et al. 2021b, Gupta et al. 2022), the power utility or aggregators request a certain amount of balancing power or power reduction explicitly, and the participants on the demand side are compensated with incentives for it (Zheng et al. 2020, Dewangan et al. 2023). Based on the identified explicit DR applications in the residential sector, we summarize the results in Table 2.2 with detailed discussion in the following.

When IBDR is implemented in a centralized scheme, it's called direct load control (DLC) (Dewangan et al. 2023), which has been offered to residential and small commercial users for decades (Deng et al. 2015, Davarzani et al. 2021). The ongoing research of DLC in DR is more focused on developing optimal DLC scheduling models and implementation mechanisms (Ozkop 2024). The authors in (Hlalele et al. 2021) propose a model, that uses DLC to find an optimal strategy which minimizes generation costs and maximizes renewable utilization. In the considered DR program, the electric water heaters from residential customers can be directly controlled by the system operator at the substation level. Specifically, a multi-objective model is formulated. To solve the non-convex and non-linear optimization problem, the ε-constraint method is used to turn the multi-objective functions into a single objective function, which finds the best compromise solution in a Pareto optimal set. Specifically, the ε is the Pareto solution set as constraints in the objective function that is increased from the minimum to the maximum value to generate a Pareto front. The results show that the introduction of DLC improves the level of renewable energy utilization and significantly reduces the system total operating cost. In another application involving DLC for 200 residential households (Saleh et al. 2022), an energy not-served (ENS)-based assessment method for evaluating smart grids (SG) functions operating residential loads is proposed and validated. The method uses a linear combination to calculate the reduction in power demands during peak-demand times. Actions of DLC in the implementation are initiated as ON/OFF

actions for certain thermostatically controlled appliances, which are also formulated with linear combinations.

Table 2.2: Identified explicit DR applications in the residential sector with a focus on modeling and optimization approaches

Research	Place	Objective Function	IBDR	Modeling	Optimization	Comments
Hlalele et al., 2021	South African	Total cost minimization	DLC	Multi-objective nonlinear models	Non-convex	ε-constraint method
Saleh et al., 2022	Canada	Peak load reduction	DLC	Combination of linear models	/	Actions of DLC are initiated as ON/OFF
Tomat et al., 2022	North America	/	DLC	Building Envelop	/	An analysis of the behavior of occupants after DLC events
Jacobsen et al., 2022	USA	/	CPR	Linear regression models	/	Randomized Controlled Trial (RCT) evaluations
Kapeller et al., 2023	Austria	/	CPR	Log-linear models	/	New variant of CPR
Hofmann et al., 2024	Norway	/	CPR	Linear fixed effect models	/	Randomized Controlled Trial (RCT) evaluations

However, the results in this study show that DLC could not provide a load-side control to reduce the power demands during peak-demand times as effective as other two implemented control methods viz. PDM (peak-demand management) and DR (demand response). Similar results regarding the efficacy of the DLC strategy are obtained in (Tomat et al. 2022). The authors create an envelope model on EnergyPlus and use the extensive real data from residential users, who are willing to share their data, to validate the model. The observed results by clustering show that the efficiency of DLC is strongly impacted by different utilization patterns. Some behavior such as a high setpoint decrease after override can reduce or even nullify the efficacy of the DLC strategy.

In contrast to DLC, another common strategy for IBDR in residential sector, namely the critical peak rebate (CPR) or peak time rebate (PTR), allows consumers to determine the demand reduction by themselves and provides them incentives for reducing usage during critical hours below a baseline level of consumption (Parrish et al. 2020, Chen et al. 2021b, Jacobsen and Stewart 2022). In a field experiment run by a vertically-integrated electric utility in the western U.S., the authors (Jacobsen and Stewart 2022) investigate nine different pricing treatments for reducing demand during critical events. The nine pricing treatments include three CPR-only, three ToU-only, and three ToU + CPR hybrids. In this process, linear regression models are utilized for the randomized controlled trial (RCT) evaluations of household energy consumption, which are based on the post-period data. The post-period data employ controls that are computed based on pre-intervention consumption. They found that the implementation of CPR alone reduces more demand in residential sector than the combined implementation of CPR and ToU. The results suggest that more complex pricing systems can overwhelm residential consumers and result in unexpected consumer decision-making in DR. In another full-scale experiment in Norway, the authors (Hofmann and Lindberg 2024) also utilize the RCT method for their implementation and the experimental setup resembles critical rebate programs. For the analysis of the households' demand response, four different linear econometric fixed effect models with individual fixed effects are formulated. In fixed effect models, the fixed effects account for unobserved characteristics that do not change over time within entities, such as living space of the household in this case. An average demand reduction of 2.92% in peak hours are observed in the experiment. The results also suggest that a shorter peak period results in a higher response and a statistically significant increased response can first be observed when the rebate is above a threshold value. In addition to the conventional CPR strategy as mentioned above, a new variant of CPR has been implemented in a field trial in Austria (Kapeller et al. 2023). Instead of providing residential consumers

incentives in peak demand hours, participants receive rebates when they remove the consumption to local peak generation hours. The authors utilize log-linear models to analyze the treatment effects viz. consumption change and load shifting. They find that on average household electricity consumption is slightly increased by 5.7% during peak generation hours and peak loads are lowered during peak demand times between 5% and 9%.

Apart from the two above mentioned common IBDR strategies in the residential sector, some other incentive-based programs such as Capacity Markets (CM), Demand Bidding/Buyback, Ancillary Service Markets (ASM), Interruptible/Curtailed Load and Emergency Demand Response have also been summarized in the literature (Nikolaidis and Poullikkas 2020, Lotfi and Ghazi 2021, Davarzani et al. 2021). However, in the course of our literature review, we found a lack of implementations with these IBDR strategies in the residential sector over the last 5 years due to different reasons. For instance, ASM refers to the capacity reserved market with a range of functions which transmission system operators (TSO) contract so that they can guarantee system security. Thus, their participants on the demand side are predominantly large and regular energy consumers (Davarzani et al. 2021), such as commercial EV fleets (Owens et al. 2024), a frequency containment reserve (FCR) market (Strömer et al. 2022) and even the power grid of a province (Ding et al. 2023).

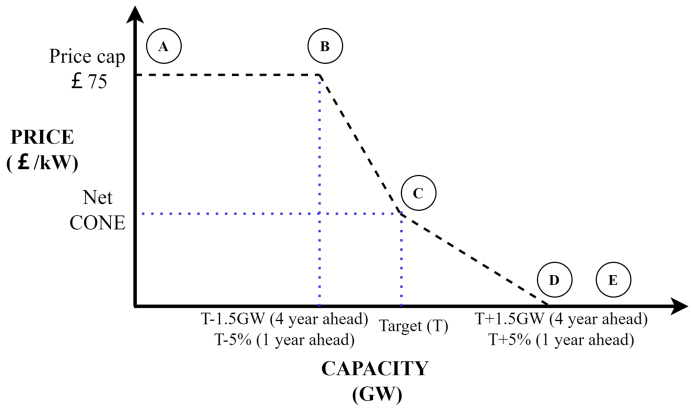


Figure 2.6: Illustrative capacity demand curve, A-E stand for different key points regarding capacity auction (DECC 2014)

Similar to ASM are the Demand Bidding/Buyback and CM programs. The former usually encourages large electricity consumers to provide load reductions at a price at which they are willing to be curtailed (Hadayeghparast and Karimipour 2020), e.g., large industrial consumers in Brazil (Dranka et al. 2021). The latter requires not only a high level of capacity but also runs over medium to long time periods. Fig. 2.6 shows an illustrative capacity demand curve under CM programs based on (DECC 2014) that is released by Department of Energy & Climate Change in the UK. It can be seen from the figure that CM has the level of capacity in GW and a time period from 1 to 4 years.

In order to deal with power outages and unexpected events in power grids, Emergency Demand Response programs are regarded as a potential method (Fan et al. 2021) in restoration. However, Emergency Demand Response programs also usually target large and industrial consumers (Cui et al. 2020, Lee et al. 2023). As for Interruptible/Curtailed Load, a recent energy economics research in Ireland states that on average, consumers are found to be mostly indifferent to curtailable contracts compared to their status quo contract (Harold et al. 2021). This could be one of the reasons that this IBDR strategy has not been the focus of relevant implementations in the last 5 years.

Focused on the modeling and optimization approaches, we derive the following outcomes:

- In the residential sector, DLC and CPR/PTR have recently gained more attention in research-oriented implementations compared to other IBDR strategies
- The focus of DLC implementations is developing optimal scheduling models and analyzing treatment effects independent of specific optimization approaches
- An impact evaluation such as RCT evaluation in the context of economics is often to be observed in field experiments and practical implementations of CPR
- For economic impact analysis, linear regression models and fixed effects models are often employed
- Other IBDR strategies, such as Demand Bidding/Buyback, CM, ASM and Emergency Demand Response programs, are more suitable for large and regular energy consumers such as industrial consumers

2.4 Commercial Demand Side Management

Similar to the residential sector, energy consumption in the commercial sector such as office buildings, warehouse and education buildings is also quite significant. According to the U.S. Energy Information Administration, the residential and commercial sectors accounted for about 19.7% and 17.2% respectively of total U.S. energy consumption in 2023, when electrical system energy losses are included (EIA 2023). However, unlike small residential loads, multi-zone commercial systems, e.g., HVAC and building automation system (BAS) could offer much higher potential for load reduction (Darwazeh et al. 2022). Besides, more commercial buildings per capita would lead to higher consumption levels (González-Torres et al. 2022), which makes it worthwhile to implement DSM in the commercial sector as well. Table 2.3 presents the identified DSM implementations in the commercial sector with more details in the following discussion.

As identical to the residential sector, price- and incentive-based DR programs that are mentioned in Section 2.3.1 are also the core of DSM applications in the commercial sector when considering cost savings. Additionally, advanced control strategies such as MPC are proven to have huge advantages for load reduction. For instance, a long-term performance evaluation of the MPC controller under ToU tariffs in a small commercial building is carried out in (Zhang et al. 2022). In this real-world operation, the authors first use the lumped resistance and capacitance approach with an R8C4 network, as shown in Fig. 2.7, to model the building thermal zones. The overall building thermal model is a fourth-order linear model. The supervisory controller in the control system is a MPC controller which optimizes the temperature setpoints in different zones, appliances and the charge/discharge power of a battery with a time interval of 5 minutes. To minimize the monthly energy cost, a convex optimization problem with an optimization horizon of 24h is formulated. The field test results show that in overall 11.7% savings of the annual energy cost are achieved compared to the baseline viz. business as usual. Moreover, the average peak reduction after utilizing MPC is 34%.

Table 2.3: Identified DSM implementations in the commercial sector with a focus on modeling and optimization approaches

Research	Place	Objective Function	Tariff	Modeling	Optimization	Comments
Drgoňa et al., 2020	Belgium	Energy use and thermal discomfort minimization	/	NL white-box model decoupled into a Hammerstein-Wiener model structure	QP	MPC combined with a NL optimization algorithm as post-processing
Freund et al., 2021	Germany	Heating energy demand and thermal discomfort minimization	/	R7C4 / fourth-order linear model	QP	A comparison between MPC and RBC
Chen et al. 2021	China	Economic benefit maximization	ToU	MDP model	RL	A non-deterministic way by using RL to deal with DR events
Zhang et al., 2022	USA	Electricity bill minimization	ToU	R8C4 / fourth-order linear model	Convex	MPC as the supervisory controller
Blum et al., 2022	USA	Daily energy consumption minimization	/	R2C2 gray-box model	Convex	MPC as the main controller
Xiong et al., 2023	China	Energy consumption and cost minimization	RTP	Linear models	Convex	A combination of an improved TC and ATES
Saxena et al., 2023	Canada	/	/	Linear models	/	A V2X program caters to participate in DR
Wang et al., 2024	China	DR target tracking viz. fixed 20% load reduction	/	/	/	Modified PSBC for coordinating TCLs during DR events

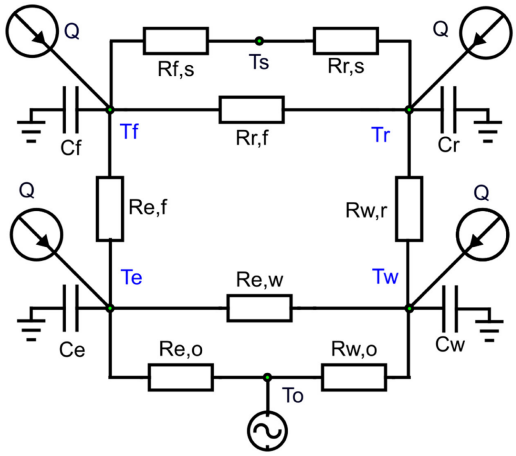


Figure 2.7: The R8C4 network to model the building thermal zones. R, C and Q indicate resistance between the connected temperature nodes, capacitance, and net heat addition to the nodes respectively (Zhang et al. 2022)

In another real-world implementation using an MPC controller to control the HVAC system of a commercial office space of approximately 6000 m^2, the authors (Blum et al. 2022) demonstrate usage of an open-source tool-chain called MPCPy (Blum and Wetter 2017). To reduce the complexity of the models, a zone-simplification scheme is applied for modeling, and RC gray-box models of thermal envelope zones with persistent estimated parameters are developed in Modelica. Besides, forecasting models with different approaches are developed as part of the implementation, such as k-Nearest Neighbors (kNN) for weather forecasting and a simple average algorithm for internal load forecasting. The optimal control problem is then defined to minimize the daily energy consumption viz. the prediction horizon is 24 hours. To ensure feasibility of the optimization, three slack variables as constraints, together with corresponding weights, are introduced for the zone temperature and supply air temperature, instead of hard constraints. This is typical for the case when the outside air temperature is too low to maintain the minimum supply air temperature. The promising test results reveal an estimated 40% energy saving by operation with MPC compared to operation without MPC over a two-month period, without significant penalty to thermal comfort.

The effects of load reduction on the demand side by implementing MPC is not just observed in traditional commercial buildings. In newly built energy-efficient commercial buildings, the improvement is also significant (Freund and Schmitz 2021). The authors demonstrate a three-month field test by operating MPC for heating energy demand and

thermal discomfort optimization in a modern large-sized, energy-efficient office building in Germany. In this process, a gray-box model with a R7C4 network for each zone, i.e., a fourth-order model, is developed in Modelica at first. To minimize the heating energy demand and thermal discomfort, a constrained non-linear optimization problem is defined. The energy-efficient building is heated by thermally activated building systems in the form of thermoactive ceilings. Thus, there is only one control input, namely the supply temperature of the corresponding thermoactive ceilings heating circuit. Compared to the default rule-based control (RBC), the 3-month field test results show a relative demand reduction of 19.2%, 24.6% and 75.2% in February, March and April respectively after implementing MPC. However, due to the slow-reacting dynamics of thermoactive ceilings, an improvement in thermal comfort could not be clearly identified in this implementation.

Similarly, in a cloud-based implementation of MPC for an office building in Belgium, the authors (Drgoňa et al. 2020) also first develop the building models in Modelica. However, unlike the RC gray-box models mentioned above, a detailed white-box non-linear model in Modelica IDEAS library is developed at first. Then, they use linearization and decoupling technique to form a Hammerstein-Wiener model structure which lowers the complexity of the control oriented building model. Fig. 2.8 presents the general model structure with the decoupling principle. In this implementation, a large GSHP coupled with two buffer tanks of 1000 liters each in the office building is the main production system for thermal comfort and also the main focus of the MPC. The purpose of the MPC is to minimize energy use and thermal discomfort, which in this case makes use of a quadratic optimization (QP) problem. Since in a practical setup some variables such as heat flows cannot be directly manipulated, the authors also introduce a post-processing algorithm to solve a non-linear optimization problem for control variables based on the heat transfer equation. The real-world operational results over 30 days show that the MPC operation can not only save 53.5% energy usage but also improve thermal comfort by 36.9% at the same time compared to the base line, namely the RBC operation.

In addition to the extensive use of MPC for load reduction in the DSM, other control strategies have also been implemented and evaluated in the commercial sector in recent research. For instance, to evaluate the performance of Priority Stack-Based Control (PSBC) for coordinating thermostatically controlled loads in real office buildings, (Wang et al. 2024) conduct a 22-day field test with the vanilla and modified PSBC strategies. They first define two distinct DR scenarios, each with a 2h duration, for weekdays and

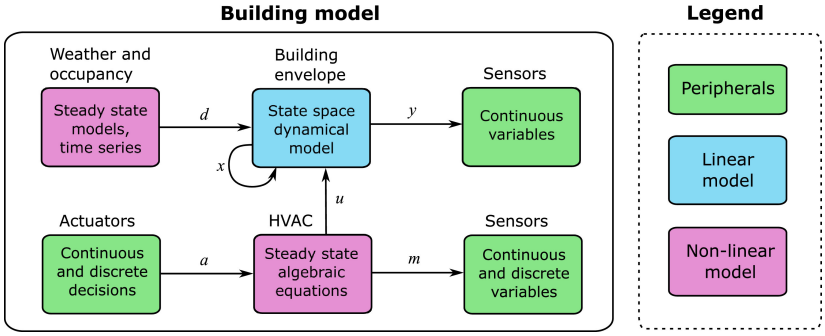


Figure 2.8: A general building model structure with decoupled non-linearities (Drgoňa et al. 2020)

weekends due to different energy patterns. Instead of optimizing load profiles directly, they establish a DR tracking target, namely a 20% reduction in power consumption compared to the baseline. This target value is utilized in the PSBC to be compared with the instantaneous aggregate power at each sample time and then to generate the regulation signal to coordinate thermostatically controlled loads . To quantify imminence, i.e., priority, the temperature distance to the lower/upper temperature bound is measured. In the vanilla PSBC strategy, the tracking target is constant, while in the modified PSBC strategy, a tracking target modified mechanism is introduced. It should be noted that this mechanism in the modified PSBC strategy is not to change the DR tracking target but to adjust the tracking power consumption. The results reveal that both strategies can optimize load profiles in terms of tracking target during DR events. Besides, the performance of the modified PSBC strategy has a notable 17.9% improvement in tracking accuracy and a significant 38.8% improvement in thermal discomfort over the vanilla PSBC strategy.

In another on-site experiment within a teaching building, (Xiong et al. 2023) test the effect of a comprehensive DR strategy that combines an improved version of transactive control, also known as market-based distributed control, and active thermal energy storage for energy storage AC systems under RTP. The core idea of this combined DR strategy is to use the two market value signals of bidding price and RTP to regulate the temperature set-points of the active thermal energy storage dynamically, where the bidding price is calculated by linear modeling on the basis of the real pricing during the past 24h. Then the authors develop an algorithm of the strategy in the form of state flow with slight differences for cooling and heating conditions. In the experimental set-up, the active thermal energy storage consists of an air source heat pump (ASHP), an energy storage tank, a variable air volume box, etc. Besides, for different conditions in summer

and winter, a 5-day experiment for cooling and a 3-day experiment for space heating are carried out separately. The results suggest that the proposed DR strategy could achieve the largest reduction in total energy consumption and cost compared to the baseline operation and other strategies such as transactive control alone or active thermal energy storage alone. Furthermore, the indoor thermal comfort is better ensured with the combined strategy. However, the advantage of the proposed DR strategy is greater in the space cooling experiment than in the heating experiment. The reasons for this phenomenon are not discussed in (Xiong et al. 2023).

Instead of the deterministic way such as modeling the DR strategy as an optimization problem, non-deterministic methods are also suitable for DR operation. (Chen et al. 2021a) propose a RL framework to formulate the optimal DR strategy and demonstrated the experiments of the proposed program for a commercial building-based virtual power plant (VPP) platform. In this framework, the authors first model VPP operation by using a Markov decision process to map the VPP state status and control action, and then formulate the monetary economic benefit based on the utility function (Pourbabak et al. 2017) with customization. The customization is designed for different commercial buildings and satisfies non-decreasing but marginal non-increasing characteristics (Mohajeryami et al. 2016). Considering the limited computational capability of installed control unit for end-users in real-world, the authors choose the simple action clipped Q-learning algorithm for training, in which the action choices are bounded or clipped by the backup power limitation. The implementation is carried out in 20 commercial buildings of different types with HVAC as the main energy consumer, in 15-min time intervals for 13 to 15 days. During this period, a ToU structure of a local electricity price scheme in Shanghai is executed, which contains DR incentives for peak and valley periods. The authors conclude that all the participating buildings increase their economic benefit between 1.2% and 12.5% compared with original gain. It proves the ability of the proposed RL method to benefit, in a non-deterministic way, the long-term operation of commercial building-based VPP with DR.

In addition to mainly optimizing the control of HVAC systems in DSM as discussed above, the urgent transition towards electric vehicle (EVs), especially the bidirectional charging ability or “Vehicle to everything” (V2X), offers another opportunity for energy management on the demand side. In a real-world demonstration in Canada, (Saxena et al. 2023) first design a survey for EV owners to consider their preferences as the pre-step for

the V2X program design. These preferences, among which the most important are owner convenience and the opportunity to earn revenue via V2X, decide which parameters are considered in the implementation. In commercial buildings, the building owner acts as an aggregator of EVs and other DERs, who dispatches the setpoints during DR events to each EV based on the signals from the power system operator feeder (Kim et al. 2021). Finally, the strategy is formulated by linear models along with linear constraints. In a 5-day field test, the EV used for testing has a maximum charging power of 6.6 kW and a maximum discharging power of 30 kW. Fig. 2.9 presents the single line diagram of the commercial building under test. The feasibility of the proposed program is validated in a real-world demonstration that achieves an averaging 30.4 kW of load reduction during a 2 hour DR event under a contracted DR capacity of 30 kW. In the meantime, the actual consumption of the building is reduced by 60.7 kWh and the EV owner earns $2.17/kWh by discharging during the DR event.

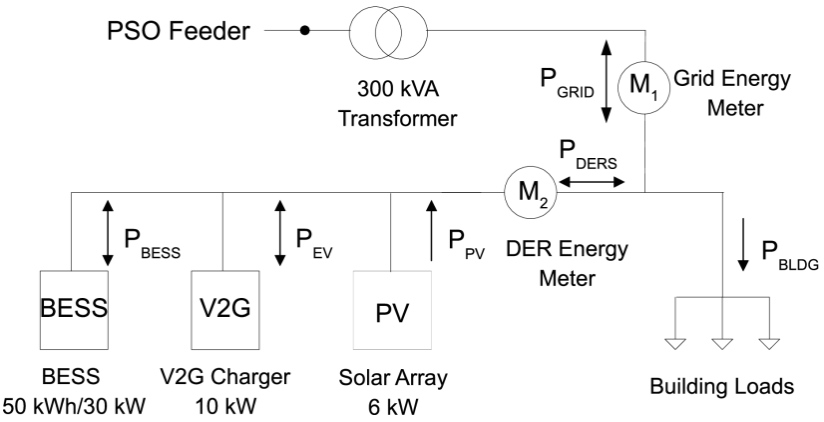


Figure 2.9: A single line diagram of the commercial building together with EV and other DERs under test (Saxena et al. 2023)

As summarized in Table 2.3, it's clear to see that advanced control methods such as MPC are more extensively used regarding optimization in the commercial sector. Focused on the modeling and optimization approaches, the following can be derived:

- RC gray-box models of the building are commonly combined with MPC in the implementation
- Both deterministic and non-deterministic ways of optimization are applicable for DSM

- Simple models such as linear models with linear constraints can already bring significant improvements in the DR implementation regarding load and cost reduction
- HVAC systems are the main object of modeling and control in most implementations due to its high energy consumption
- V2X technology has a high potential as an important part of DSM in the commercial sector
- Long-term implementation needs to be further researched

2.5 Industrial Demand Side Management

Unlike the applications and implementations in the residential and commercial sectors, most industries that take part in DR programs are participating through ancillary services programs (Shoreh et al. 2016), which have different priorities compared to those in other sectors. For instance, ancillary services are typically needed year round not just during peak hours (Shoreh et al. 2016). Besides, due to the greater impact of the industrial sector on the grid, the technical barriers such as open standards are necessary to be investigated. Although standards like OpenADR (OpenADR 2011) and ZigBee (Deese and Daum 2018) have been adopted in several industrial applications, the development of associated standards for DR interoperability is still ongoing. However, these open standards are not the focus of this review. In this section, we discuss the unique characteristics and improvements of DR in the industrial sector regarding modeling and optimization in recent years.

In order to improve the traditional structure of ToU tariffs in the industrial sector, in which the utility determines and communicates peak, mid-peak and off-peak hours to industrial customers, (Kholerdi and Ghasemi-Marzbali 2021) propose a modified method for determining the ToU based on the interactive structure. In this interactive ToU tariff, the utility determines the peak months and hours for one year based on the obtained yearly grid load curve. Meanwhile, the voluntary industrial customers select the off-peak months and hours by themselves. By introducing the price elasticity of demand (Aalami et al. 2010), a linear model of demand response is built which defines the relationship between the load and the price in each hour of each month. In the implementation, eight large industrial companies covering four businesses (cement, paper, steel and MDF) have participated. The results reveal the improvement of the proposed interactive ToU.

Compared to the conventional ToU, the average daily load reduction during peak hours has improved from 5.49 MW to 13.59 MW, which equals a reduction of 18.6% instead of 7.5%.

In another demonstration in a ceramic manufacturing company, (Ma et al. 2020b) design a data-driven sustainable intelligent manufacturing architecture based on DR. This framework contains three layers: a perception layer, a management layer and an application layer. In the implementation, the perception layer is responsible for real-time monitoring of energy consumption, i.e., 1 minute in this case, and the other two layers are dealing with energy efficiency assessment and optimization respectively. For ceramic manufacturing, the authors have developed linear models for electrical energy peak load shifting based on a typical serial production line with n shop-floors. However, the model in each floor that considers the time of peak load shifting is non-linear. They utilized PSO for searching the optimal time of peak load shifting in order to minimize the average energy cost. In a 3-day comparative demonstration under a typical ToU tariff, the optimized ball mill consumes 15% more energy with only 7.6% more cost, which makes the average cost 6.5% lower after optimization.

Aside from these two implementations mentioned above, as the review progressed, we found that almost all published research articles in the last 5 years regarding DR implementation in the industrial sector are based on simulations or numerical results. For instance, the authors of (dos Santos et al. 2023) investigate 53 DR application studies in industrial scenarios from 2013 to 2022. However, all 16 studies in the recent years, i.e., covering from 2020 to 2022 in their review, are carried out only with simulations or numerical analysis. This review result also lets us conclude that there is comparatively few literature that focus on DSM implementation compared to the works that focus on theoretical and numerical analysis.

Possible reasons for this finding are, for example, the efforts and costs to model and optimize a large industry equipment are too high for a research work and the companies care mostly on costs and profits (Ma et al. 2020b). Implementations in the industrial sector may involve confidentiality issues and conflict of interest (dos Santos et al. 2023). Besides, there is a variety of barriers, such as informational barriers, slowing participation in industrial DR (Scharnhorst et al. 2024), which could result in increasing the difficulty of drawing findings in industrial DR implementations. Other barriers to implement DR

in the industrial sector are well summarized in (Shoreh et al. 2016, Lashmar et al. 2022, Scharnhorst et al. 2024).

Based on the results above, our findings for industrial DSM are summarized as follows:

- Dominant industries that have high potential for implementing DSM are energy-intensive industries such as cement, metals, mineral products, etc.
- The number of research-oriented implementations of industrial DSM in the last five years are very limited.[1]
- Linear models at a high level, i.e., without consideration of detailed physical behavior of specific equipment, can already contribute to industrial DSM remarkably
- For medium- and small-sized industrial companies, it's important to ensure that the cost savings of implementing DR is higher than updating equipment such as for real-time monitoring

2.6 Discussion

DSM as one efficient strategy for exploiting load management on the demand side can not only bring financial benefits for customers in different sectors, but also contribute to stabilizing the grid as renewable energy sources become more prevalent. Moreover, with the development of ICT on one side and modeling techniques as well as optimization algorithms on the other side, it is important to investigate research-oriented implementations of DSM to provide deeper insights into its potential.

Based on the review results, we find that the active research-oriented implementations in the last five years have focused on the residential and commercial sector whereas the works in the industrial sector are mostly simulation- or numerical analysis-based. In the residential sector, RTP and ToU have attracted more research attention in the implicit DR compared to the other two tariff schemes. This result is supported by their high dynamics, especially RTP, which makes them more suitable for the combination with novel optimization methods. However, it cannot be concluded that RTP is the "cure-all"

[1] For that reason, we can not provide a summarized table

solution in all situations. As for the explicit DR in the residential sector, DLC and CPR have gained more research attention compared to other IBDR schemes. In the commercial sector, the RC gray-box models are extensively used due to its good trade-off between the required modeling efforts and accuracy. Moreover, the combination with MPC is a common choice for implementation. In addition, HVAC systems are still the main object of the implementation in the commercial sector due to its high energy consumption. In the industrial sector, most of the literature from the last years focused on simulations or numerical analysis, which makes the number of implementation-related literature much less than theoretical works. Lastly, it is worth noting that there is a common feature of findings in the different sectors, namely, simple models such as linear models can already achieve a lot of improvement in the DSM implementations.

While this chapter has provided a comprehensive survey of recent practical DSM applications in the literature, several limitations must be acknowledged. First, the selection of a literature database inevitably reflects a certain degree of subjectivity. The focus was placed primarily on studies published in English and in peer-reviewed journals in Scopus, which may exclude relevant insights from other databases or languages. In addition, the review concentrated mainly on modeling and the corresponding optimization approaches, while other important aspects during implementation, such as ICT technologies and standards, were only briefly addressed. This selective concentration may limit the completeness of the overall picture and some innovative findings in several practical applications. Future review should therefore aim at integrating a broader set of databases, even in languages other than English, and optimizing the review methodology to gain access to richer information, such as in the industrial sector, that cannot be obtained through conventional methods.

As for future implementation, we propose several improvement suggestions with a focus on modeling and optimization based on our findings. First of all, it would be interesting to investigate the impact of models with different complexities, which constitutes the main focus of the following chapters of this thesis. In the review process, we have seen that both complex models and simple models can bring remarkable benefits in DSM. The second suggestion is to widen the focus of real-world implementations, especially in the commercial sector, from primarily HVAC to also include V2X technologies. The third one is to investigate the reliability of the models and optimization methods in a long-term implementation. Most of the implementation works that we have reviewed have been evaluated for a duration between several days and one month.

3 Hypothesis and Quantification

Parts of this chapter are adapted or reproduced from the author's own previous works:

Chang Li, Kevin Förderer, Tobias Moser, Luigi Spatafora and Veit Hagenmeyer.
"Gossen's first law in the modeling for demand side management: a first heat pump case study".
In: *2023 Springer Energy Informatics Academy Conference (EI.A 2023)*, 2023, pp. 111–125. doi: 10.1007/978-3-031-48652-4_8.

This chapter systematically addresses RQ1 and RQ2: "How to quantify the complexity and the utility of DER models in the context of DSM?", based on the hypothesis, that the famous Gossen's First Law in economics applies in the domain of DSM with DERs. The first section introduces the theoretical foundation of Gossen's Law and the DERs, namely heat pump systems and electric vehicle charging equipments, that are the reference systems for the validation in the following Chapter 4. Then the exact hypothesis is proposed in Section 3.2. The following Section 3.3 discusses different methods for quantifying model complexity and model utility within the considered domain and indicates as well as explains the methods chosen for the validation, which is conducted in Chapter 4 with three different use cases.

3.1 Theoretical Foundation

This section aims at introducing the necessary theoretical foundation for the following sections. This includes Gossen's Law in economics, the working principles of heat pump systems and the characteristics of the charging process for electric vehicles.

3.1.1 Gossen's Law

In 1854, Hermann Heinrich Gossen has introduced three laws of economics in his book *Entwickelung der Gesetze des menschlichen Verkehrs, und der daraus fliessenden Regeln für menschliches Handeln* (Gossen 1854), which are also known as Gossen's Law. Gossen's first Law describes the law of diminishing marginal utility, which states that marginal utility decreases across an increasing quantity of ranges relevant to decision-making. This range could be consumption of goods (e.g. food) or services. In this relationship, there is a so-called saturation point for marginal utility. It decides whether marginal utility is positive or negative. When the quantity of ranges is less than the value of the saturation point, marginal utility is positive. Mathematically, the following equations (Gossen 1854) are used to formally characterize this relationship for use cases of only positive marginal utility, where U is the quantified utility and C is the quantified consumption. The theoretical relationship for use cases of positive marginal utility is illustrated in Fig. 3.1.

$$\frac{\partial U}{\partial C} > 0 \tag{3.1}$$

$$\frac{\partial^2 U}{\partial C^2} < 0 \tag{3.2}$$

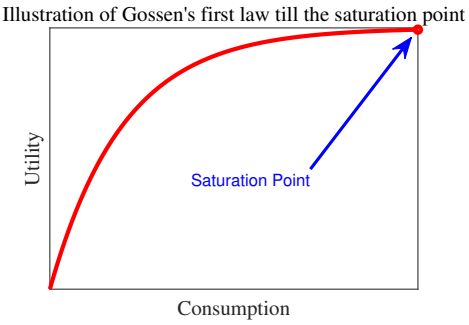


Figure 3.1: Illustration of Gossen's first law till the saturation point, given quantification

Gossen's second law is also known as the law of equi-marginal utility, which indicates that the optimum satisfaction, i.e., utility of a combination of several commodities is reached when marginal utility relative to its price is equal across all commodities. This law applies for use cases that contain at least two commodities. Formally, it can be represented by Equation (3.3) (Gossen 1854), where P is the price and (i, j) represents different commodities.

$$\frac{\partial U/\partial C_i}{P_i} = \frac{\partial U/\partial C_j}{P_j}, \forall(i,j) \tag{3.3}$$

Unlike the mentioned laws above, Gossen's third law is a defined concept, which indicates that the subjective scarcity is source of economic value. This concept, however, is difficult to be formally described by mathematical equations.

3.1.2 Heat Pump Systems

In the residential sector, heat pump systems usually contain two major subsystems: a heat pump and a hot water storage. The operating principle of a heat pump used for house space heating and hot water supply is based on a reverse Carnot thermodynamic cycle (Sarbu and Sebarchievici 2014), which means the thermal energy at the output consists of two parts, i.e., environmental energy and operating energy, namely the consumed electricity. Fig. 3.2 shows the basic operating principle, based on (Bundesverband Wärmepumpe e.V. 2019), of a typical heat pump system in the residential sector.

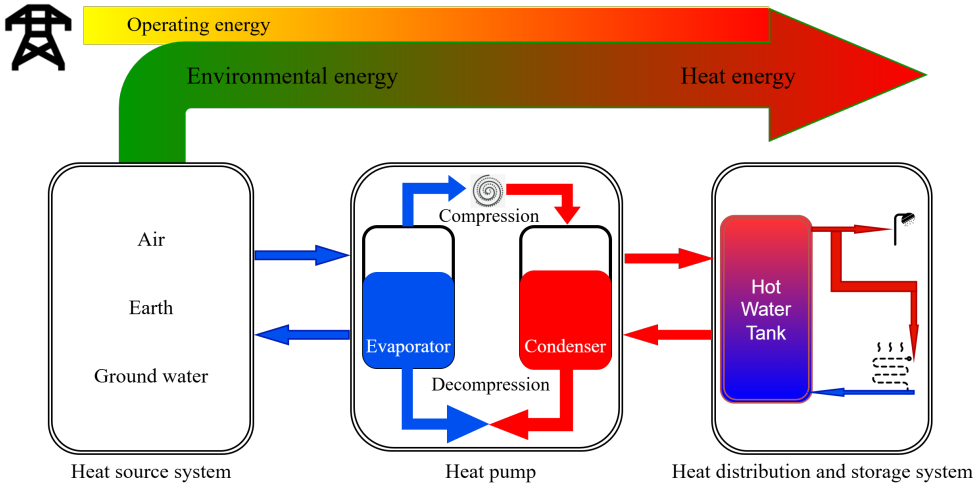


Figure 3.2: Operating principle of a typical residential heat pump system based on (Bundesverband Wärmepumpe e.V. 2019)

In the operation process for heating, the heat exchanger extracts environmental energy from heat sources to the brine, which is then delivered to a refrigerant by another heat exchanger. The compressor uses most of the electricity to move a refrigerant through a refrigeration cycle. The

heat is then passed on to the home heat distribution or/and storage system through another heat exchanger. Based on the different heat sources for the environmental energy, heat pumps are further subdivided into Air Source Heat Pump (ASHP), Ground Source Heat Pump (GSHP) and Water Source Heat Pump (WSHP).

In addition to the heat pump itself, the integration of a hot water storage, i.e., thermal energy storage plays a pivotal role in enhancing operational flexibility, particularly for addressing the intermittency of renewable energy sources. By decoupling energy production from demand, thermal energy storage enables efficient load-shifting and stabilization of grid interactions. There exist a variety of storage systems for thermal energy, such as hot water tank, Phase Change Materials (PCMs)-based thermal energy storage system, seasonal thermal energy storage, etc. In the residential sector, due to heat power requirements and cost considerations, a hot water tank is the most common configuration. As the name indicates, a hot water tank utilizes hot water as the thermal energy storage medium and is encased in a tank that has excellent insulation, so that the heat loss is small.

The operating principle and the configuration of a heat pump system determines the main heat transfer processes in the system as well, which takes place in three subsystems. First of all, there is the heat exchange with the environment, the modeling of which is determined by the heat sources. For instance, an ASHP uses an outdoor evaporator heat exchanger coil with fans to extract heat from the ambient air forcedly, while a GSHP could use a Ground Heat Exchanger (GHE) in borehole to absorb heat from the ground source. Considering the use cases in the following sections, which involve GSHP systems, the focus of further discussions is then on a typical GSHP system. To build and validate a detailed theoretical thermal model of a borehole GHE for heat transfer requires too many geometric and thermal properties (Śliwa and Gonet 2005, Ruiz-Calvo et al. 2015), which are beyond the scope of the present work. Alternatively, if only the inlet temperature T^{in} and outlet temperature T^{out} of the borehole GHE are focused on, as shown in Fig. 3.3, based on the First Law of Thermodynamics, the specific heat capacity of brine c_b and the mass flow of brine $\dot{m}_b$ can be used to calculate the temperature change at different time steps as follows:

$$T_t^{out} = T_t^{in} + \frac{P_{Q,t}^{abs}}{c_b \cdot \dot{m}_b} \tag{3.4}$$

where $P_{Q,t}^{abs}$ is the absorbed thermal power at time t, which is also the difference between thermal power $P_{Q,t}^{hp}$ and electrical power of the GSHP $P_{el,t}^{hp}$ at the same time step and calculated in Equation (3.5).

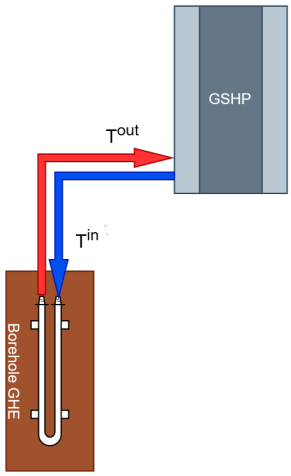


Figure 3.3: Illustration of a borehole GHE

$$P_{Q,t}^{abs} = P_{Q,t}^{hp} - P_{el,t}^{hp} \tag{3.5}$$

The second heat exchange process takes place in the heat pump itself. Similar to the modeling of the GHE, the thermal model of a heat pump could also have a varying degree of detail and complexity. As mentioned before, heat pumps operate upon a reverse Carnot cycle involving four steps. In a detailed, physical model, each step and the associated physical processes can be described. However, in the context of DSM, models are much more abstract. Therefore, the attention is focused on abstract models that are based on the Coefficient of Performance (COP), which is typically used to characterize heat pump's instantaneous overall performance at each time step. This value is influenced by many factors in actual operation such as the supply/return temperature, outdoor temperature, etc.. Equation (3.6) provides the definition of the COP, where the thermal power $P_{Q,t}^{hp}$ of the heat pump is measured and calculated based on the First Law of Thermodynamics in Equation (3.7), in which c_w, ρ_w and $\dot{V_w}$ are the specific heat capacity, density and volume rate of water respectively. The difference between supply temperature and return temperature is represented by $(T^{supply} - T^{return})$.

$$COP_t = \frac{P_{Q,t}^{hp}}{P_{el,t}^{hp}} \tag{3.6}$$

$$P_{Q,t}^{hp} = c_w \cdot \dot{V}_w \cdot \rho_w \cdot (T_t^{supply} - T_t^{return}) \tag{3.7}$$

The third heat exchange occurs in the hot water storage. The temperature and the corresponding thermal energy variations within the hot water storage significantly influence the performance of the entire system. Assuming that the density and specific heat capacity of hot water inside the storage are constant and the heat loss is negligible throughout the considered time frame, based on the heat capacity formula, the thermal energy change in the storage between two successive time steps is calculated as follows:

$$\Delta Q_s = c_w \cdot V_s \cdot \rho_w \cdot (T_t^{mean} - T_{t-1}^{mean}) \tag{3.8}$$

where V_s is the volume of the hot water tank and $(T_t^{mean} - T_{t-1}^{mean})$ denotes the average temperature change of the hot water, which are determined in Equation (3.9) for multiple layers with different temperatures within the hot water storage, assuming that the temperature is evenly distributed in each layer at every time step:

$$T_t^{mean} = \frac{1}{n} \sum_{i=1}^{n} T_t^i \tag{3.9}$$

By modeling the main heat transfer processes in these three subsystems, the simulation and validation of the heat pump system's performance is feasible and will be carried out in detail with two use cases in Section 4.1 and Section 4.2.

3.1.3 Electric Vehicle Supply Equipment

An Electric Vehicle Supply Equipment (EVSE), also known as charging station or chargepoint, is a power supply device that is capable of recharging plug-in Electric Vehicle (EV) with different power output in Alternating Current (AC) or Direct Current (DC). With the development of electric vehicles in the last years, the number of electric chargepoints is expanding rapidly, which therefore gains more attention and importance to act as DERs and has an increasing impact on grid integration (Das et al. 2020).

In terms of the modeling of charging behavior, simplified or linear models are utilized at an early stage for theoretical research purposes. For instance, the authors (Hõimoja et al. 2012) use a simplified linear Equation (3.10), that is based on the State of Charge (SoC), to calculate the charging time t_{ch}, where E_{EV} the rated battery capacity and P_{ch} the charging power are assumed to be constant.

$$t_{ch} = E_{EV} \cdot \frac{SoC_{start} - SoC_{stop}}{P_{ch}} \tag{3.10}$$

However, in a more practical context, due to the uncoordinated deployment of charging stations (Chang et al. 2021) and the system fluctuations regarding charging behaviors (Gong et al. 2020), the charging environment is dynamic (Alsabbagh et al. 2020). This makes the use of traditional modeling methods such as white-box models difficult for an accurate analysis and forecast (Chen et al. 2022) of charging behaviors, which are essential for operational decision making such as DSM. For instance, the results of short-term load forecasting can help utilities to optimize generation and to ensure grid stability in the short term. More details will be discussed together with the use case in Section 4.3.2.

3.2 Hypothesis

We propose the applicability of Gossen's First Law for the relationship between model utility and model complexity within the domain of DSM. The proposed hypothesis is stated as follows (Li et al. 2023):

Hypothesis. *In general, the complexity-utility relationship in the modeling of Distributed Energy Resources for Demand Side Management could be represented by a diminishing marginal utility curve.*

In order to validate this hypothesis, some preliminary questions, namely the quantification of the model complexity and the model utility, needs to be discussed and answered first in the following Section 3.3.

3.3 Quantification Methods

The proposed hypothesis assumes a diminishing marginal utility curve to represent the complexity-utility relationship in the modeling of DERs for DSM. As with Gossen's First Law, the marginal utility itself is an inherently abstract concept and needs to be quantified first in order to illustrate its relationship with consumption or other properties, such as income (Layard et al. 2008) in economics. Similarly, the method for quantifying the complexity and utility of DERs modeling in DSM is also crucial to visualize the interaction between them. This section discusses separately what kinds of quantitative options for complexity and utility are available and then explains the rationale for the choices made in this thesis.

3.3.1 Quantification of the Model Complexity

In computer science, complexity can be measured in various ways, such as required time, number of operations, required memory and big O notation. They do depend on the specific algorithms, their implementation, and the hardware they are running on. For instance, big O notation is often used to classify the efficiency or complexity of algorithms according to how their run time grows as the input size increases (Cormen et al. 2001). However, this famous approach primarily refers to algorithms or functions without the consideration of the involved models, which are of more importance in the context of DERs modeling in DSM. Therefore, in our scenario an appropriate approach should consider the models and the functions together. To determine the approaches, a short review of different methods for quantifying the model complexity is conducted first, followed by a discussion and analysis of different approaches. The review methodology is as same as described in Section 2.2 with a focus on the review works from different domains. Finally, the methods adopted in this thesis is selected based on this evaluation

For a better understanding of what is model complexity in environmental and climatological research, the authors (Malmborg et al. 2024) propose a more explicit definition of model complexity from an ecological perspective. This definition incorporates four facets, namely model class complexity, parameters complexity, input complexity and computational complexity. The model class complexity determines the structure of a model such as statistical or mechanistic models. The parameters complexity contains the number of parameters included in a model and the estimability of parameters. The input complexity touches on a diverse range of aspects such as challenges when collecting data, which is usually depending on the natural environmental process. The computational complexity relates to the several aspects of a model workflow such as software and model fitting degree.

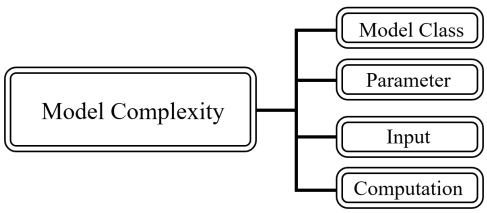


Figure 3.4: Definition of model complexity from an ecological perspective based on (Malmborg et al. 2024)

Fig. 3.4 shows the structure of this definition of model complexity. However, the authors propose this definition primarily as a conceptual paradigm, without providing concrete examples or detailed explanations.

In a review of urban building energy modeling (UBEM) approaches and methods, (Ali et al. 2021) discuss several modeling approaches and methods that could increase or decrease the model complexity. For instance, in data-driven models, which is a type of bottom-up modeling approach, feature extraction could reduce the model's complexity by eliminating unnecessary or irrelevant variables. Therefore, the number of extracted features can be utilized to compare the complexity of different data-driven models. Another type of bottom-up modeling approach, the reduced-order models have less complexity compared to the physics-based models due to fewer required inputs for the model formulation process. In this context, the number of required inputs can be used to compare the complexity of different reduced-order models. Besides, the reduced-order models provide more flexibility for optimization and reduce computational complexity.

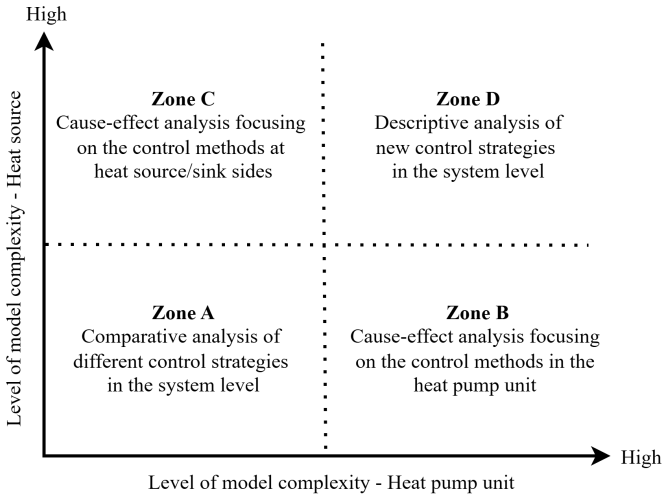


Figure 3.5: Scheme for different levels of model complexity depending on the type of study, based on (Clauß and Georges 2019)

In a more specific use case, (Clauß and Georges 2019) investigate the influence of model complexity in heat pump control on different key performance indicators for the energy efficiency, the DR potential and the heat pump operation. The authors adopt a scheme, outlined in (Madani et al. 2011), for selecting the necessary level of model complexity for a given type of analysis. Fig. 3.5 presents this adopted scheme and the authors investigate in Zone D when investigating models for the water storage tank, the heat pump control, the auxiliary heater control and the heat pump system control.

In the case study, the authors first investigate 6 control schemes with different model complexity. However, the model complexity is not directly quantified at the beginning, but instead inferred

through the number of HP cycles and average HP run time. Regarding the model complexity of auxiliary heater control, the authors also do not quantify the complexity, but rather approaches it by combining different components.

In some specific modeling environments or languages, such as Modelica, the configuration of a modeling component has a direct impact on the number of unknowns and equations of the system represented by the Differential-Algebraic System of Equations (DAE) (Qiu et al. 2024) and this results in a change in model complexity (Kovačević et al. 2025). In this context, the notion of "configuration" has two layers of meaning. The first refers to different modeling components that have similar functions in a system. For instance, Fig. 3.6 shows a complex stratified thermal storage model for domestic hot water (DHW), which could also be reconfigured by a simple non-stratified thermal storage model for the same purpose. The second meaning concerns the configuration menu of a modeling component, which could be adjusted based on the inputs. This approach offers the advantage of facilitating a clear and intuitive quantification and comparison of model complexity based on the number of unknowns and equations of DAE, similar to the third approach. However, this approach still has certain limitations. For example, modeling components may depend on specific libraries, making cross-library comparisons challenging, as components from different libraries are not necessarily compatible with each other. Additionally, while modifying certain modeling components, the extent to which the number of unknowns and equations of DAE is reduced is often difficult to predict in advance before the start of simulation. This could limit its feasibility and interpretability in some scenarios.

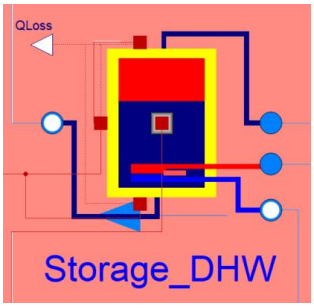


Figure 3.6: A stratified thermal storage model for domestic how water in Modelica (Kovačević et al. 2025)

Another approach to quantify the model complexity could involve assessing the number of parameters required in models, which is similar to one of the discussed facets in (Malmborg et al. 2024). In general, on a structural basis, any model is a combination of different inputs and outputs. In addition, for the same model, the number of parameters could be adjusted according to the study objectives or experimental conditions, so that models of different complexity can be built and identified accordingly. In particular, white-box and gray-box models are developed

based on physical laws or physical structures of a system (Li et al. 2023), which makes the parameters possess clear physical interpretability. Such interpretability can further facilitate the understanding and explanation of differences between models of varying complexity.

In the context of deep learning algorithms, model complexity is also a challenge. In a systematic overview of recent studies on model complexity in deep learning, (Hu et al. 2021) divide model complexity of deep learning into two categories: model expressive capacity and model effective complexity. The first category, model expressive capacity, captures the capacity of deep learning models in approximating complex problems, namely the upper bound. On the other hand, the model effective complexity reflects the practical complexity (Hanin and Rolnick 2019) of the functions represented by deep learning models with specific parameterization. It means that the effective model complexity could only be quantified for a model with fixed parameters. Based on this categorization, Equation (3.11) indicates the relationship between the expressive model complexity (denoted by $MEC(\mathcal{H})$ and the effective model complexity (denoted by EMC), where h and $\mathcal{H}$ represents a deep learning model and its corresponding hypothesis space respectively.

$$MEC(\mathcal{H}) = sup\{EMC(h) : h \in \mathcal{H}\} \tag{3.11}$$

In practical applications, more attention is given to EMC and its quantitative measures. In deep learning models, an effective complexity measure is expected to be sensitive to different parameter values used in models with the same structure (Hu et al. 2021). The authors point out that a well-known method is based on the piecewise linear property, which is suitable for deep learning models with piecewise linear activation functions. Other quantification metrics, such as the double descent phenomenon and Fisher-Rao norm, are primarily based on ideas and with respect to a certain category of deep learning models. Therefore, they do not receive detailed discussions by the authors. In general, detecting effective model complexity during training helps to investigate the usefulness of optimization algorithms, the role of regularizations, and generalization capability (Hu et al. 2021).

In addition to the methods mentioned above, when the scope of deep learning models is further extended to black-box models, there exists another approach specifically designed for them. In contrast to white-box models and gray-box models, black-box models do not focus on the physical properties or structures of a system and instead are data-driven by large empirical or measurement datasets (Li et al. 2023). Therefore, the physical details of a system, such as energy system, in black-box models are usually neglected and massive highly resolved datasets such as load profiles are necessary for good modeling performance. Typically, black-box models are integrated with advanced algorithms such as Machine Learning (ML) algorithms to form a comprehensive model system. Due to the unique structures and characteristics of different ML algorithms, such as

hyperparameter settings, input sizes, etc., which differ significantly from traditional physics-based models, quantification of their complexity also requires a different treatment depending on the specific algorithms utilized in the black-box models. In this thesis, an ML algorithm, namely the Random Forest (RF), is utilized in one use case. For this algorithm, its input size can be directly employed for quantifying model complexity for regression problems with black-box models. More details are further discussed in combination with concrete use cases in Section 4.2.1.3 and Section 4.3.

Table 3.1: Summary of the discussed approaches for quantification of the model complexity

Approach	**Use-case Scenario**	**Scope of Application**
Big O notation	Classification of algorithmic complexity	+++
Feature extraction	Data-driven models in UBEM	+
Reduced order	Gray-box models in UBEM	++
The number of unknowns and equations of DAE	Models in Modelica language	++
The number of parameters	White-box and gray-box models	+++
Piecewise linear property	Deep learning models with piecewise linear activation functions	++
Input sizes	Black-box models	+++

Table 3.1 provides a summary of several different approaches discussed above for quantification of the complexity of models and algorithms. Each approach is evaluated in terms of its use-case scenario and the corresponding scope of application, where more + means a wider scope of application. It's worth noting that not all of them are listed in Table 3.1 while some of the approaches are hard to be quantified, such as by using different control schemes. Among them, the approach using the number of parameters demonstrates the highest degree of applicability and feasibility for white-box and gray-box models based on the comparative analysis. The ecological perspective in (Malmborg et al. 2024) also indicates that parameter complexity is widely applicable. It is particularly advantageous in that it facilitates a deeper understanding of model behavior, as the associated parameters are explicitly defined and possess physically interpretable meaning. This interpretability allows for clear insight into the internal mechanisms of different models and the complexity they represent. It overcomes the limitation focusing solely on algorithms while neglecting the models themselves, as identical models may be utilized with different algorithms. Consequently, the number of required parameters, has been chosen to quantify the complexity for the first two use cases in this thesis, where models are constructed based on physical laws. In contrast, the last listed approach, which is specifically tailored for

black-box models, is selected for the third use case, which contains data-driven models, in this thesis. A detailed discussion and application of the selected approaches will be provided in the following Chapter 4.

3.3.2 Quantification of the Model Utility

One major goal of DSM applications is to reduce the cost of energy acquisition and the associated penalties by continuously monitoring energy use and managing appliance schedules (Bakare et al. 2023). Moreover, DSM not only aims to help customers reduce energy consumption or costs, but also seeks to assist grid operators in addressing generation-side imbalances due to variable generation. In doing so, DSM could contribute significantly to enhancing the resilience, operational efficiency, and flexibility of the overall energy system. Based on this definition, the methods for quantification of the model utility can be derived from the practical applications of DSM. In (Péan et al. 2019), the authors explain and summarize four typical ways for utilizing DSM, namely load-shifting, peak shaving, reduction of energy use and valley filling. Based on these, the amount of shifted energy can be used as a measure for quantification.

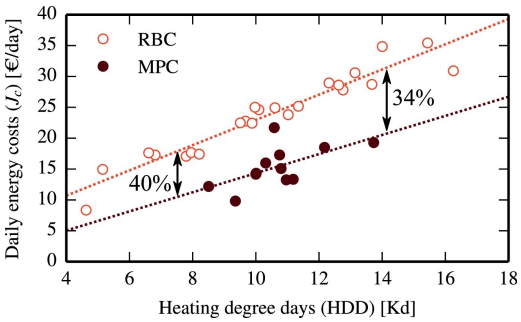


Figure 3.7: Quantified flexibility by using daily energy costs (De Coninck and Helsen 2016)

In (De Coninck and Helsen 2016), two more specific measures, namely daily primary energy use and daily energy costs are defined to show the improved and quantified utility in a comparative study when implementing MPC and rule-based control (RBC) in the same objective. Fig. 3.7 and Fig. 3.8 provide a visual representation of these two measures.

In addition, it is worth noting that the accuracy of a model must first be verified through offline simulations before the model is used to analyze utility and optimize flexibility with optimization approaches, such as MPC, in DSM applications. Models with high predictive and simulation accuracy can assist grid operators or DSM participants in optimizing resource allocation, such as regulating the required demand by matching the energy resources and energy availability without

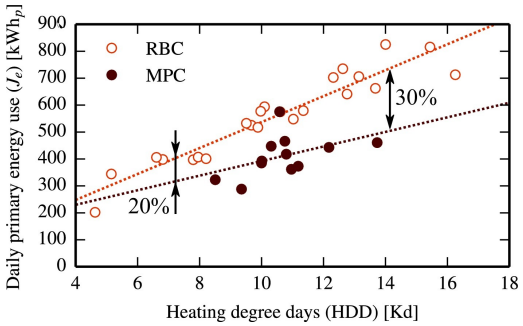


Figure 3.8: Quantified flexibility by using daily primary energy use (De Coninck and Helsen 2016)

adding new sources to the present system, reducing unnecessary energy waste and effectively lowering operational costs (Panda et al. 2022), thereby improving the overall efficiency and profitability of DSM applications. Furthermore, models with high predictive accuracy can further enhance the effectiveness of optimization approaches that rely on predictive steps such as MPC. According to ISO 5725-1, the general term "accuracy" describes the closeness of a measurement to the true value.

Based on this definition, the accuracy of a model with the help of some useful metrics in descriptive statistics, such as normalized Root Mean Square Error (nRMSE), normalized Mean Absolute Error (nMAE), Mean Absolute Percentage Error (MAPE) and Maximum Absolute Percentage Error (MaxAPE), can be quantitatively described. One major focus of this thesis is on the accuracy of different models in an offline simulation and uses quantified accuracy to represent utility of models. In order to reduce the impact of absolute values on the accuracy analysis, three descriptive statistics namely nRMSE, nMAE, MAPE and MaxAPE are defined in Equation (3.12), (3.13), (3.14) and (3.15), where $\hat{Y}$ is the generated or simulated value and Y_a is the ground truth.

$$\text{nRMSE}(\hat{Y}) = \frac{\text{RMSE}(\hat{Y})}{max(Y_a) - min(Y_a)} \tag{3.12}$$

$$\text{nMAE}(\hat{Y}) = \frac{\text{MAE}(\hat{Y})}{max(Y_a) - min(Y_a)} \tag{3.13}$$

$$\text{MAPE}(\hat{Y}) = \frac{1}{n} \sum_{1}^{i=1} \left| \frac{Y_{a,i} - \hat{Y}_i}{Y_{a,i}} \right| \tag{3.14}$$

$$\text{MaxAPE}(\hat{Y}) = max \left| \frac{Y_{a,i} - \hat{Y}_i}{Y_{a,i}} \right| \tag{3.15}$$

3.4 Discussion

In this chapter, the theoretical foundation for the hypothesis of diminishing marginal utility in DER modeling for DSM is first introduced. Then the focus lies on quantifying model complexity and utility in our scope, in order to answer the RQ1 and RQ2. As the first step, a brief survey of other review and application works related to complexity of models and algorithms across different domains is carried out. Table 3.1 presents the summary of different approaches for model complexity quantification. Considering the use-case scenario and scope of application, the number of parameters and input size are selected for white-box and black-box DER models respectively. For the quantification of model utility, it is first defined as the accuracy of a model, considering that the subsequent validation focuses on offline simulations. Afterwards, different useful metrics in descriptive statistics according to the definition of "accuracy" in ISO 5725-1 are introduced. The metrics constitute the foundation for quantifying model utility.

Given the results of this chapter, there are still limitations that can be addressed in future work. Firstly, the development of DER models in Modelica language is well established (Kovačević et al. 2025). Therefore, quantification method for model complexity using the number of unknowns and equations of DAE can be employed to investigate DER models in this language framework. Moreover, model utility is defined as accuracy in this thesis due to offline simulations. For online experiments and practical applications, other quantification approaches for quantifying model utility can provide important new insights, such as energy consumption, energy cost and CO_2 emission.

4 Gossen's First Law in the Modeling of Distributed Energy Resources for Demand Side Management

This chapter aims to address RQ3 systematically, namely "what is the relationship between the model complexity and the model utility of DERs models in the context of DSM?", by validating the proposed hypothesis in Section 3.2 through three different use cases. Key validation metrics are defined and analyzed to assess the reliability of the proposed hypothesis. The three use cases, including a one-day GSHP experiment at the Energy Lab at Karlsruhe Institute of Technology (KIT), a real-world GSHP dataset from a well constructed residential house at Switzerland and a massive real-world dataset of domestic EV chargepoints statistics in the UK, which are discussed in detail from Section 4.1 to Section 4.3. Finally, the results of this chapter are discussed and concluded in Section 4.4 and a brief outlook on a potential further development is given in the same section.

4.1 Heat Pump Case Study

Major Parts of this section are adapted or reproduced from the author's own previous works:

Chang Li, Kevin Förderer, Tobias Moser, Luigi Spatafora and Veit Hagenmeyer.
"Gossen's first law in the modeling for demand side management: a first heat pump case study".
In: *2023 Springer Energy Informatics Academy Conference (EI.A 2023)*, 2023, pp. 111–125. doi: 10.1007/978-3-031-48652-4_8.

In this section, the focus of the first case study lies on a GSHP system with a hot water storage (HWS) that is responsible for both space heating and domestic hot water supply. The system is part of the Living Lab Energy Campus (LLEC) from the Energy Lab at KIT. This section consists

of four parts, namely the fundamentals, including the experimental setting and the methodology, the utilized dataset, modeling process and the results as well as their discussion.

4.1.1 Fundamentals

This subsection introduces the experimental setting of the GSHP system first and then describes the methodology for modeling the heat and the energy transfer in each subsystem of a GSHP system, which constitutes the foundation for the subsequent model classification in Section 4.1.4.

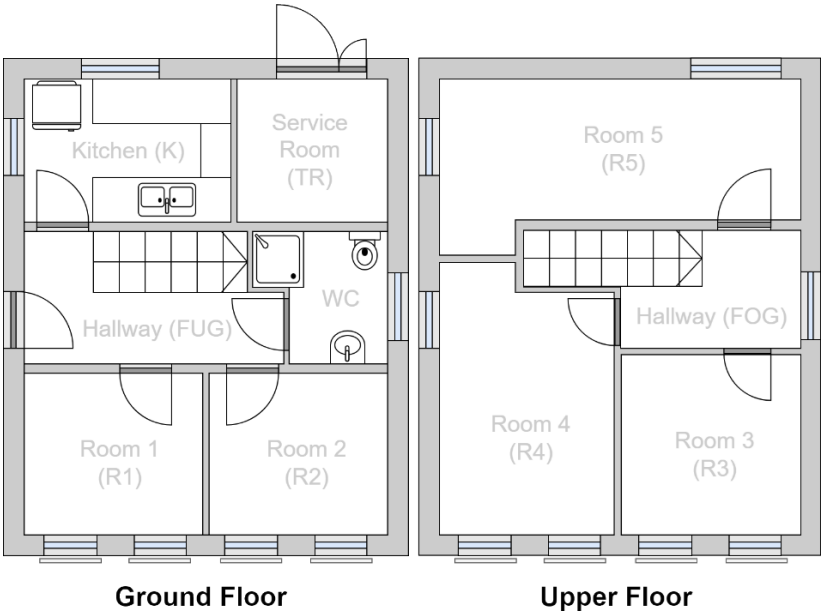
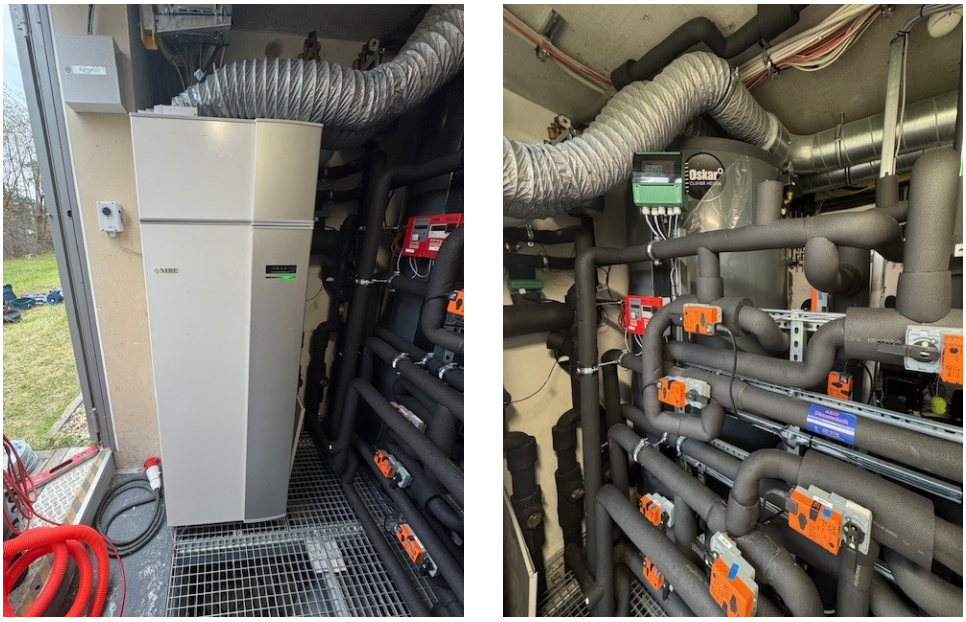


Figure 4.1: Schematic floor plan of the LLEC house (Li et al. 2023)

Analyzing the utility of different models relies on accurate real-world measurement data, which are provided by our measurement system infrastructure. Regarding the experimental setting, the floor plan and setup of our LLEC buildings (Hagenmeyer et al. 2016) is briefly introduced first to give an overview of their spatial layout and their heating matrix setup. Subsequently, the technical and software infrastructure that makes time series data from LLEC available is described.

Within the facilities of the Energy Lab (KIT 2014), three free-standing single-family houses have been purpose-built and equipped with a large variety of automation hardwares to allow for the use of advanced modern energy solutions (Hagenmeyer et al. 2016). All houses share the same two-story floor plan with several normal rooms, kitchen and restroom, as shown in Fig. 4.1. But

they differ in the choice of heating source. Each room is equipped with a range of different kinds of sensors and actuators, e.g., room climate sensors, heat flow meters, motorized windows, and smart thermostats, with only minor differences in equipment between the houses. The hardware is addressed by a programmable logic controller (PLC), which handles control requests from external systems and automatically collects, filters and stores measurement data in our database. The latter process will be explained in more detail in the following.



(a) The NIBE heat pump in the experimental LLEC house

(b) The Oskar hot water storage in the experimental LLEC house

Figure 4.2: The hardware of heat pump and hot water storage in the experimental house

This use case focuses on the LLEC house that uses a GSHP together with a HWS for the house heating and domestic hot water supply. Fig. 4.2 shows the HP and HWS in the experimental house and Fig. 4.3 presents the schematic heat matrix of the overall heating system along with the electrical circuit installed in the experimental LLEC house. The GSHP has its own internal sensors to determine the hot water flow rate, the supply and return temperature of water and brine as well as other data points. Furthermore, additional external sensors for water temperature and flow rate are installed to build a redundant measurement system to ensure more useful characteristics such as self-calibration, error compensation and the recovery of lost information.

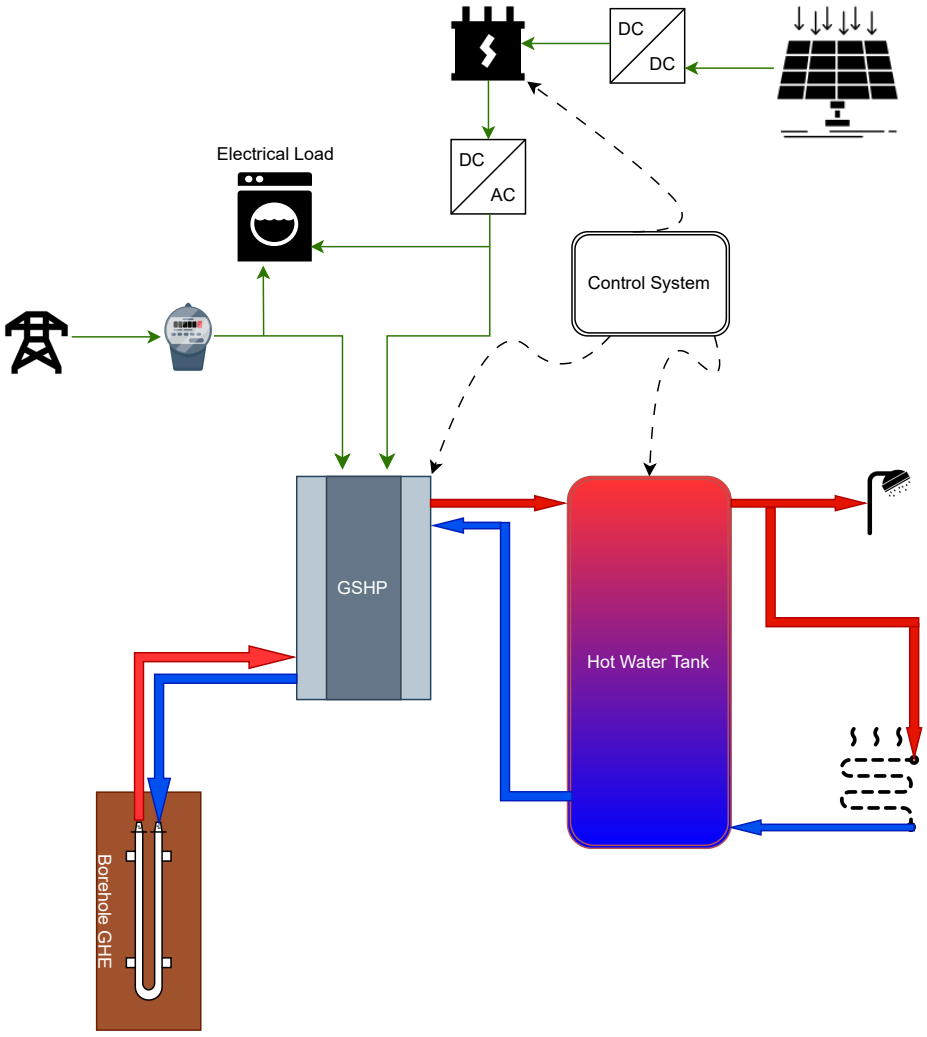


Figure 4.3: Schematic heat matrix and electrical circuit of the experimental LLEC house (Li et al. 2023)

The measurement system architecture for measurement acquisition, processing and storage is shown in Fig. 4.4 and consists of separate tools for collecting, processing, and storing timeseries data, with the relevant systems being shown in green. The data source in our experiments is the PLC, which acquires data in regular intervals via different bus systems or analogue inputs. The data is then formatted as JSON and sent to a Message Queuing Telemetry Transport (MQTT) Broker with a data point specific topic. The formative data is subsequently received by a data logger service subscribed to the MQTT topic, parsed, and pushed to our InfluxDB database via its Representational State Transfer (REST) interface. This new architecture allows to easily select the required inputs for different models via their respective topics and to choose the datasets for the validation and analysis of the simulation results.

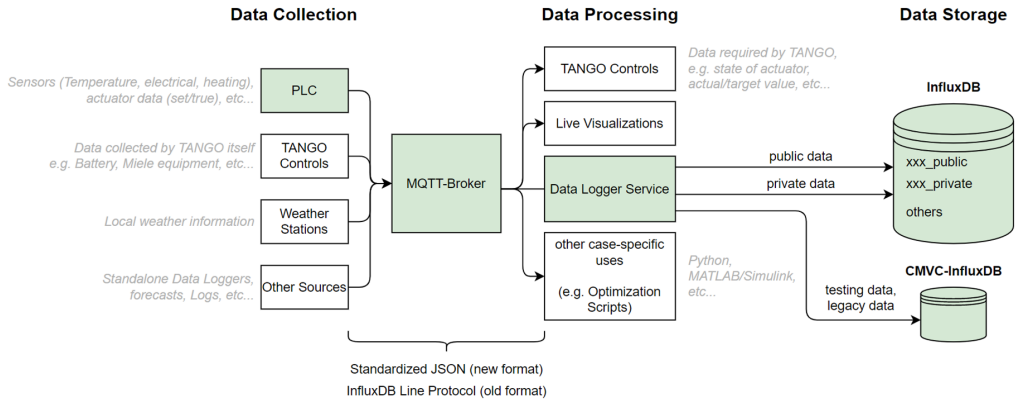


Figure 4.4: Technical infrastructure for data acquisition, processing and storage at LLEC (Li et al. 2023)

The operating principle of a GSHP for house space heating and hot water supply is based on a reverse Carnot thermodynamic cycle (Sarbu and Sebarchievici 2014), which means the thermal energy at the output consists of two parts, i.e., environmental energy and electrical driving energy. Depending on the source of different energy, models can vary accordingly. This use case focuses on a Ground-Coupled Heat Pump, a subset of GSHP, with vertical U-tubes as Ground Heat Exchanger (GHE). In this case, heat transfer takes place in three subsystems: GHE in borehole, heat pump and hot water storage, as described in Section 3.1.2. The structure of the system, as presented in Fig. 4.3, provides an overview of the subsystems that are to be modeled. The exact mathematical models with different complexity for each subsystem are developed and explained in the following.

4.1.2 Methodology

In order to validate the hypothesis, the modeling of each subsystem represents the initial step. For each subsystem, white-box models are developed using mathematical models according to its physical and functional characteristics. In this process, a key aspect lies in investigating whether models of different complexity can be applied to the same subsystem. These models are then parametrized and calibrated with experimental data obtained from LLEC. Based on the defined quantification methods in Chapter 3, different model classes with quantified complexity are then introduced. The model classes are built upon the models of the various subsystems and constitute the foundation for validation within offline simulations, where the quantified model utility, namely the accuracy, is systematically compared.

4.1.3 Dataset

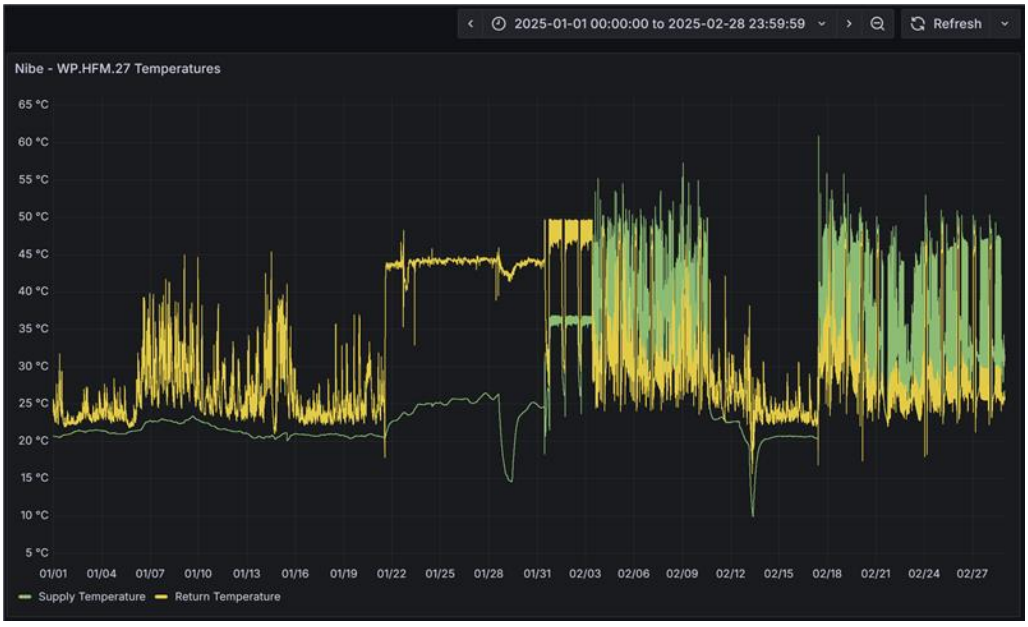


Figure 4.5: An illustrative example of the supply and return temperature of the experimental GSHP from January 2025 to February 2025 on Grafana

For the analysis and validation, a typical winter day in February 2023 is selected and the sensor data for 24 h is recorded continuously. The reason for initially selecting only a single day, a relatively short time frame, for experimentation and data collection lies in the fact that this

represents the first attempt to validate the proposed hypothesis. This limitation is addressed in the following Section 4.2.

The raw data is automatically collected and stored in a 1-min interval, which can be directly read and represented on the open and composable observability platform Grafana (Grafana 2014). Fig. 4.5 serves as an illustrative example to show raw data of the supply and return temperature of the experimental GSHP from January to February 2025.

The data are then collated to 60-min intervals by using the first value in each hour for the offline simulation of load profile. The initial value of the electrical power of the GSHP is set to the first value of the measurement.

4.1.4 Modeling

In this section, the employed model classification, i.e., model class is introduced and different options for modeling a GSHP and HWS with varying degree of complexity, which are used for the evaluation, are discussed. Four different model classes (A, B, C, and D) with decreasing complexity in terms of the number of parameters required are introduced. An overview of the classification and the number of required parameters is first presented in Table 4.1.

Table 4.1: Model classification with respect to parameters

Model Class	**Combination**	**Number of required Parameters**
Model A	(4.1)(4.2)(4.7)(4.9)(4.10)	11
Model B	(4.7)(4.8)	6
Model C	(4.5)(4.8)	5
Model D	(4.3)(4.6)	3

For the heat exchange occurring in the U-tubes GHE, Equation (4.1) and (4.2) model the heat exchange process based on the physical laws mentioned in Section 3.1.2.

$$T_t^{out} = T_t^{in} + \frac{P_{Q,t}^{abs}}{c_b \cdot \dot{m}_b} \tag{4.1}$$

$$P_{Q,t}^{abs} = P_{Q,t}^{hp} - P_{el,t}^{hp} \tag{4.2}$$

In terms of the thermal model of the heat pump itself, models that are based on COP are considered, as explained in Section 3.1.2. In (Peralta et al. 2021), a linear data-fitting equation is utilized for calculating the COP of the GSHP, taking the outlet temperature of the borehole GHE and indoor temperature into consideration. Similarly, in this use case, a second-order polynomial regression in consideration of outdoor air temperature T_t^{env} is employed, to calculate the declared COP and declared heat power P_{dh} in kW at medium partial load based on the data sheet of the heat pump manufacturer (NIBE 2023) in Equation (4.3) and (4.4). The R-squared value of the regression is 0.9902 and 0.9972 respectively, which represent a very good fit of the model to the data.

$$COP_t^{cal} = -0.0007 \cdot (T_t^{env})^2 + 0.0983 \cdot (T_t^{env}) + 3.8429 \tag{4.3}$$

$$P_{dh} = 0.0009 \cdot (T_t^{env})^2 - 0.1992 \cdot (T_t^{env}) + 3.4164 \tag{4.4}$$

Furthermore, Equation (4.3) and (4.4) can be simplified by neglecting the effect of outdoor temperature and consider $COP^{const.}$ as well as $P_{dh}^{const.}$ as a constant as follows:

$$COP^{const.} = 3.8429 \tag{4.5}$$

$$P_{dh}^{const.} = 3.4164 \tag{4.6}$$

Alternatively, the COP^{avg} can be calculated directly with the measured thermal and electrical data over a period of time and obtain an average value as follows:

$$COP^{avg} = \frac{1}{n} \sum_{1}^{n} \frac{P_{Q,t}^{hp}}{P_{el,t}^{hp}} \tag{4.7}$$

The thermal power $P_{Q,t}^{hp}$ is measured and calculated with Equation (4.8), where c_w, ρ_w and $\dot{V}_w$ are the specific heat capacity, density and volume rate of water respectively. The difference between supply temperature and return temperature is represented by $(T^{supply} - T^{return})$.

$$P_{Q,t}^{hp} = c_w \cdot \dot{V}_w \cdot \rho_w \cdot (T_t^{supply} - T_t^{return}) \tag{4.8}$$

As the central storage for thermal energy, the temperature and corresponding energy changes in the HWS have a significant impact on the overall system. Regarding the thermal model of a HWS,

the structure of a HWS determines its temperature calculation. In this use case, a multi-layer hot water tank with negligible heat loss as the central storage for domestic hot water and space heating is implemented. Assuming the density and the specific heat capacity of hot water as constant, the thermal energy change in the storage between two successive time steps is calculated as follows:

$$\Delta Q_s = c_w \cdot V_s \cdot \rho_w \cdot (T_t^{mean} - T_{t-1}^{mean}) \tag{4.9}$$

where V_s is the volume of the hot water tank and $(T_t^{mean} - T_{t-1}^{mean})$ denotes the average temperature change of hot water, which are determined in Equation (4.10) with the assumption that the temperature is evenly distributed in each layer at every time step:

$$T_t^{mean} = \frac{T_t^u + T_t^m + T_t^l}{3} \tag{4.10}$$

In Equation (4.10), three temperature sensors placed in the upper, middle and lower layer of the HWS are used to measure the temperature of each layer, assuming that each of these temperatures represents one-third of the total capacity.

Based on the aforementioned models, a more detailed interpretation of the model classifications presented in Table 4.1 can now be provided. As shown in Table 4.1, Model A utilizes Equation (4.1) and (4.2) to calculate the absorbed thermal power in the brine directly. The result is then combined with Equation (4.7) to calculate the electrical power of the heat pump. In addition, the energy change in the hot water storage is also taken into account by calculating Equation (4.9) and (4.10). Model B combines Equation (4.7) and (4.8) to compute the consumed electrical power of the heat pump without consideration of the energy change in the storage. Furthermore, Model C uses a constant $COP^{const.}$ to estimate the $P_{el,t}^{hp}$. Finally, a constant declared heat power is obtained in the simplest Model D. An overview of the individual parameters that apply to the four model classes is given in Table 4.2.

Additionally, Table 4.3 lists the fixed thermal and other parameters, where the specific heat capacity of brine is taken from the technical diagram provided by the heat pump manufacturer (NIBE 2023) and based on a 20 vol% mixture of Tyfocor.

4.1.5 Results and Discussions

In this section, based on the model classification in Section 4.1.4, various models are used to perform offline simulations to calculate the load profile, namely the electrical power of the GSHP system, and to analyze the results against the measurement.

Table 4.2: Overview of the applied parameters to each model class

Parameter	Model A	Model B	Model C	Model D
T^{supply}		■	■	
T^{return}		■	■	
$\dot{V}_w$		■	■	
COP^{avg}	■	■		
$\dot{m}_b$	■			
T^{out}	■			
T^{in}	■			
T^{u}	■			
T^{m}	■			
T^{l}	■			
ρ_w	■	■	■	
c_w	■	■	■	
c_b	■			
V_s	■			
T^{env}				■
$P_{dh}^{const.}$				■
COP^{cal}				■

Table 4.3: Fixed thermal and volume parameters

Parameter	Value
Hot water tank volume V_s	920 L
Specific heat capacity of water c_w	4186 J/(kg·°C)
Density of water ρ_w	0.988 kg/L
Specific heat capacity of brine c_b	3940 J/(kg·°C)

For the simulation and analysis, the initial value of the electrical power of the GSHP is set to the first value of the measurement. Fig. 4.6 shows the simulation results of the different model classes and their deviation from the measurement. The diagram shows that the results of Model A, which is the most complex one in this use case, are almost identical to the measured results since it considers not only the energy transfer in the heat pump itself in detail but also the thermal energy change in the HWS. In contrast, the less complex Models B and C show large deviations at some points in time, such as $t = 10$ and $t = 15$. This behavior could be caused by ignoring

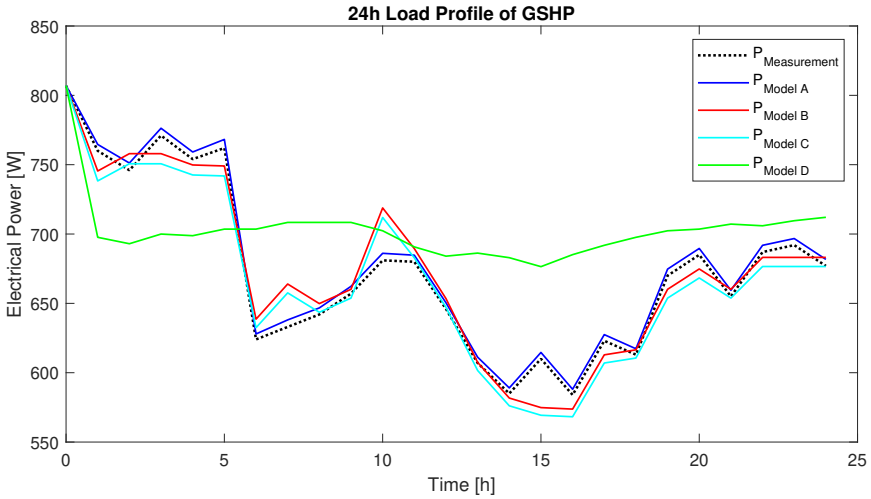


Figure 4.6: Comparison between model simulation results and measured values (Li et al. 2023)

the energy change in the HWS in both models. Besides, it's worth noting that the deviations of Model B are not always smaller than the deviations of Model C at all time steps, although Model B is slightly more complex than Model C in terms of the number of required parameters. Finally, Model D consistently exhibits a relatively large deviation from the measured values, as it relies solely on declared constant $P_{dh}^{const.}$ and does not account for more complex dynamic energy transformations.

In order to describe the overall statistic features of the simulation results, the MAPE and the MaxAPE are calculated, yielding the results presented in Table 4.4. Model A, with the most parameters, has the lowest MAPE of 0.68% as compared to the other three simplified models. Furthermore, the MaxAPE by using Model A is also the minimum in all models. Between Model B and Model C with a slight difference in the model complexity, it is worth noting that although the MaxAPE of Model B is larger than that of Model C, the MAPE of Model B is still smaller than the MAPE of Model C. As for Model D, the largest simplification of it leads to the largest error among the other models.

As explained in Section 3.3.2, quantification of the model utility by utilizing accuracy is dependent on the particular application being considered. In order to quantify the utility of these models, the following formula Equation (4.11) is defined to calculate and describe the accuracy, i.e., utility of the model, where U represents the utility of a model in natural number.

$$U = (1 - (\mathrm{MAPE})) \cdot 100 \tag{4.11}$$

Table 4.4: Comparison of MAPE and MaxAPE

	MAPE	**MaxAPE**
Model A	0.68%	0.82%
Model B	1.65%	5.55%
Model C	1.80%	4.54%
Model D	8.07%	17.32%

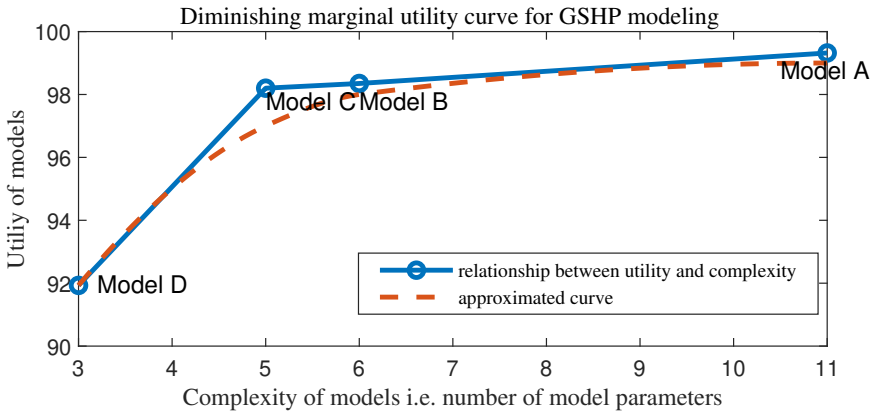


Figure 4.7: Diminishing marginal utility curve based on model complexity (based on (Li et al. 2023))

With this definition, the relationship between model complexity and model utility of four different GSHP model classes is illustrated in Fig. 4.7. This demonstrates that the results are basically in line with our hypothesis of diminishing marginal utility. However, regarding the limitations of data points, the graph line is not as smooth as the approximated diminishing marginal utility curve determined via regression methods which is also presented in orange dashed line in Fig. 4.7. On one hand, this difference between the simulation results and the approximated curve is evident in the initial stage of the curve. On the other hand, based on the substantial improvement in terms of the utility from Model D to Model C, it indicates that a highly satisfactory outcome can be achieved without a substantial increase in the model complexity in this use case. However, this first study is only limited in scope and therefore provides only limited evidence. In the following sections, additional studies are conducted to add further evidence to our claim.

4.2 A Thorough Heat Pump Case Study

Major Parts of this section are adapted or reproduced from the author's own previous works:

Chang Li, Gina Brecher, Jovana Kovačević, Hüseyin Çakmak, Kevin Förderer, Jörg Matthes and Veit Hagenmeyer.
"Gossen's first law in the modeling for demand side management: a thorough heat pump case study with deep learning based partial time series data generation".
In: *Springer Energy Informatics*, 2024, 4:47. doi: 10.1186/s42162-024-00353-z.

The use case in Section 4.1 is a first attempt to validate the proposed hypothesis. However, there are two major limitations in this first use case. The first is that only one day, i.e., 24 h in February has been selected for validation, which could limit the robustness and generalizability of the proposed hypothesis, since different days might have different patterns. Secondly, potential applications of the findings, especially in real-world scenarios rather than experiments, should be discussed and summarized in more detail (Li et al. 2024a). Therefore, this second thorough case study is focused on a GSHP system with two HWSs and their operation over a longer time period. In this process, a large real-world dataset with the predicted flow rate data is utilized as the input. The system is installed in a stand-alone house in real-world scenarios in Switzerland, which is equipped with corresponding automatic data measurement and data storage equipments (Meyer 2024). To handle the problem of the absence of time series data in the dataset, three different machine learning algorithms together with our modified persistence model, which serves as the baseline, are first utilized and compared. With the generated flow rate, the heat pump system modeling, model classification based on the complexity namely the number of required parameters and load profile simulation for a time horizon of 7 days with different patterns are then carried out. This section consists of five parts, namely the fundamentals, including the methodology and the experimental setting, the utilized dataset, the data generation results, the modeling process and the results as well as discussions.

4.2.1 Fundamentals

Similar to Section 4.1.1, in this section the fundamentals for modeling the heat and the energy transfer in each subsystem of a GSHP system are also introduced, which constitutes the foundation for the subsequent model classification in Section 4.2.3. In the second part, the related work and

the algorithms for time series data generation are introduced, which are essential for ensuring the completeness of the dataset that is utilized in this use case for the subsequent simulations.

4.2.1.1 Heat and Energy Transfer

In Section 4.1.4, the modeling of the ground source heat pump is carried out based on three main subsystems for heat transfer, namely the thermal model of the borehole GHE, the thermal model of the heat pump itself and the thermal model of the HWS. However, due to the new structure of the selected system in the current use case, it's necessary to modify the models. In addition, it's worth noting that in order to improve readability of equation numbering in each use case, the tags of some equations below are renumbered, although they remain unchanged from the use case in the last section.

The selected system uses a GSHP together with a smaller hot water tank for the domestic hot water supply and a larger hot water tank for the house heating. The schematic heat matrix of the overall heating system as well as different positions of the installed temperature sensors is shown in Fig. 4.8.

The heat transfer in the HP itself, namely the thermal power P_Q is measured and calculated with Equation (4.12) without change, where c_w, ρ_w and $\dot{V}_w$ are the specific heat capacity, density and volume rate of water respectively. The difference between supply temperature and return temperature is represented by $(T^{supply} - T^{return})$.

$$P_Q = c_w \cdot \dot{V}_w \cdot \rho_w \cdot (T^{supply} - T^{return}) \tag{4.12}$$

The heat transfer in the borehole GHE is unchanged modeled in Equation (4.13) and (4.14), where T^{in} and T^{out} are the inlet and outlet temperature of the borehole GHE as shown in Fig. 3.3. c_b is the specific heat capacity of the brine and $\dot{m}_b$ is the mass flow of the brine. Besides, P_Q^{abs} is the absorbed thermal power, which is also the difference between P_Q and P.

$$T^{out} = T^{in} + \frac{P_Q^{abs}}{c_b \cdot \dot{m}_b} \tag{4.13}$$

$$P_Q^{abs} = P_Q - P \tag{4.14}$$

To model the performance of the heat pump itself, one simple way is to calculate the COP directly with the measured thermal and electrical power over a period of time to obtain an average value

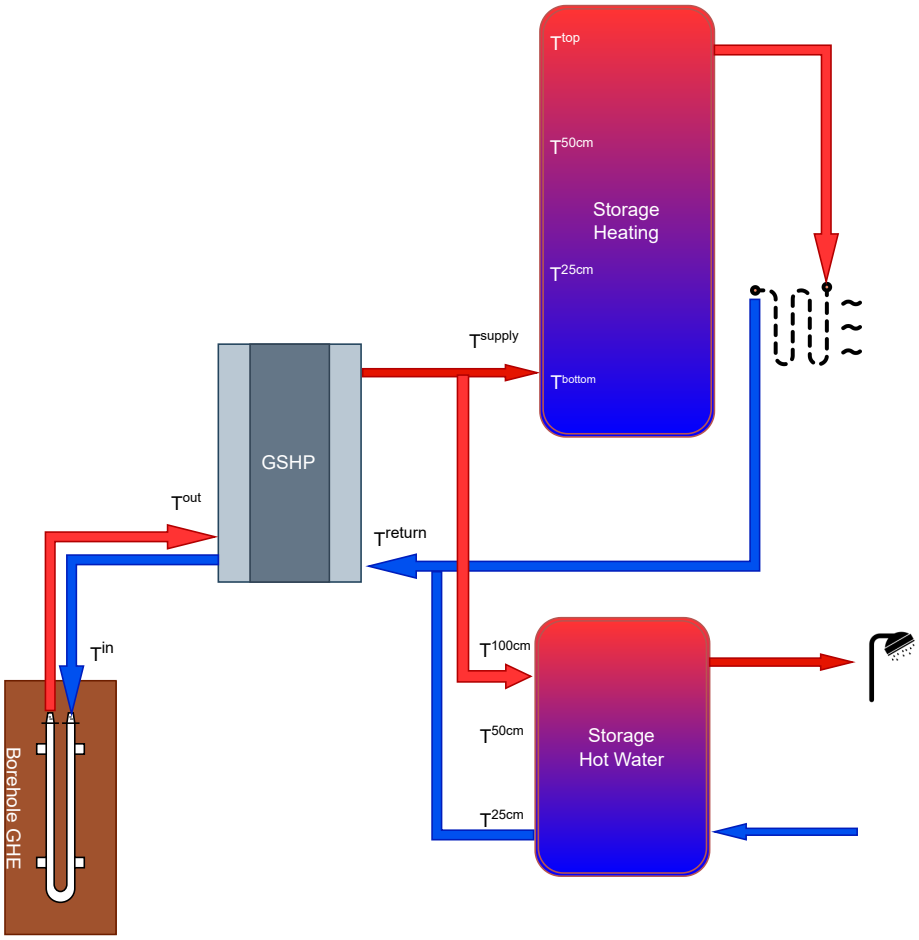


Figure 4.8: Schematic heat matrix consisting of positions of installed temperature sensors

as presented in Equation (4.15). Moreover, the thermal power can be obtained as mentioned in Equation (4.12).

$$\mathrm{COP}^{avg} = \frac{1}{n}\sum_{t=1}^{n}\frac{P_{Q,t}}{P_t} \tag{4.15}$$

In this use case, the system contains two different hot water tanks for different purposes as shown in Fig. 4.8. As the central storage for thermal energy, the temperature and corresponding energy

changes have a significant impact on the overall system. Therefore, it's necessary to consider the energy changes of the storage separately. Moreover, it shows that 4 temperature sensors are installed at different layers in the large heating storage tank and 3 temperature sensors are placed for the smaller one with equal distance from each other. This layout leads to the modification of the thermal model of the HWS. In general, the thermal energy change in the storage between two successive time steps can be calculated with Equation (4.16) under the assumption that the density and the specific heat capacity of hot water is constant. In Equation (4.16), the V_s is the volume of the hot water tank and $(T_t^{mean} - T_{t-1}^{mean})$ denotes the average temperature change of the hot water, which are determined in Equation (4.17) and (4.18) for the small and the large storage respectively with the assumption that the temperature is evenly distributed in each layer at every time step.

$$\Delta Q_s = c_w \cdot V_s \cdot \rho_w \cdot (T_t^{mean} - T_{t-1}^{mean}) \tag{4.16}$$

$$T_t^{mean,s} = \frac{T_t^{25cm} + T_t^{50cm} + T_t^{100cm}}{3} \tag{4.17}$$

$$T_t^{mean,l} = \frac{T_t^{bottom} + T_t^{25cm} + T_t^{50cm} + T_t^{top}}{4} \tag{4.18}$$

4.2.1.2 Related Work: Time Series Data Generation

Table 4.5: Excerpt from the raw data

Date and Time	T^{supply} [°C]	T^{return} [°C]	P_Q [W]	P [W]	COP
2021-01-02 03:00:00	42.4	34.6	13898	3440	4.04
2021-01-02 03:15:00	51.9	44.3	13494	4040	3.34
2021-01-02 03:30:00	58.3	51.1	12775	4612	2.77
...	...	...	...	...	...

The data in the dataset is provided with an interval of 30 seconds to 60 seconds (Meyer 2024). In this use case, the historical raw data with a time interval of 15 minutes are extracted for the years 2021 and 2022. Table 4.5 shows an excerpt from the extracted raw data, where T^{supply}, T^{return}, P_Q and P have the same meaning as explained in Equation (4.12). However, one key variable is missing in the raw data, which is the flow rate, i.e., $\dot{V}_w$ in Equation (4.12). This variable is used

for calculating the thermal power and thus needs to be generated first for the following comparison and simulation.

Dealing with partially missing data in modeling when utilizing large datasets for validation, has been an important topic, not only in engineering but also in other fields such as medicine for a long time. In order to address this problem more accurately and reliably, different approaches, from common statistical techniques to machine learning based methods in recent years, are explored based on different use cases in many publications. In (Zhang 2016), single imputation of missing data such as mean, median and mode imputations is conducted. However, no quantified results are summarized in the article. The authors in (Austin et al. 2021) have developed a model based on multiple imputation to create imputed data and proven that the created values by using multiple imputation are plausible for their use case. Another new technique, which is a hybrid approach of single and multiple imputation techniques, is proposed in (Khan and Hoque 2020) in two variations to impute categorical and numeric data. The experimental results show that the proposed algorithm achieves a higher F-measure around 20%, which is a measure of predictive performance, for binary data imputation and around 11% improvement in terms of error reduction for numeric data. To handle the nonlinear associations between the variables in multilevel models, a flexible sequential approach based on Bayesian estimation techniques is proposed in (Grund et al. 2021), which outperforms the conventional MI methods for multilevel models with nonlinear effects. In (Weber et al. 2021), the authors have introduced a new Copy-Paste Imputation method for imputing energy and power time series. The method takes into account the total energy of each gap and outperforms the selected three benchmark imputation methods in their work.

In addition to using statistical methods to reconstruct missing data, machine learning imputation methods are also widely used for imputation of missing data. For instance, the authors in (Jerez et al. 2010) compare the performance of machine learning based techniques such as multi-layer perceptron (MLP) and k-nearest neighbor (KNN) with statistical techniques such as MI. The results reveal that the machine learning techniques lead to a significant enhancement of accuracy compared to statistical procedures. Similarly, eight statistical and machine learning imputation methods are compared based on real data and predictive models in (Li et al. 2024c). The most effective results are attained by KNN and Random Forest (RF). In the survey paper (Emmanuel et al. 2021), the authors aggregate different imputation methods, particularly focusing on machine learning techniques. They evaluate the performance of KNN and missForest, which is an iterative

method based on RF, by utilizing a power plant fan dataset. Besides the common machine learning techniques, deep learning methods are also explored for dealing with missing data such as Long Short-Term Memory (LSTM). In (Tian et al. 2018), a new model named as LSTM-M is proposed for managing missing data in the traffic flow, which outperforms several other methods such as Support Vector Regression (SVR) in terms of accuracy. Likewise, the authors in (Ma et al. 2020a) propose a LSTM-BIT model, which is a hybrid LSTM model with Bi-directional Imputation and Transfer Learning (BIT). The results show that the proposed model achieves a 4.24% to 47.15% RMSE under different missing rates.

Moreover, since Transformer was proposed in 2017 (Vaswani et al. 2017), the exploration of applications based on its architecture is still ongoing. The huge success of this architecture in natural language processing (NLP) and computer vision (CV) motivates the exploration of its further potential such as handling time series data (Hertel et al. 2023). However, there have been very few works that focus on utilizing Transformer for handling data generation.

4.2.1.3 Selected Approaches

Based on the related works presented above, three different approaches are selected for data generation in this section, namely RF, LSTM and Transformer. Furthermore, a modified persistence model as the baseline is proposed for a better quantitative comparison and discussion. These algorithms for forecasting the flow rate in the heat pump modeling are first briefly described, together with the definition of our modified persistence model as the baseline.

As an ensemble learning method for classification and regression problems (Breiman 2001), RF has been widely used in many classification and regression problems. When dealing with data generation, it also shows promising results as stated in (Emmanuel et al. 2021). When the data is presented through time series, it requires transforming the time series dataset into a supervised learning problem first. Fig. 4.9 shows this transformation process, a sliding window with an input size of one as an example, where Y is the value at each time step.

In detail, the value Y at the next time step is used as the ground truth for prediction of the current time step. For instance, Y2 and Y3 are used as the prediction value for the first and the second

time step and so on. In this way, the time series dataset is transformed from a single array into two separate arrays, thereby constituting a supervised learning problem.

However, there is a limitation of this method that cannot be ignored, i.e., random forest cannot extrapolate. It means that predicted values are always within the range of the training set. In this use case, different input sizes are tested to find an ideal parameter. Finally, a bagged regression ensemble object with an input size of 5 together with the temporal features such as Monday, Tuesday etc. as the 6th input is created, to use the bootstrap aggregation method for model training, since significant improvements with further increased input sizes can not be achieved within our simulation setting.

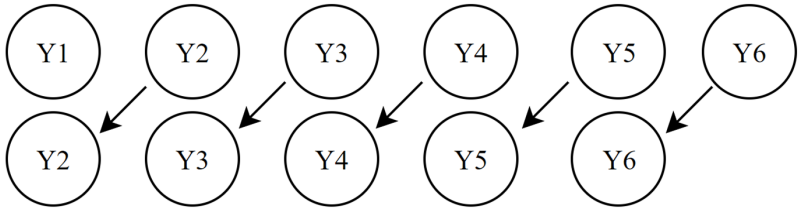


Figure 4.9: Transformation of time series into a supervised learning problem with input size of one

For predicting data based on time series while avoiding the vanishing gradient problem, LSTM has been developed as a modified version of traditional RNN. By introducing the so-called gates, LSTM can regulate the flow of information and maintain valuable information. In comparison to other RNN, LSTM can deal with large amounts of data and time steps easily (Zhu et al. 2019). Besides, it's also powerful when managing missing data as presented in (Tian et al. 2018, Ma et al. 2020a). Based on these advantages, it's been chosen as one of the algorithms in the use case for predicting the flow rate.

For all RNNs, one major limitation is that the computations must be performed in the sequence's order, which makes parallel computation difficult and thus limits the efficiency when dealing with long sequences. The proposed Transformer architecture in (Vaswani et al. 2017), which relies on the self-attention and multi-head attention mechanism, solved this limitation, making it more efficient than RNNs. While there is still debate about the advantages of Transformer in time series as remarked in (Wen et al. 2022), the consideration and introduction of this new architecture to deal with time series data generation is worthwhile.

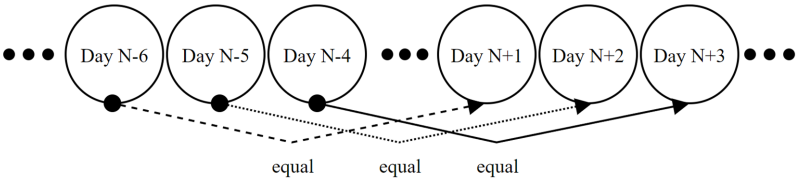


Figure 4.10: Modified persistence model

The persistence model (Notton and Voyant 2018) is often used as a trivial reference model when different forecast models are compared. In this work, a modified version of the persistence model is defined by considering the temporal impacts. Instead of generating the future value by assuming that no changes happen between the current time step and next time step, the values from the previous week are used, as presented in Fig. 4.10.

4.2.2 Dataset and Generation Results

According to the date and time, the raw data are pre-processed into time series data by hour at first. Besides, it's assumed that the thermal power and the electrical power are constant throughout each time interval. Moreover, it's worth noting that the thermal power equals zero when the heat pump is turned off, which means that the frequency of zeros in the pre-processed data should be as small as possible to avoid the case of sparse data.

It should be noted that there is no explicit definition specifying the frequency of zeros required in a dataset for it to be classified as sparse data. However, machine learning often performs poorly in prognosis, especially when dealing with sparse data (Peng et al. 2021). Therefore, it's preferable to choose datasets with fewer zeros.

Based on these three conditions mentioned above, the data from January 4th to February 7th in 2021 and from January 31st to March 6th in 2022 are selected for the calculation of the average flow rate by hour. Each time period starts on Monday and ends on Sunday. The reason for choosing another month in 2022 is that several days of data are completely missing in January. Fig. 4.11 shows the results of calculated flow rate of the selected 5 weeks in 2021 and 2022. The frequency of zeros of the selected time period in 2021 and 2022 are 23.57% and 32.38% respectively. It shows that the data in 2021 are denser than

the data in 2022. Therefore, the chosen time period in 2021 will be determined for the following analysis and validation.

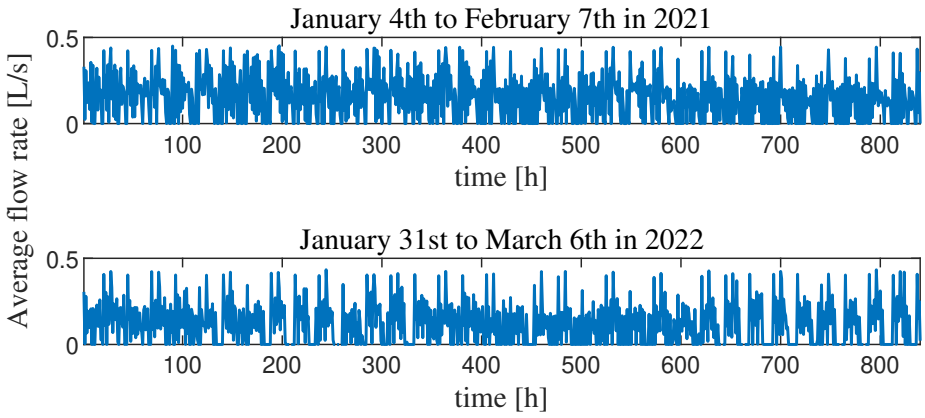


Figure 4.11: Average flow rate during the selected time period in 2021 and 2022

Based on the discussed approaches in Section 4.2.1.2, the flow rate generation results are given and compared in the following steps. As mentioned above, the selected time period in 2021 contains 5 weeks. The flow rate in the first 4 weeks are calculated and used as training set with cross-validation. The subsequent week of 7 days serves as the ground truth for validation of the generated flow rate. In contrast to predicting multiple subsequent time steps in a closed loop forecasting, an open loop forecasting for generating the data at the next time step is utilized in this case. This means that for subsequent time steps, the ground truth, which is the calculated flow rate in this case, is collected and used as input till the last time step.

The conventional approach is to create forecast models for each measured variable in Equation (4.12), namely the thermal power, the supply and return temperatures and then use the predicted values of these variables for calculating the flow rate. Compared to this conventional approach, the proposed pre-processing approach is more straightforward and less complex. The proposed approach calculates the flow rate in the past explicitly and only needs to create a forecast model for the flow rate directly.

To optimize the forecast results of each method, the hyperparameters are tuned in different approaches separately, where the hyperparameters for RF are automatically optimized in MATLAB and the tuned hyperparameter settings for LSTM and Transformer in PyTorch are shown in Table 4.6. It's worth noting that hyperparameters such as epoch and number of layers in LSTM and Transformer, which have a significant impact on the complexity and the run time of both

approaches, are set to be the same in order to ensure that the complexity of both methods does not differ too much within the range of tuned values.

Table 4.6: Hyperparameter setting for LSTM and Transformer

LSTM		Transformer	
Input Size	1	Num Heads	16
Hidden Size	256	Hidden Dimension	512
Hidden Layer	1	Num Layers	1
Batch Size	128	Batch Size	4
Epoch	200	Epoch	200
Learning Rate	0.01	Learning Rate	0.001
Optimizer	Adam	Regularization	Dropout = 0.1

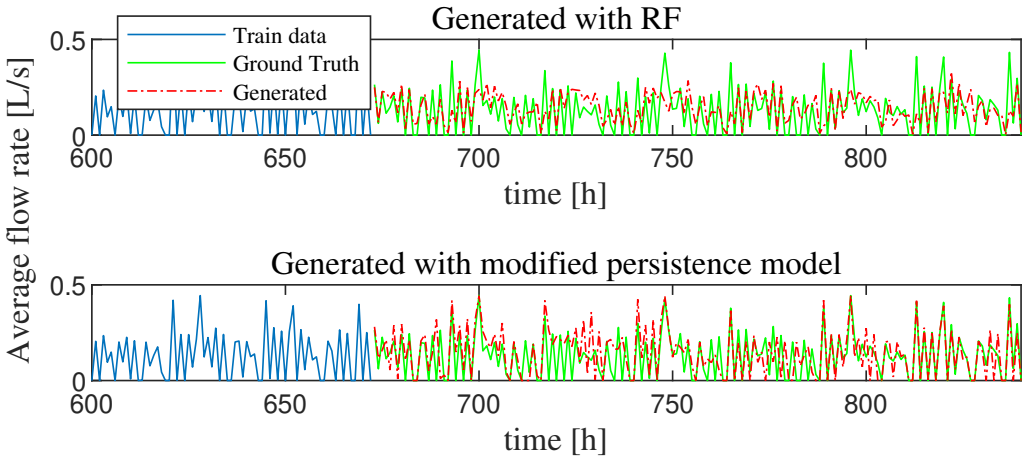


Figure 4.12: Generated average flow rate with RF and the modified persistence model (Li et al. 2024a)

Two descriptive statistics, as defined in Section 3.3.2, are summarized in Table 4.7. The detailed plots are presented in Fig. 4.12 and Fig. 4.13. It should be noted that only an excerpt of the training data is plotted in order to better visualize the comparison between the ground truth and the generated data.

According to the results in Table 4.7, the minimum error of the generated data is given by LSTM with a nRMSE of 10.56% and a nMAE of 7.47%. On the other hand, the results of RF are no better than the baseline with the modified persistence model. This demonstrates the limitation of

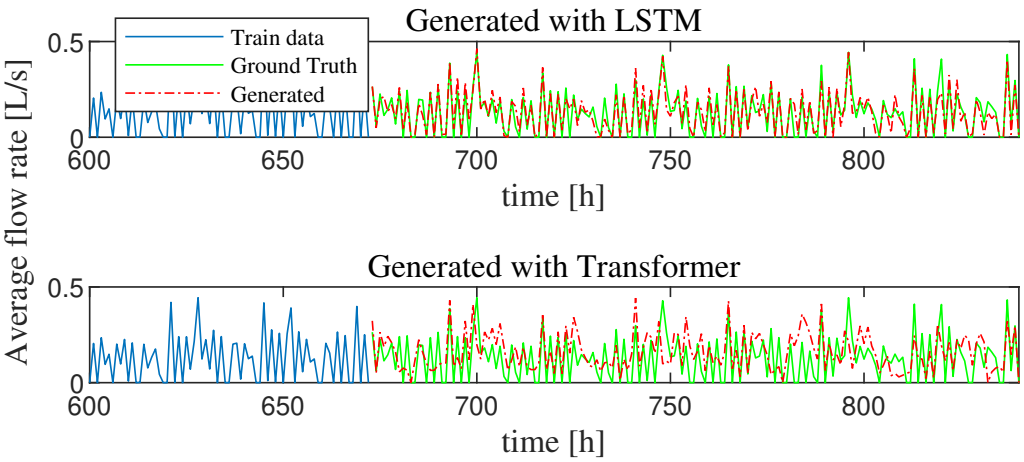


Figure 4.13: Generated average flow rate with LSTM and Transformer (Li et al. 2024a)

Table 4.7: Summary of descriptive statistics for each algorithm

	Persistence Model	RF	LSTM	Transformer
nRMSE	22.08%	22.00%	10.56%	24.46%
nMAE	15.14%	16.26%	7.47%	19.73%

RF when dealing with sparse data, although the input size of RF is longer compared to LSTM. In addition, attention should be drawn to the fact that the summarized results represent the capability of each machine learning algorithm under the current tuned hyperparameter settings in this scenario. For the model classification and utility comparison in Section 4.2.3, the LSTM generated results with the smallest error will be utilized.

4.2.3 Modeling

In this section, the employed model classification is introduced and different options for modeling a GSHP with two HWSs with varying degree of complexity are discussed, which are used for the evaluation. By using the modified models, five different model classes (A, B, C, D and E) with decreasing complexity in terms of the number of required parameters are defined. All model classes utilize Equation (4.12) to calculate the thermal power with the generated average flow rate to further obtain the electrical power, while Model A considers the energy changes in both HWSs, Model B and Model C neglect the impact of the small and the large hot water tank respectively. Moreover, Model D is further simplified by ignoring the energy changes in both storage. The

last Model E directly uses the average COP to calculate the consumed electrical power. Table 4.8 presents the model classification and the number of required parameters and an overview of the individual parameters that apply to each model class is given in Table 4.9.

Table 4.8: Model classification with respect to parameters

Model Class	**Combination**	**Number of required Parameters**
Model A	(4.12)(4.13)(4.14)(4.16)(4.17)(4.18)	16
Model B	(4.12)(4.13)(4.14)(4.16)(4.18)	14
Model C	(4.12)(4.13)(4.14)(4.16)(4.17)	13
Model D	(4.12)(4.13)(4.14)	9
Model E	(4.12)(4.15)	7

Table 4.9: Overview of the applied parameters to each model class

Parameter	**Model A**	**Model B**	**Model C**	**Model D**	**Model E**
T^{supply}	■	■	■	■	■
T^{return}	■	■	■	■	■
$\dot{V}_w$	■	■	■	■	■
$\dot{m}_b$	■	■	■	■	
T^{out}	■	■	■	■	
T^{in}	■	■	■	■	
T^{bottom}	■	■			
T^{25cm}	■	■	■		
T^{50cm}	■	■	■		
T^{100cm}	■		■		
T^{top}	■	■			
ρ_w	■	■	■	■	■
c_w	■	■	■	■	■
c_b	■	■	■	■	
V_s	■	■	■		
V_l	■				
P_Q					■
P					■

4.2.4 Results and Discussions

As explained in Section 3.3.2, quantification of the model utility by utilizing accuracy is dependent on the particular application being considered. In this use case, the quantification of the utility of the models is modified with the new definition in Equation (4.19), where U represents the utility of a model in percentage. The reason to use nMAE instead of MAPE as described in Section 4.1.5 is that the ground truth in this use case contains zeros, which makes the calculation of MAPE infeasible.

$$U = (1 - (\text{nMAE})) \cdot 100[\%] \tag{4.19}$$

As mentioned in Section 4.2.2, a time horizon of 168 hours, namely a consecutive 7-day period in winter, is determined for the simulation and analysis, while this application scenario focuses on space heating and hot water supply. Besides, different from the initialization in the first use case in Section 4.1, the initial value of the consumed electrical power is calculated by utilizing the generated flow rate. Fig. 4.14 shows the results of different models along with the differences between them and the ground truth.

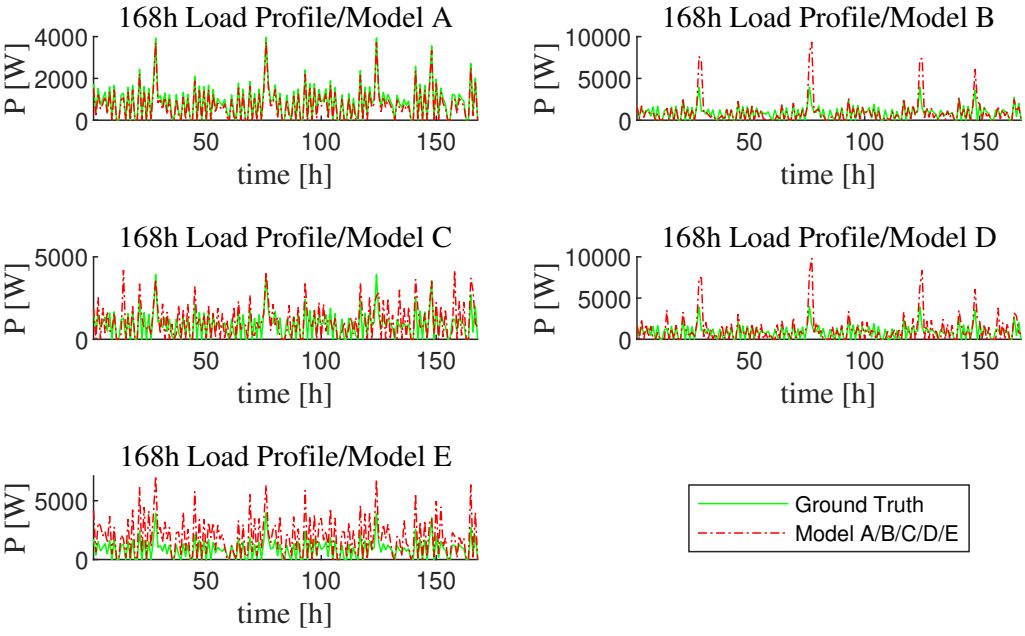


Figure 4.14: Comparison between model results and measured results (Li et al. 2024a)

The diagram shows that the results of Model A are the closest to the measured results, whereas Model B and Model D show several large deviations at some time steps as shown in some tips of the curve. What these two models have in common is that neither considers the energy changes in the small storage for domestic hot water. Therefore, one possible reason for this behavior is that the usage patterns of the domestic hot water are more dynamic than heating. In addition, the simplest Model E in our case presents a larger value than the ground truth in most cases, which could be caused by the underestimated average COP in Equation (4.15), since COP is equal zero when the heat pump is turned off.

In order to describe the overall statistic features of the simulation results and the utility of the models as defined above, the nMAE and the corresponding U are calculated, yielding the results presented in Table 4.10. Model A, with the highest complexity in terms of the required parameters, has the lowest nMAE of 3.77% compared to other four model classes and thus has the highest utility among all models. Besides, it's worth noting that Model B has a lower nMAE than Model C despite the large deviations at some time steps, which means the overall impact of the large hot water storage is greater than that of the small one.

Table 4.10: nMAE and Utility of each model class

	Model A	Model B	Model C	Model D	Model E
nMAE	3.77%	10.62%	16.99%	20.78%	26.79%
U	96.23%	89.38%	83.01%	79.22%	73.21%

With the definition in Equation (4.19), the relationship between the utility and the complexity of all five model classes is illustrated in Fig. 4.15. This demonstrates that the results with a longer time horizon of 7 days are further verifying the proposed hypothesis in Section 3.2, which is that the complexity-utility relationship in the field of DSM modeling could be represented by a diminishing marginal utility curve. However, it is important to note that the graph line is not as smooth as an approximated diminishing marginal utility curve described by a polynomial curve of degree 2, which is also presented in orange dashed line as a reference in Fig. 4.15. Moreover, the large deviation between the simulation and the approximation results, such as the data point of Model C, reveals that there could exist gaps between the simulation and an ideal value by approximation when the model might neglect key elements or parameters that are essential to system accuracy during simplification. For instance, Model C neglects the impact of the larger HWS during model simplification and this results in the large deviation in Fig. 4.15. In comparison to the results in Section 4.1.5, it can be observed that the consideration of a longer time horizon, namely more data points, inherently mitigates the possible negative effect of individual outliers on the overall results. The main reason why the curve in Fig. 4.7 exhibits a significantly steeper

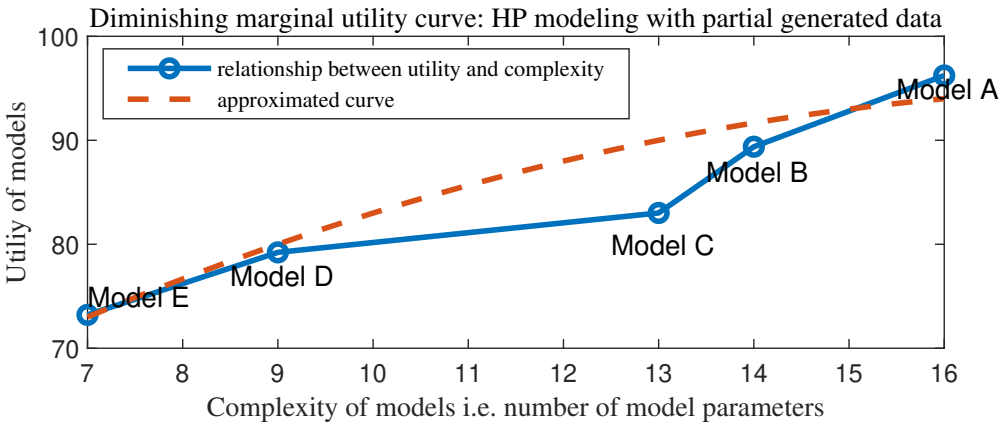


Figure 4.15: Diminishing marginal utility curve based on the complexity of models (Li et al. 2024a)

increase at the beginning compared to the curve line in this use case is that the improvement of the utility is more significant between the first two relatively simplified model classes in the first use case than in this use case. And this could potentially be caused by the limited number of data points in the first use case. Moreover, compared to the first use case, one key variable, namely the flow rate, in this use case is generated based on LSTM, which could also have an impact on the marginal utility curve. Furthermore, due to the zero values of electrical power in our dataset, the definition of utility of models is modified in the present use case compared to Section 4.1. Both results underline that quantification of model utility is dependent on the utilized datasets. Based on the results in Fig. 4.15, which are basically in line with the diminishing marginal utility curve, the proposed hypothesis is further verified.

4.3 Aggregated Load of Charging Stations Case Study

Parts of this chapter are adapted or reproduced from the author's own previous works:

Chang Li, Miao Zhang, Kevin Förderer, Jörg Matthes and Veit Hagenmeyer.
"A comparative analysis of machine learning algorithms for aggregated electric chargepoint load forecasting".
In: *9th International Conference on Sustainable and Renewable Energy Engineering*, 2024, 545. doi: 10.1051/e3sconf/202454501004.

In the two use cases mentioned above, modeling is carried out based on white-box models. However, as discussed in Section 3.3.1, this modeling approach is not applicable to all scenarios, such as black-box models for forecasting. Therefore, it is necessary to investigate another type of use case based on black-box models and validate the proposed hypothesis in this scenario.

In this section, a use case based on aggregated load forecasting of charging stations for EVs is investigated. Due to the uncoordinated deployment of charging stations (Chang et al. 2021) and the system fluctuations regarding charging behaviors (Gong et al. 2020), the charging environment is dynamic (Alsabbagh et al. 2020). This makes the use of traditional modeling methods such as white-box models difficult for an accurate analysis and forecast (Chen et al. 2022), which is essential for operational decision making such as demand side management. For instance, the results of short-term load forecasting can help utilities to optimize generation and to ensure grid stability in the short term. This section consists of three parts, namely the utilized dataset, the modeling process based on machine learning models and the results as well as discussions. The utilized methodology for forecasting has been described in Section 4.2.1.2 and will not be repeated here.

4.3.1 Dataset

> The raw data for the experimental statistics on the usage of domestic electric vehicle chargepoints are released by Department for Transport in the UK in December 2018. The raw data contain 3.2 million charging events recorded across approximately 25,000 funded domestic chargepoints through the whole year of 2017 (Department for Transport 2018). This dataset is found to be useful in other research topics such as the performance

analysis of battery-assisted charging (Ali et al. 2020) and demand response (Fang et al. 2020). Table 4.11 shows a part structure of the raw data.

Table 4.11: Part structure of the raw data

Start Date	Start Time	End Date	End Time	Energy [kWh]	Plugin Duration [h]
2017-12-31	23:59:23	2018-01-01	18:20:23	8.8	18.35
2017-12-31	23:59:00	2018-01-01	00:03:00	10.2	0.07
2017-12-31	23:59:00	2018-01-01	13:40:00	6.2	13.68
...	...	...	...	...	...

Based on the date and time, the raw data are pre-processed to generate aggregated load demand by hour. In the preprocessing, any plug-in events that were less than 3 minutes in length are treated as anomalies and therefore excluded. Besides, it's assumed that the charging power is constant throughout the plugin duration based on empirical observation and the simplified piecewise-linear charging profile model in (Zhang et al. 2012).

In addition, attention should be drawn to the fact that in actual charging scenarios, the power drawn by each individual vehicle is not constant through the whole charging process. However, since this use case focuses on the statistical results of a large volume of raw data rather than the data of an individual vehicle, the simplification can be considered reasonable for practical DSM applications.

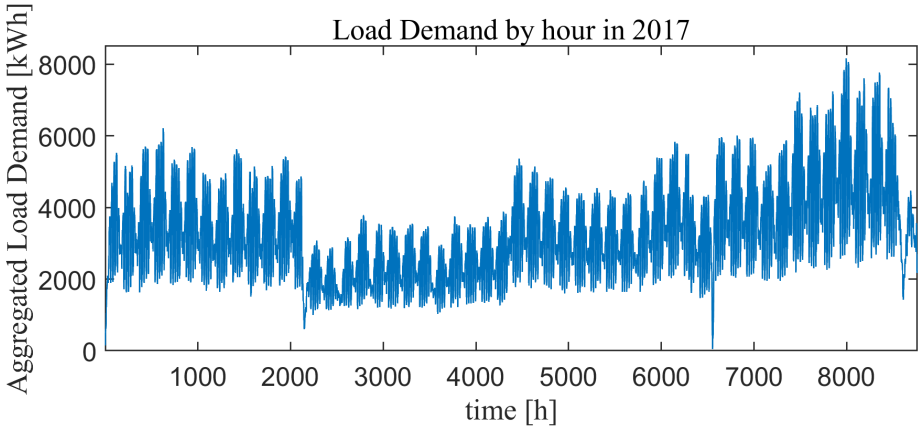
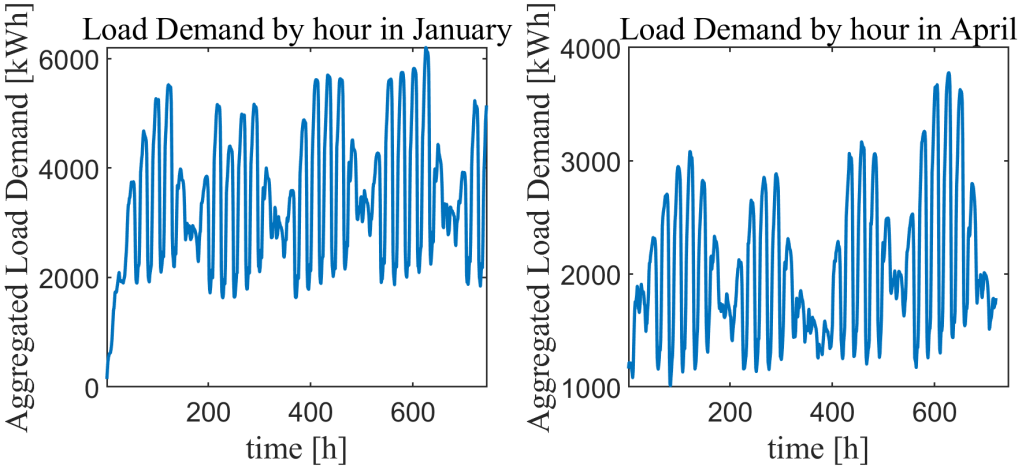
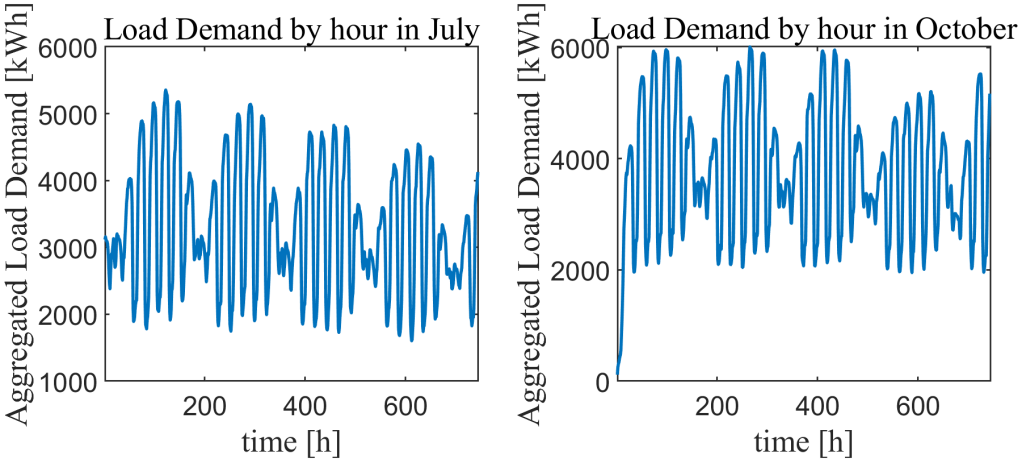


Figure 4.16: Aggregated Load Demand by hour in the year 2017 (Li et al. 2024b)

All pre-processed data for the year 2017 are presented in Fig.4.16. To enhance data diversity while maintaining the efficiency of subsequent model training and forecast, the data of January, April, July and October as representative months in each season are extracted separately as inputs, which are presented in Fig.4.17.



(a) Aggregated Load Demand by hour in January and April 2017 (Li et al. 2024b)



(b) Aggregated Load Demand by hour in July and October 2017 (Li et al. 2024b)

Figure 4.17: Aggregated Load Demand by hour in January, April, July and October of 2017 as representative months in each season

Then the first 744 hours in January, April, July and October respectively are used as training set with cross-validation for model training and the next 72 hours of data are reserved for test. By

utilizing the descriptive statistics, i.e., nRMSE and MAPE, the results of forecasting models with different complexity are evaluated.

4.3.2 Modeling

As mentioned in Section 4.2.1.2, RF is a commonly-used and powerful machine learning algorithm in many classification and regression problems. Moreover, as discussed in Section 3.3.1, another advantage of RF compared to other machine learning algorithms is that its input size can be directly employed for quantifying model complexity for regression problems. For time series forecasting, it requires that the time series dataset is transformed into a supervised learning problem first (Li et al. 2024b). Fig. 4.9 has already shown this transformation process, i.e., sliding window, with an input size of one as an example, where Y is the value at each time step. In this use case, the forecasting model is built on RF and the complexity of different models is quantified by using different input sizes, as discussed in Section 3.3.1, to create a bagged regression ensemble object. In order to investigate the relationship between the black-box model complexity and the model utility, RF with varying input sizes are applied. Finally, five bagged regression ensemble objects are created, with an input size from 1 to 5 together with the temporal features of days such as Monday, Tuesday etc. as the extra input, to use the bootstrap aggregation method for model training. Equation (4.20) presents the structure of the features of different input sizes, where n is the number of input sizes and 1 to 10 mean each month respectively. These models are utilized to support the investigation of the hypothesis by generating the utility-complexity curve.

$$X = \begin{bmatrix} Feature_{11} & Feature_{12} & ... & Feature_{1n} & Feature_{temporal1} \\ Feature_{41} & Feature_{42} & ... & Feature_{4n} & Feature_{temporal4} \\ Feature_{71} & Feature_{72} & ... & Feature_{7n} & Feature_{temporal7} \\ Feature_{101} & Feature_{102} & ... & Feature_{10n} & Feature_{temporal10} \end{bmatrix} \tag{4.20}$$

4.3.3 Results and Discussions

In this use case, the quantification of the utility of the forecast models is modified with the new definition in Equation (4.21), where U represents the utility of a forecast model in percentage and avgMAPE is the mean value of MAPE in each representative month of the year.

$$U = (1 - (\text{avgMAPE})) \cdot 100[\%] \tag{4.21}$$

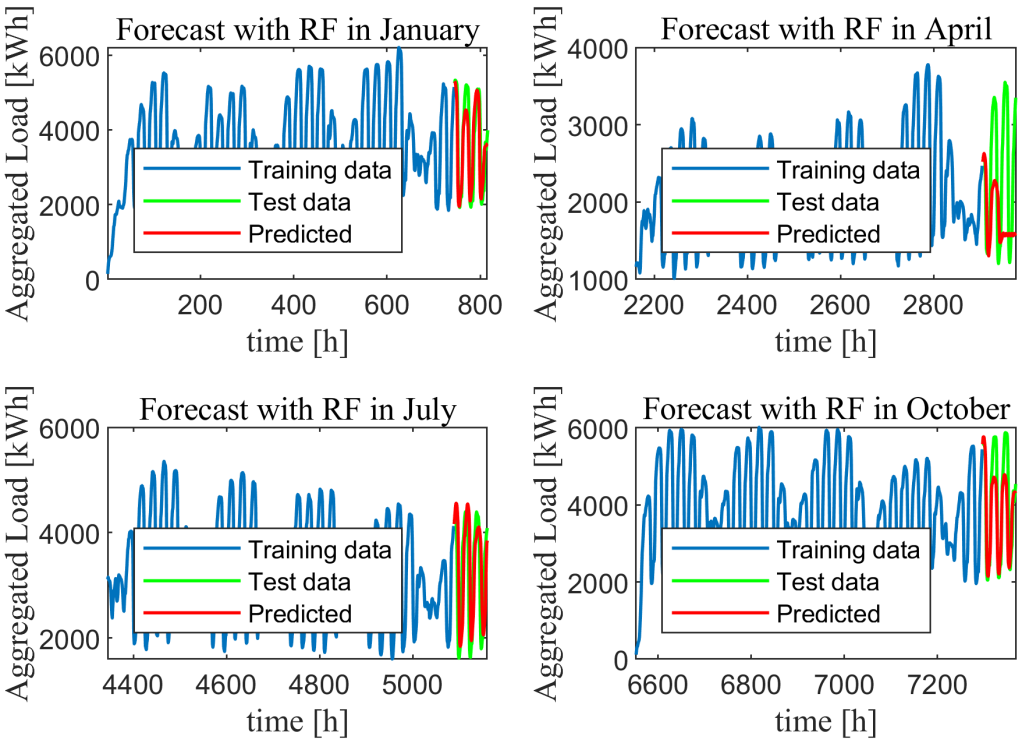


Figure 4.18: Load demand forecast results with an input size of 5

As mentioned in Section 4.3.1, the predictive horizon of 72 hours is set up for each month. In order to illustrate the performance of the forecasting models, Fig. 4.18 shows the results when the input size is set to 5, i.e., $n = 5$ in Equation (4.20). First of all, the diagram shows that the patterns of the data have a huge impact on the forecasting accuracy of forecast models. For instance, in this use case, the most complex model demonstrates strong predictive performance in January, July and October, yet its accuracy significantly deteriorates during April. Therefore, the average of all months in Equation (4.21) effectively attenuates the influence of these pattern variations. Besides, RF cannot extrapolate. It means that predicted values are always within the range of the training set, which is also confirmed by the results.

In order to describe the overall statistic features of the forecasting results and the utility of the models as defined above, the avgMAPE and the corresponding U are calculated, yielding the results presented in Table 4.12. The model in the rightmost column, with the highest complexity in terms of the input sizes (excluding the temporal feature), has the lowest avgMAPE of 19.37% compared to other four forecast models and thus has the highest utility among all forecast models.

Table 4.12: avgMAPE and Utility of forecast models with different input sizes, excluding the temporal feature

Input Size	1	2	3	4	5
avgMAPE	46.46%	36.60%	28.43%	24.10%	19.37%
U	53.54%	63.40%	71.57%	75.90%	80.63%

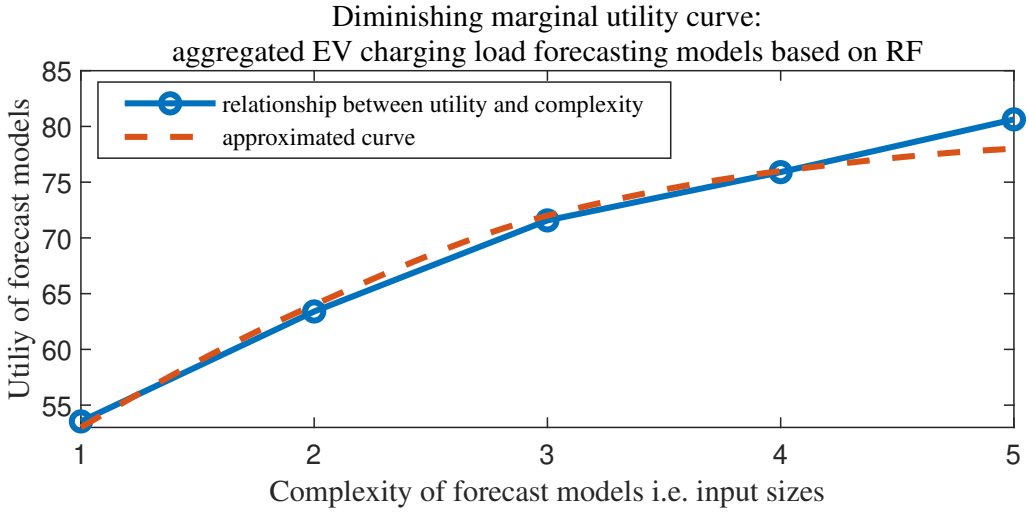


Figure 4.19: Diminishing marginal utility curve based on the complexity of models

With the definition in Equation (4.21), the relationship between the utility and the complexity of all five forecast models are illustrated in Fig. 4.19. This demonstrates that the results in the context of black-box models are further verifying the proposed hypothesis in Section 3.2, which is that the complexity-utility relationship in the field of DSM modeling could be represented by a diminishing marginal utility curve.

However, it should be noted that the graph line is again not as smooth as an approximated diminishing marginal utility curve given as a polynomial curve of degree 2, which is presented in orange dashed line as a reference in Fig. 4.19. The relative large deviation between the simulation and the approximation results, such as the data point when the input size is equal one, reveals that there could exist gaps between the simulation and an ideal value by approximation, which is reasonable for an over simplified forecast model.

Furthermore, as the same algorithm and simulation hardware are employed for different forecasting models across this use case, it is also worthwhile to investigate the relationship between model utility and the computation time. Fig. 4.20 presents this relationship, where the horizontal axis indicates the computation time in seconds and the respective input sizes of the models are

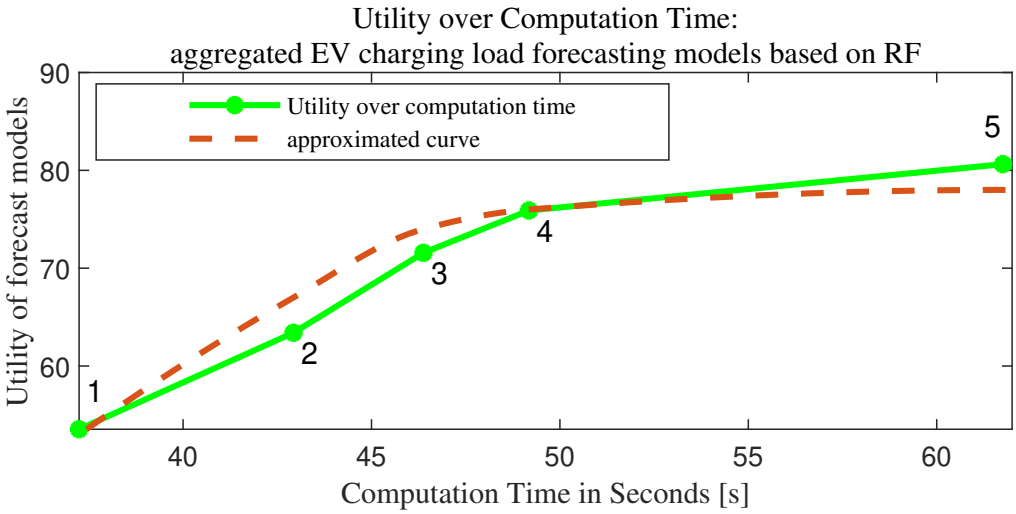


Figure 4.20: Utility over computation time based on RF models

annotated in the figure. It can be seen that the results also exhibit a diminishing marginal utility trend. However, as explained in Section 3.3.1, there is a limitation of utilizing the computation time compared to using the input size as the horizontal axis. The computation time does depend on specific algorithms and the simulation environment, which makes it challenging to quantify model complexity before conducting simulations or experiments. In contrast, the complexity of forecast models can be already quantified during modeling by utilizing the input size.

In addition, it can be observed that when the input size increases from 4 to 5, the computation time rises sharply, whereas the previous computation time's increases are nearly linear regarding the increasing input sizes. This indicates that the internal computational complexity of the algorithm RF, when used to construct forecasting models, does not grow linearly with the input size. This finding can further assist decision-makers in selecting a model of appropriate complexity, particularly when interpreted in conjunction with the diminishing marginal utility curve presented in Fig. 4.19.

4.4 Discussion Summary

In this chapter, three use cases are carried out to validate the proposed hypothesis and answer the RQ3: What is the relationship between the model complexity and the model utility of DERs models in the context of DSM?, based on the accuracy in off-line simulations. The first two use cases focus on white-box DER models for GSHP systems. The evaluation of the models by

using various data sources demonstrates the validity of the proposed hypothesis for presenting the relationship between the model complexity and the model utility in the modeling of GSHP, as a type of DERs, for DSM. However, the results also show that the real graph lines in both use cases are not as smooth as an ideal diminishing marginal utility curve. One reason for this is that the quantification of the model complexity is not mathematically continuous, which limits the smoothness of the real curve. Another reason is that the approximated diminishing marginal utility curve is an idealized curve itself, which is more intended to express a trend of a relationship between two variables. In addition, the curve in the second use case aligns more closely with the ideal curve compared to that in the first use case. This could be caused by the different number of data points that are used in the simulations. This additional result could provide suggestions for future experiments that a longer time horizon of investigation might help to minimize the effect of individual outliers on the overall analysis.

After validating the hypothesis with white-box DER models, the third use case looks into modeling with black-box DER models in order to deal with the scenarios in DSM where the white-box models are less suited, such as load forecasting of EV charging stations. By leveraging an extensive real-world dataset for model training, the results are promising and in line with the proposed hypothesis as well. It's worth noting that the computation time can also be used as a quantification method for complexity when utilizing the same algorithm for forecast models. In addition, both real curves in this use case also have gaps to their respective approximated ideal curves, which can be attributed to the same reasons mentioned above.

Overall, the evaluation with three use cases shows the validity of the proposed hypothesis to present the relationship between model complexity and model utility for DERs in DSM, which answers the RQ3 directly. The proposed hypothesis can be used to identity the relationships across various practical applications by developers, engineers or operators in order to provide a quantified reference for different DSM applications.

Lastly, given the results of this chapter, there are several areas that future works can improve on. As the comparison between the first and the second GSHP use cases shows, the time horizon of the experiment, i.e., the number of data points, have an impact on the results. To further improve the validity of the hypothesis, future experiments or investigations should cover a time horizon that is longer than 7 days. Furthermore, the hypothesis has to be further investigated and validated with other possible DERs on the demand side, e.g. battery systems, to broaden its scope of application. Moreover, new explainable methods for quantifying the complexity of other black-box models need to be investigated to further validate the hypothesis in the scenarios with black-box models.

5 Quantification of the Contribution of Distributed Energy Resources Submodel in Demand Side Management

This chapter takes a step further and aims at addressing RQ4: "How to quantify the contributions of a DER's submodel to model utility in DSM?" systematically by introducing a quantification method for determining the contribution of a DER's submodel in the context of DSM based on the Shapley value. As mentioned in Section 1.3, the goal of DSM in general is to reschedule energy consumption on the demand side to electricity production and electrical load on the power grid. In other words, DSM should contribute to improve the flexibility in the overall energy system. However, the value of flexibility in DSM is a measure that is challenging to quantify. Furthermore, for DERs models composed of several submodels, assessing the individual contribution of each submodel to the overall accuracy is of significant value for informed decision-making. To deal with this challenge, this chapter introduces an approach for determining the contribution of DERs submodels by considering the maximal power of a DER in the context of DSM based on the Shapley value. This quantification method is another step forward for confirming the hypothesis of diminishing marginal utility of model complexity in practical DSM applications. The remainder of this chapter is structured as follows: Section 5.1 introduces the theoretical foundation of Shapley values and then related works are discussed in Section 5.2. In Section 5.3, a case study for a heat pump system is carried out to demonstrate the application of the proposed approach. Finally, the results are discussed in Section 5.4, providing insights into potential applications.

5.1 Theoretical Foundation

In 1951, the mathematician Lloyd Shapley has introduced a solution concept in n-Person Game, i.e., cooperative game theory, for quantifying each player's contribution by evaluating the marginal

impact they make when added to all possible subsets of players, and computing the average of these impacts (Shapley 1951). This quantified value is also known as the Shapley value. This solution concept has proved to be the only solution that satisfies all four foundational properties, namely efficiency, symmetry, linearity and null player, in 1953 (Kuhn and Tucker 1953). This establishes the theoretical groundwork for the method's future use in a wide range of disciplines.

According to (Winter 2002), the Shapley value φ of player i is calculated by using the following equations, where S is a subset or a coalition of the whole set N of n players and v is the value function to determine the value or worth of a coalition. And $|S|$ represents the number of players in this coalition S, whose cardinality is between 0 and $n-1$.

$$\varphi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} (v(S \cup \{i\}) - v(S)) \tag{5.1}$$

$$= \frac{1}{n} \sum_{S \subseteq N \setminus \{i\}} \binom{n-1}{|S|}^{-1} (v(S \cup \{i\}) - v(S)) \tag{5.2}$$

As mentioned above, the Shapley value satisfies four foundational properties. The first property is efficiency, which means the sum of Shapley values of all players equals the value of the whole set N. This property guarantees that the gain is distributed among all players. Equation (5.3) (Kuhn and Tucker 1953) formulates this property.

$$\sum_{i \in N} \varphi_i(v) = v(N) \tag{5.3}$$

The second property is symmetry, which means that if two different players i and j contribute equally to all possible coalitions, then their Shapley values are also equal. This property guarantees that the players are treated equally if contribute equally. Equation (5.4) (Kuhn and Tucker 1953) formulates this property, where S is an arbitrary coalition without i and j.

$$v(S \cup \{i\}) = v(S \cup \{j\}) \Rightarrow \varphi_i(v) = \varphi_j(v) \tag{5.4}$$

The third property is linearity, which means the Shapley values of the player i in different value functions are linear and additive. This property considers mainly the situation where more value functions are involved in the analysis. The following equations (Kuhn and Tucker 1953) formulate this property, where v and w are different value functions and a is any real number.

$$\varphi_i(v + w) = \varphi_i(v) + \varphi_i(w) \tag{5.5}$$

$$\varphi_i(av) = a\varphi_i(v) \tag{5.6}$$

The last foundational property is the so-called null player or dummy player, which means that if player i does not contribute to any coalition in terms of the value function v, then the Shapley value of i is zero and this player i is the so-called dummy player or null player. This property further guarantees the fair contribution of each player when evaluating the marginal impact they make. Equation (5.7) (Kuhn and Tucker 1953) formulates this property, where S is an arbitrary coalition without i.

$$v(S \cup \{i\}) = v(S) \Rightarrow \varphi_i(v) = 0 \tag{5.7}$$

It is precisely because the Shapley value satisfies these four foundational properties that it has found applications not only in game theory, but also in a wide range of other fields with the idea of "players" being replaced with other domain concepts. Some of them are discussed in the following section.

5.2 State of the Art

In recent years, the concept of the Shapley value, initially proposed within game theory, has found widespread applications in various domains beyond its original context, such as explainable Artificial Intelligence (AI), cost allocation, energy systems and so on (Varenhorst et al. 2024). For instance, the authors (Lundberg and Lee 2017) have proposed a unified framework named as SHapley Additive exPlanations (SHAP), which is based on Shapley value, for interpreting predictions. In 2020, they further improve the interpretability of tree-based models, such as random forests, decision trees, and gradient boosted trees, by presenting TreeExplainer, an explanation method for trees in these models based on Shapley value (Lundberg et al. 2020). However, other authors (Kumar et al. 2020) point out that some further assumptions, such as the feature independence, are needed to use this framework for interpretation. This could lead to an inaccurate approximation to the true Shapley values. To deal with the dependence of features in machine learning models, (Aas et al. 2021) have improved the Kernel SHAP method by estimating the conditional distribution to incorporate dependence. This improvement has led to a more accurate approximations to the true Shapley values than the original Kernel SHAP approach. In 2023, the team that developed SHAP has published a review article further discussing

algorithms to generate Shapley value explanations (Chen et al. 2023). By comparing 24 distinct algorithms, the authors highlight the key innovations in recent approaches for estimating Shapley value explanations and give recommendations, such as the importance of the feature-removal approach and the specific estimation strategy, for the utilization of Shapley value explanations in both industry and academia (Chen et al. 2023). Another survey carried out by (Luo and Pei 2024) focuses on methods for computing the Shapley value in the domain of databases and machine learning, where they also point out the idea of "players" are being tuples, tables, features, data samples and models in specific applications.

In the domain of energy systems, the Shapley value has also proven to be highly applicable as well. For approximating the Shapley value in realistic community energy settings, a team in the Netherlands has compared several existing methods and proposed a more efficient one by clustering consumers into a smaller number of demand profiles (Cremers et al. 2022). In other application, the authors from Norway (Cai et al. 2023) have presented an Energy Management (EM) strategy for residential microgrid systems using Model Predictive Control (MPC)-based Reinforcement Learning (RL) and Shapley value, where the Shapley value approach is applied as a feasible solution for distributing the collective cost. Considering energy system operators, a distributed approach of Shapley value calculation for the redispatch congestion cost allocation to deal with data privacy concerns with respect to real-world implementation is briefly discussed in (Bauer et al. 2023). For device flexibility in DSM, the authors (Varenhorst et al. 2024) have proposed a method to determine the value of flexible assets using average marginal contributions, based on the Shapley value. In their scenario, the optimization goal is to minimize the Euclidean norm of a single smart house profile using DSM.

By reviewing the above numerous recent publications related to the Shapley value, it's observable that most of them are focused either on the methods for computing the Shapley value or on the application in black-box models, which have already made a lot of progress. However, in the context of white-box models for DERs in DSM, the potential of the Shapley value remains to be explored. Therefore, in the next section a novel approach for quantifying the contribution of DERs submodels based on the Shapley value together with a use case will be presented and discussed.

5.3 Heat Pump and Heat Pump Storage Case Study

The goal of this section is to propose a new approach for quantifying the contribution of DERs submodels in the context of white-box DER models based on the Shapley value and to show how the approach can be applied in the context of DSM with the use case mentioned in Section 4.2.

5.3.1 Quantification Approach

In Chapter 4, the first two use cases both employ white-box models to describe GSHP systems. In defining the model utility U, the simulation results derived from different model classes are compared with the actual measurements by using metrics such as MAPE and nMAE. In order to quantify the contribution of submodels based on the Shapley value in the following analysis, an essential prerequisite lies in how the value function v is defined. Considering that, a device's ability to contribute effectively to load profile adjustment, i.e., flexibility adjustment, in DSM is closely related to its maximum power P_{max}. Therefore, the following definition for the value function v is first proposed in our scenario, where S is the coalition of $submodel_i$. It is important to note that the $submodel_i$ here has a different meaning than the model class mentioned in Chapter 4. It represents different mathematical models which are used to build a model class. In other words, equations such as Equation (4.12), Equation (4.13) form the $submodel_i$ here. Under this definition, the value function v represents the realistically reliable maximum adjustable power of a coalition S to load adjustment in DSM.

$$v(S) = v(\sum_{S} submodel_i) = P_{max} \cdot U(S) \tag{5.8}$$

Then the Shapley value of a coalition S, i.e., a white-box model with submodels in our scenario, can be calculated according to Equation (5.2).

5.3.2 The Shapley Value for Heat Pump and Hot Water Storage Submodels

In this section, the attention is focused on how the proposed approach can be applied to the use case with white-box DERs models. As mentioned at the beginning of Section 5.3, the detailed set up of the use case is already described in Section 4.2, where the $P_{max} = 12.5\ kW$. Table 4.8 presents the model classification results with five model classes, among which Model E is too abstract and neglects the impact of HWSs. Therefore, the objectives of the following valuation include Model A to Model D. For comprehensibility of the following explanation, an excerpt of Table 4.8 is presented below.

As discussed in Section 5.1 and Section 5.2, the idea of "player" in game theory can be replaced with other concepts depending on the use case. As shown in Table 5.1, it is evident that Equations (4.12) (4.13) (4.14) are required for all four model classes. Based on the proposed concept, the model classes can be conceptualized as a n-Person Game, where n is equal to 4. The

detailed mapping is presented in Table 5.2. Table 5.3 then shows the conceptualization of Model Class as a coalition and its corresponding value based on Equation (5.8).

Table 5.1: Model classification with respect to parameters, an excerpt of Table 4.8

Model Class	**Combination**	**Number of required Parameters**
Model A	(4.12)(4.13)(4.14)(4.16)(4.17)(4.18)	16
Model B	(4.12)(4.13)(4.14)(4.16)(4.18)	14
Model C	(4.12)(4.13)(4.14)(4.16)(4.17)	13
Model D	(4.12)(4.13)(4.14)	9

Table 5.2: Mapping of $submodel_i$ based on the proposed concept

Equations	**"Player"** $submodel_i$
(4.12)(4.13)(4.14)	$submodel_1$
(4.16)	$submodel_2$
(4.17)	$submodel_3$
(4.18)	$submodel_4$

Table 5.3: Conceptualization of Model Class based on the proposed concept

Model Class	**Coalition**	**Value of coalition** $v(S)$
Model A	$submodel_1$, $submodel_2$, $submodel_3$, $submodel_4$	12.03 kW
Model B	$submodel_1$, $submodel_2$, $submodel_4$	11.17 kW
Model C	$submodel_1$, $submodel_2$, $submodel_3$	10.38 kW
Model D	$submodel_1$	9.90 kW

According to Equation (5.2), the Shapley value of each $submodel_i$ can be computed step by step. Given that n is equal to 4 in this use case, the number of computational steps remains manageable, thereby enabling a visualization of the full calculation procedure. For instance, Table 5.4 presents the full calculation procedure of the Shapley value of $submodel_1$, where $\varphi_{submodel_1} = 10.58kW$ when the final result is rounded with two decimal precision. Similarly, the Shapley value of the rest $submodels$ can be computed and the full steps will not be repeated here. The results of the Shapley value of all four $submodels$ are printed in Fig. 5.1.

Table 5.4: The Shapley value kW of $submodel_1$

Subset S	$v(S \cup \{i\})$	$v(S)$	$\lvert S \rvert$	**Weight**	**Weighted Contribution**
ϕ	9.90	0	0	1/4	2.475 kW
$submodel_2$	9.90	0	1	1/12	0.825 kW
$submodel_3$	9.90	0	1	1/12	0.825 kW
$submodel_4$	9.90	0	1	1/12	0.825 kW
$submodel_2$, $submodel_3$	10.38	0	2	1/12	0.865 kW
$submodel_2$, $submodel_4$	11.17	0	2	1/12	0.931 kW
$submodel_3$, $submodel_4$	9.90	0	2	1/12	0.825 kW
$submodel_2$, $submodel_3$, $submodel_4$	12.03	0	3	1/4	3.008 kW
$\varphi_{submodel_1}$					10.58 kW

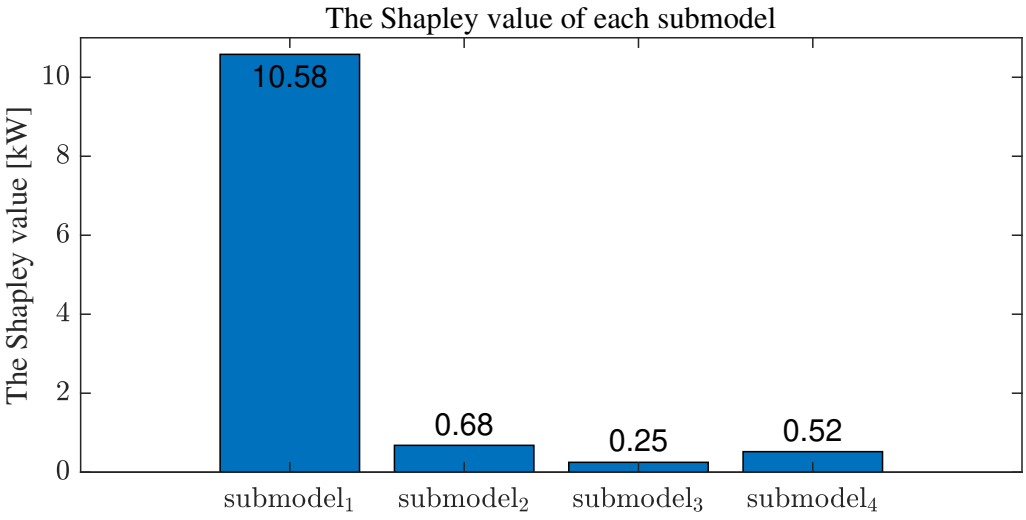


Figure 5.1: The Shapley value of each model in kW

With the calculated Shapley value, the contribution of each $submodel$ in the coalition is quantified. Based on the results, the satisfaction of four foundational properties mentioned in Section 5.1 needs to be verified first. The first and most important property, namely efficiency, can directly be verified by aggregating the Shapley value of all submodels, which is equal to 12.03 kW exactly.

Table 5.5 shows the verification. The second and the third property, symmetry and linearity, do not apply to this use case, as no two submodels' Shapley values are equal and the value function is uniquely defined. The last property is also not applicable to this case since each $submodel$ ("player") has a certain contribution to the coalition. More detailed discussions are presented in the following section.

Table 5.5: Verification of the efficiency property

$submodel_i$	**Shapley value** kW	
$submodel_1$	10.58	
$submodel_2$	0.68	$\sum_{i \in \text{Model } A} \varphi_i(v) = 12.03$
$submodel_3$	0.25	
$submodel_4$	0.52	
Value of coalition Model A kW		$v(Model\ A) = 12.03$

5.4 Discussion

Based on the results in Fig. 5.1, it's clear to see that the Shapley value of $submodel_1$ is markedly higher than that of the remaining $submodels$ in the whole set, i.e., Model A. This finding indicates that $submodel_1$ has a dominant influence in the formulation of Model A when using Model A to evaluate the flexibility of DSM assets. This result is also expected, as this submodel is essential across all model classes. Furthermore, the Shapley value of $submodel_4$ is more than twice that of $submodel_3$. And in this scenario, $submodel_3$ and $submodel_4$ interpret the impact of the small and the large HWS respectively. Based on the approach, the contribution of $submodel_4$, that is, the large HWS, could be quantified as 2.08 times that of $submodel_3$, the small HWS. Combined with their Shapley values, their impacts on the whole coalition, which is defined considering the maximal power of the GSHP system, are also quantified.

In addition, it is important to highlight how this methodology differs from another approach, namely sensitivity analysis, which studies the uncertainty in the output of a mathematical model or system can be apportioned to different sources of uncertainty in its inputs (The Joint Research Centre: EU Science Hub 2025). It should be more considered as a pre-requisite for statistical model (Saltelli et al. 2004) or used to test the robustness of the results of a model or system (Gutierrez-Garcia et al. 2022).

Overall, the evaluation with the use case shows the feasibility of the proposed methodology to quantify the contribution of submodels within white-box DER models in DSM based on the Shapley value, which answers the RQ4 directly. The methodology can be used together with the proposed hypothesis in Chapter 4 to further assist developers, engineers or operators, not only in selecting a model with an appropriate level of complexity based on utility requirements, but also facilitating a deeper understanding of which $submodel$ within the model class contributes most to the overall flexibility in DSM, particularly when the model class can be viewed as a coalition.

Lastly, given the results of this chapter, there are different areas that future works can improve on. In defining the value function in this case study, the maximum power of the GSHP system was taken into account. This is because the maximum power determines the theoretical upper limit of the adjustable capacity that a DER can provide when participating in DSM. However, the value function could be redefined according to different scenarios, allowing it to be integrated with other variables such as energy costs. This could further extend the applicability of the approach. Furthermore, an extension of this approach to encompass other model types, such as gray-box models, might serve as a potential improvement.

6 Conclusion and Outlook

6.1 Conclusion

In the present dissertation, the influence of DER model complexity on their utility for DSM is investigated. The study begins with a comprehensive survey on the current state of DSM implementations in the residential, commercial and industrial sector, focusing on their models and approaches. One important finding of the review is that, in many practical DSM/DR applications, simple models such as linear models can already yield highly satisfactory results. To support better decision-making for developers, engineers or operators by providing a quantified reference, it's worthwhile to investigate this relationship between the model complexity and the model utility quantitatively, which leads to:

RQ1: How to quantify the complexity of DER models in the context of DSM? In Chapter 3, a brief review of other review and application works related to complexity of models and algorithms across different domains is first carried out. In the literature, there are many existing measures to quantify the complexity of models or functions in different scenarios. The quantifiable measures resulting from the review analysis and discussions are summarized in Table 3.1, with their scope of application specified. After comparing the limitations and advantages of different measures and considering our scenarios, the number of parameters and input sizes are chosen for white-box DER models and black-box DER models respectively due to their different model characteristics.

RQ2: How to quantify the utility of DER models in the context of DSM? In Chapter 3, after establishing the quantification method for the model complexity, different quantification methods for the model utility are further analyzed. Since the subsequent validation process is primarily based on offline simulations, the model utility is defined here in terms of model accuracy. Moreover, utility is a more general concept and its definition depends strongly on the specific applications. Therefore, different useful metrics in descriptive statistics according to the definition of "accuracy" in ISO 5725-1 are introduced. These metrics, such as MAPE and nMAE, are subsequently employed in the quantification of the model utility, based on the requirements of different use cases.

After the quantification of model complexity and utility, it leads to:

RQ3: What is the relationship between the model complexity and the model utility of DERs models in the context of DSM? To address the core research question of this thesis, a hypothesis in Chapter 3 is proposed, mainly inspired by Gossen's First Law in economics. The hypothesis states that the complexity-utility relationship in the modeling of distributed energy resources for demand side management could be represented by a diminishing marginal utility curve in general. To validate the hypothesis, three use cases are investigated in Chapter 4, including two heat pump systems based on white-box models and one electric vehicle supply equipment use case based on black-box models. The results of all use cases are basically in line with the hypothesis, although the derivation is relatively large at few individual data points. The simulation results are strongly dependent on the length and quality of the data. A lack of rich data could result in individual outliers exerting an excessive influence on the calculation of the model utility, which could result in large derivation such as excessive curve growth rate. Therefore, it is noted in the discussion of Chapter 4 that extending time horizon represents a significant and necessary avenue for future work and improvement.

For white-box models of distributed energy resources, they could consist of several different submodels. In addition to quantify the model complexity as one whole model, it's also worthwhile to further investigate the specific contributions of each submodel, which leads us to:

RQ4: How to quantify the contributions of a DER's submodel to model utility in DSM? For DERs models composed of several submodels, assessing the individual contribution of each submodel to the overall model utility, namely accuracy in our case study, is of significant value for informed decision-making. In Chapter 5, a new approach based on the Shapley value to quantify the contribution of submodels within a whole white-box model is introduced. In this approach, the maximum power of the ground source heat pump system is incorporated through the definition of the value function. The approach is applied to the heat pump and hot water storage case study, which is described in Chapter 4, to demonstrate its feasibility.

In conclusion, by answering the four research questions discussed above, this thesis contributes to providing energy system developers, modeling engineers and operators with a quantified reference to support decision-making by proposing a novel hypothesis and a new quantification method for the contribution of submodels. The thesis provides a new direction for exploring the relationship between the model complexity and the model utility. This potentially enables a more efficient choice of models for an application. When it comes to a deeper understanding of the contribution of submodels, the presented quantification approach based on the Shapley value enables more straightforward and transparent comparison and analysis.

6.2 Outlook

The proposed hypothesis is validated with three use cases in this thesis. However, there are three main facets that could be further researched and improved in the future work. First of all, the datasets for validation of the first and the second use case cover 1 day and 7 days respectively. In a future experiment, the time horizon should be longer than 7 days to minimize the impact of specific outliers on the overall results. Moreover, datasets from different seasons or years could enrich the data and potentially provide more insights. Secondly, modeling of other energy systems on the demand side, such as home battery system, could provide more validation results and further enable the investigation of the synergy of different energy equipments on the demand side. Thirdly, for systems that are not well-suited to white-box modeling, such as treating a house as one node in the energy network and modeling district heating networks, other model types, that are gray-box and black-box models, should be implemented. To investigate the relationship between the model complexity and the model utility in these scenarios, new quantification methods could be further defined and identified.

Nomenclature

Abbreviations

AI	Artificial Intelligence
AC	Air Conditioner
ASM	Ancillary Service Markets
ASHP	Air Source Heat Pump
BAS	Building Automation System
BIM	Building Information Modeling
CPP	Critical Peak Pricing
CPR	Critical Peak Rebate
CM	Capacity Markets
DSM	Demand Side Management
DR	Demand Response
DER	Distributed Energy Resource
DRA	Demand Response Aggregator
DLC	Direct Load Control
EMPC	Economic Model Predictive Controller
EDR	Emergency Demand Response
EVSE	Electric Vehicle Supply Equipment
FCR	Frequency Containment Reserve
FFM	Fat-free Mass

GA	Genetic Algorithm
GSHP	Ground Source Heat Pump
HP	Heat Pump
HWS	Hot Water Storage
ICT	Information and Communication Technology
IDR	Industrial Demand Response
IEA	International Energy Agency
IBR	Inclining Block Rates
IBDR	Incentive-Based Demand Response
KIT	Karlsruher Institute of Technology
LSTM	Long Short-Term Memory
MPC	Model Predictive Control
MDP	Markov Decision Process
MILP	Mixed-Integer Linear Programming
MIQP	Mixed-Integer Quadratic Programming
MAPE	Mean Absolute Percentage Error
MaxAPE	Maximum Absolute Percentage Error
MDP	Markov Decision Process
nRMSE	normalized Root Mean Square Error
nMAE	normalized Mean Absolute Error
P2H	Power-to-Heat
PSO	Particle Swarm Optimization
PDM	Peak-Demand Management
PTR	Peak Time Rebate
PSBC	Priority Stack-Based Control

RTP	Real-Time Pricing
RL	Reinforcement Learning
RCT	Randomized Controlled Trial
RBC	Rule-Based Control
RF	Random Forest
SHAP	SHapley Additive exPlanations
SG	Smart Grid
TES	Thermal Energy Storage
ToU	Time-of-Use
TSO	Transmission System Operators
V2X	Vehicle to Everything
VPP	Virtual Power Plant

List of Figures

List of Tables

Bibliography

HA Aalami, M Parsa Moghaddam, and GR Yousefi. Demand response modeling considering interruptible/curtailable loads and capacity market programs. *Applied energy*, 87(1):243–250, 2010.

Kjersti Aas, Martin Jullum, and Anders Løland. Explaining individual predictions when features are dependent: More accurate approximations to shapley values. *Artificial Intelligence*, 298: 103502, 2021.

Junade Ali, Vladimir Dyo, and Sijing Zhang. Battery-assisted electric vehicle charging: Data driven performance analysis. In *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 429–433. IEEE, 2020.

Usman Ali, Mohammad Haris Shamsi, Cathal Hoare, Eleni Mangina, and James O'Donnell. Review of urban building energy modeling (ubem) approaches, methods and tools using qualitative and quantitative analysis. *Energy and buildings*, 246:111073, 2021.

Amro Alsabbagh, Brian Wu, and Chengbin Ma. Distributed electric vehicles charging management considering time anxiety and customer behaviors. *IEEE Transactions on Industrial Informatics*, 17(4):2422–2431, 2020.

Peter C Austin, Ian R White, Douglas S Lee, and Stef van Buuren. Missing data in clinical research: a tutorial on multiple imputation. *Canadian Journal of Cardiology*, 37(9):1322–1331, 2021.

Mutiu Shola Bakare, Abubakar Abdulkarim, Mohammad Zeeshan, and Aliyu Nuhu Shuaibu. A comprehensive overview on demand side energy management towards smart grids: challenges, solutions, and future direction. *Energy Informatics*, 6(1):4, 2023.

Arkasama Bandyopadhyay, Benjamin D Leibowicz, Emily A Beagle, and Michael E Webber. As one falls, another rises? residential peak load reduction through electricity rate structures. *Sustainable Cities and Society*, 60:102191, 2020.

Zhejing Bao, Qin Zhou, Zhihui Yang, Qiang Yang, Lizhong Xu, and Ting Wu. A multi time-scale and multi energy-type coordinated microgrid scheduling solution—part i: Model and methodology. *IEEE Transactions on Power Systems*, 30(5):2257–2266, 2014a.

Zhejing Bao, Qin Zhou, Zhihui Yang, Qiang Yang, Lizhong Xu, and Ting Wu. A multi time-scale and multi energy-type coordinated microgrid scheduling solution—part ii: Optimization algorithm and case studies. *IEEE transactions on power systems*, 30(5):2267–2277, 2014b.

Lukas Barth, Nicole Ludwig, Esther Mengelkamp, and Philipp Staudt. A comprehensive modelling framework for demand side flexibility in smart grids. *Computer Science-Research and Development*, 33(1):13–23, 2018.

Rebecca Bauer, Xinliang Dai, and Veit Hagenmeyer. A shapley value-based distributed ac opf approach for redispatch congestion cost allocation. In *Proceedings of the 14th ACM International Conference on Future Energy Systems*, e-Energy '23, page 109–113, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400700323. doi: 10.1145/3575813.3576881. URL https://doi.org/10.1145/3575813.3576881.

BC Hydro. Just and reasonable: Inclining block rates. https://justandreasonable.com/knowledge-base/inclining-block-rate/, 2024. Accessed: 2024-07-12.

David Blum, Zhe Wang, Chris Weyandt, Donghun Kim, Michael Wetter, Tianzhen Hong, and Mary Ann Piette. Field demonstration and implementation analysis of model predictive control in an office hvac system. *Applied energy*, 318:119104, 2022.

David H. Blum and Michael Wetter. *MPCPy: An open-source software platform for model predictive control in buildings*, volume 15. IBPSA, 2017.

Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

Jan vom Brocke, Alexander Simons, Bjoern Niehaves, Bjorn Niehaves, Kai Reimer, Ralf Plattfaut, and Anne Cleven. Reconstructing the giant: On the importance of rigour in documenting the literature search process. 2009.

Bundesverband Wärmepumpe e.V. Wie funktioniert die Wärmepumpe? https://www.waermepumpe.de/waermepumpe/funktion-waermequellen/, 2019. Accessed: 2025-01-31.

Jesse Burkhardt, Kenneth T Gillingham, and Praveen K Kopalle. Field experimental evidence on the effect of pricing on residential electricity conservation. *Management Science*, 69(12):7784–7798, 2023.

Wenqi Cai, Arash Bahari Kordabad, and Sébastien Gros. Energy management in residential microgrid using model predictive control-based reinforcement learning and shapley value. *Engineering Applications of Artificial Intelligence*, 119:105793, 2023. ISSN 0952-1976. doi: https://doi.org/10.1016/j.engappai.2022.105793. URL https://www.sciencedirect.com/science/article/pii/S0952197622007837.

Munseok Chang, Sungwoo Bae, Gilhwan Cha, and Jaehyun Yoo. Aggregated electric vehicle fast-charging power demand analysis and forecast based on lstm neural network. *Sustainability*, 13 (24):13783, 2021.

Hugh Chen, Ian C Covert, Scott M Lundberg, and Su-In Lee. Algorithms to estimate shapley value feature attributions. *Nature Machine Intelligence*, 5(6):590–601, 2023.

Tao Chen, Qiushi Cui, Ciwei Gao, Qinran Hu, Kexing Lai, Jianlin Yang, Ran Lyu, Hao Zhang, and Jinyuan Zhang. Optimal demand response strategy of commercial building-based virtual power plant using reinforcement learning. *IET Generation, Transmission & Distribution*, 15 (16):2309–2318, 2021a.

Wen Chen, Jing Qiu, and Qingmian Chai. Customized critical peak rebate pricing mechanism for virtual power plants. *IEEE Transactions on Sustainable Energy*, 12(4):2169–2183, 2021b.

Yongbao Chen, Mingyue Guo, Zhisen Chen, Zhe Chen, and Ying Ji. Physical energy and data-driven models in building energy prediction: A review. *Energy Reports*, 8:2656–2671, 2022.

Andrew D Chiasson. Advances in modeling of ground-source heat pump systems. *Ph.D. Thesis*, 1999.

John Clauß and Laurent Georges. Model complexity of heat pump systems to investigate the building energy flexibility and guidelines for model implementation. *Applied Energy*, 255: 113847, 2019.

T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein. *Introduction To Algorithms*. Mit Electrical Engineering and Computer Science. MIT Press, 2001. ISBN 9780262032933. URL `https://books.google.de/books?id=NLngYyWFl_YC`.

Sho Cremers, Valentin Robu, Daan Hofman, Titus Naber, Kawin Zheng, and Sonam Norbu. Efficient methods for approximating the shapley value for asset sharing in energy communities. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, e-Energy '22, page 320–324, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450393973. doi: 10.1145/3538637.3538861. URL `https://doi.org/10.1145/3538637.3538861`.

Hongbo Cui, Wei Xia, Shanlin Yang, and Xiaojia Wang. Real-time emergency demand response strategy for optimal load dispatch of heat and power micro-grids. *International Journal of Electrical Power & Energy Systems*, 121:106127, 2020.

Darwish Darwazeh, Jean Duquette, Burak Gunay, Ian Wilton, and Scott Shillinglaw. Review of peak load management strategies in commercial buildings. *Sustainable Cities and Society*, 77: 103493, 2022.

Himadry Shekhar Das, Mohammad Mominur Rahman, Shuhui Li, and Chee Wei Tan. Electric vehicles standards, charging infrastructure, and impact on grid integration: A technological review. *Renewable and Sustainable Energy Reviews*, 120:109618, 2020.

Sima Davarzani, Ioana Pisica, Gareth A Taylor, and Kevin J Munisami. Residential demand response strategies and applications in active distribution network management. *Renewable and sustainable energy reviews*, 138:110567, 2021.

Moacyr AG De Brito, Leonardo P Sampaio, G Luigi, Guilherme A e Melo, and Carlos A Canesin. Comparative analysis of mppt techniques for pv applications. In *2011 International Conference on Clean Electrical Power (ICCEP)*, pages 99–104. IEEE, 2011.

Roel De Coninck and Lieve Helsen. Practical implementation and evaluation of model predictive control for an office building in brussels. *Energy and Buildings*, 111:290–298, 2016.

DECC. Implementing Electricity Market Reform (EMR). `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/324176/Implementing_Electricity_Market_Reform.pdf`, 2014. Accessed: 2024-08-01.

Anthony S Deese and Julian Daum. Application of zigbee-based internet of things technology to demand response in smart grids. *IFAC-PapersOnLine*, 51(28):43–48, 2018.

Ruilong Deng, Zaiyue Yang, Mo-Yuen Chow, and Jiming Chen. A survey on demand response in smart grids: Mathematical models and approaches. *IEEE Transactions on Industrial Informatics*, 11(3):570–582, 2015.

Department for Transport. Electric chargepoint analysis 2017: Domestics, 2018. URL `https://www.data.gov.uk/dataset/`.

Deutsche Energie-Agentur. Studie: Roadmap demand side management. `https://www.dena.de/infocenter/roadmap-demand-side-management`, 2016. Accessed: 2024-11-01.

Chaman Lal Dewangan, Vineeth Vijayan, Devesh Shukla, S Chakrabarti, SN Singh, Ankush Sharma, and Md Alamgir Hossain. An improved decentralized scheme for incentive-based demand response from residential customers. *Energy*, 284:128568, 2023.

Yihong Ding, Qinliang Tan, Zijing Shan, Jian Han, and Yimei Zhang. A two-stage dispatching optimization strategy for hybrid renewable energy system with low-carbon and sustainability in ancillary service market. *Renewable Energy*, 207:647–659, 2023.

C. Dornellas, E. Neves, L. Barroso, J. Mello, J.W.M. Lima, J. Dutra, A. Pimenta, and H. Sousa. Evaluation of demand side management mechanisms and opportunities for their development in the brazilian power industry. *CIGRE Session 46*, 2016-August, 2016.

Stephanie Alencar Braga dos Santos, José Marques Soares, Giovanni Cordeiro Barroso, and Bruno de Athayde Prata. Demand response application in industrial scenarios: A systematic mapping of practical implementation. *Expert Systems with Applications*, 215:119393, 2023.

Géremi Gilson Dranka, Paula Ferreira, and A Ismael F Vaz. Integrating supply and demand-side management in renewable-based energy systems. *Energy*, 232:120978, 2021.

Ján Drgoňa, Damien Picard, and Lieve Helsen. Cloud-based implementation of white-box model predictive control for a geotabs office building: A field test demonstration. *Journal of Process Control*, 88:63–77, 2020.

EIA. Consumption & Efficiency - U.S. Energy Information Administration (EIA). `https://www.eia.gov/tools/faqs/faq.php?id=86&t=1`, 2023. Accessed: 2024-08-16.

Rasha Elazab, Omar Saif, Amr MA Amin Metwally, and Mohamed Daowd. New smart home energy management systems based on inclining block-rate pricing scheme. *Clean Energy*, 6 (3):503–511, 2022.

Tlamelo Emmanuel, Thabiso Maupong, Dimane Mpoeleng, Thabo Semong, Banyatsang Mphago, and Oteng Tabona. A survey on missing data in machine learning. *Journal of Big data*, 8:1–37, 2021.

Dongming Fan, Yi Ren, Qiang Feng, Yiliu Liu, Zili Wang, and Jing Lin. Restoration of smart grids: Current status, challenges, and opportunities. *Renewable and Sustainable Energy Reviews*, 143:110909, 2021.

Chen Fang, Xiaojin Zhao, Qin Xu, Donghan Feng, Haojing Wang, and Yun Zhou. Aggregator-based demand response mechanism for electric vehicles participating in peak regulation in valley time of receiving-end power grid. *Global Energy Interconnection*, 3(5):453–463, 2020.

Christian Finck, Rongling Li, and Wim Zeiler. Optimal control of demand flexibility under real-time pricing for heating systems in buildings: A real-life demonstration. *Applied energy*, 263:114671, 2020.

Moritz Frahm, Thomas Dengiz, Philipp Zwickel, Heiko Maaß, Jörg Matthes, and Veit Hagenmeyer. Occupant-oriented demand response with multi-zone thermal building control. *Applied Energy*, 347:121454, 2023.

Alejandro Fraija, Kodjo Agbossou, Nilson Henao, Sousso Kelouwani, Michaël Fournier, and Sayed Saeed Hosseini. A discount-based time-of-use electricity pricing strategy for demand response with minimum information using reinforcement learning. *IEEE Access*, 10:54018–54028, 2022.

Svenne Freund and Gerhard Schmitz. Implementation of model predictive control in a large-sized, low-energy office building. *Building and environment*, 197:107830, 2021.

Hessam Golmohamadi. Demand-side management in industrial sector: A review of heavy industries. *Renewable and Sustainable Energy Reviews*, 156:111963, 2022.

Lili Gong, Wu Cao, Kangli Liu, Yue Yu, and Jianfeng Zhao. Demand responsive charging strategy of electric vehicles to mitigate the volatility of renewable energy sources. *Renewable Energy*, 156:665–676, 2020.

M González-Torres, Luis Pérez-Lombard, Juan F Coronel, Ismael R Maestre, and Da Yan. A review on buildings energy information: Trends, end-uses, fuels and drivers. *Energy Reports*, 8:626–637, 2022.

Hermann Heinrich Gossen. *Entwickelung der Gesetze des menschlichen Verkehrs, und der daraus fliessenden Regeln für menschliches Handeln.* Friedrich Vieweg und Sohn, 1854. URL `https://books.google.de/books?id=BzFGAAAAYAAJ`.

Grafana. Grafana: The open and composable observability platform, 2014. URL `https://grafana.com/`.

Simon Grund, Oliver Lüdtke, and Alexander Robitzsch. Multiple imputation of missing data in multilevel models with the r package mdmb: A flexible sequential modeling approach. *Behavior Research Methods*, 53(6):2631–2649, 2021.

Sandip Kumar Gupta, T Ghose, and Kalyan Chatterjee. Coordinated control of incentive-based demand response program and bess for frequency regulation in low inertia isolated grid. *Electric Power Systems Research*, 209:108037, 2022.

Francisco Gutierrez-Garcia, Angel Arcos-Vargas, and Antonio Gomez-Exposito. Robustness of electricity systems with nearly 100% share of renewables: A worst-case study. *Renewable and Sustainable Energy Reviews*, 155:111932, 2022.

Shahrzad Hadayeghparast and Hadis Karimipour. Comprehensive modeling of demand response programs. *Demand Response Application in Smart Grids: Concepts and Planning Issues-Volume 1*, pages 21–78, 2020.

Veit Hagenmeyer, Hüseyin Kemal Çakmak, Clemens Düpmeier, Timm Faulwasser, Jörg Isele, Hubert B Keller, Peter Kohlhepp, Uwe Kühnapfel, Uwe Stucky, Simon Waczowicz, et al. Information and communication technology in energy lab 2.0: Smart energies system simulation and control center with an open-street-map-based power flow simulation example. *Energy Technology*, 4(1):145–162, 2016.

Boris Hanin and David Rolnick. Complexity of linear regions in deep networks. In *International Conference on Machine Learning*, pages 2596–2604. PMLR, 2019.

Jason Harold, Valentin Bertsch, and Harrison Fell. Preferences for curtailable electricity contracts: Can curtailment benefit consumers and the electricity system? *Energy Economics*, 102:105454, 2021.

Matthias Hertel, Maximilian Beichter, Benedikt Heidrich, Oliver Neumann, Benjamin Schäfer, Ralf Mikut, and Veit Hagenmeyer. Transformer training strategies for forecasting multiple load time series. *Energy Informatics*, 6(Suppl 1):20, 2023.

Thabo G Hlalele, Jiangfeng Zhang, Raj M Naidoo, and Ramesh C Bansal. Multi-objective economic dispatch with residential demand response programme under renewable obligation. *Energy*, 218:119473, 2021.

Matthias Hofmann and Karen Byskov Lindberg. Evidence of households' demand flexibility in response to variable hourly electricity prices–results from a comprehensive field experiment in norway. *Energy Policy*, 184:113821, 2024.

Hardi Hõimoja, Alfred Rufer, Grzegorz Dziechciaruk, and Andrea Vezzini. An ultrafast ev charging station demonstrator. In *International Symposium on Power Electronics Power Electronics, Electrical Drives, Automation and Motion*, pages 1390–1395. IEEE, 2012.

John Horowitz, John List, and Kenneth E McConnell. A test of diminishing marginal value. *Economica*, 74(296):650–663, 2007.

Xia Hu, Lingyang Chu, Jian Pei, Weiqing Liu, and Jiang Bian. Model complexity of deep learning: A survey. *Knowledge and Information Systems*, 63(10):2585–2619, 2021.

IEA. Demand Response Overview. `https://www.iea.org/energy-system/energy-efficiency-and-demand/demand-response`, 2023. Accessed: 2024-07-12.

Semich Impram, Secil Varbak Nese, and Bülent Oral. Challenges of renewable energy penetration on power system flexibility: A survey. *Energy Strategy Reviews*, 31:100539, 2020. ISSN 2211-467X.

Grant D Jacobsen and James I Stewart. How do consumers respond to price complexity? experimental evidence from the power sector. *Journal of Environmental Economics and Management*, 116:102716, 2022.

Heesun Jang, Seongman Moon, and Jihyo Kim. Effects of time-of-use pricing for residential customers and wholesale market consequences in south korea. *Energy Economics*, 134:107557, 2024. ISSN 0140-9883. doi: https://doi.org/10.1016/j.eneco.2024.107557.

José M Jerez, Ignacio Molina, Pedro J García-Laencina, Emilio Alba, Nuria Ribelles, Miguel Martín, and Leonardo Franco. Missing data imputation using statistical and machine learning methods in a real breast cancer problem. *Artificial intelligence in medicine*, 50(2):105–115, 2010.

Quanyuan Jiang and Haijiao Wang. Two-time-scale coordination control for a battery energy storage system to mitigate wind power fluctuations. *IEEE Transactions on Energy Conversion*, 28(1):52–61, 2012.

Rui Jing, Mei Na Xie, Feng Xiang Wang, and Long Xiang Chen. Fair p2p energy trading between residential and commercial multi-energy systems enabling integrated demand-side management. *Applied Energy*, 262:114551, 2020.

Rudolf Kapeller, Jed J Cohen, Andrea Kollmann, and Johannes Reichl. Incentivizing residential electricity consumers to increase demand during periods of high local solar generation. *Energy Economics*, 127:107028, 2023.

Shahidul Islam Khan and Abu Sayed Md Latiful Hoque. Sice: an improved missing data imputation technique. *Journal of big Data*, 7(1):37, 2020.

Somayeh Siahchehre Kholerdi and Ali Ghasemi-Marzbali. Interactive time-of-use demand response for industrial electricity customers: A case study. *Utilities Policy*, 70:101192, 2021.

Hyung Joon Kim, Hyun Joon Kang, and Mun Kyeom Kim. Data-driven bidding strategy for der aggregator based on gated recurrent unit–enhanced learning particle swarm optimization. *IEEE Access*, 9:66420–66435, 2021.

KIT. Energy lab 2.0, 2014. URL `https://www.elab2.kit.edu/`.

Peter Kohlhepp, Hassan Harb, Henryk Wolisz, Simon Waczowicz, Dirk Müller, and Veit Hagenmeyer. Large-scale grid integration of residential thermal energy storages as demand-side flexibility resource: A review of international field studies. *Renewable and Sustainable Energy Reviews*, 101:527–547, 2019.

Jovana Kovačević, Chang Li, Gina Brecher, Kevin Förderer, Hüseyin K Çakmak, and Veit Hagenmeyer. The impact of model complexity reduction on system dynamics: A modelica case study of a residential heating system. In *2025 IEEE Conference on Technologies for Sustainability (SusTech)*, pages 1–8. IEEE, 2025.

Harold W Kuhn and Albert William Tucker. *Contributions to the Theory of Games*. Number 28. Princeton University Press, 1953.

I Elizabeth Kumar, Suresh Venkatasubramanian, Carlos Scheidegger, and Sorelle Friedler. Problems with shapley-value-based explanations as feature importance measures. In *International conference on machine learning*, pages 5491–5500. PMLR, 2020.

N Lashmar, B Wade, L Molyneaux, and P Ashworth. Motivations, barriers, and enablers for demand response programs: A commercial and industrial consumer perspective. *Energy Research & Social Science*, 90:102667, 2022.

Richard Layard, Guy Mayraz, and Stephen Nickell. The marginal utility of income. *Journal of Public Economics*, 92(8-9):1846–1857, 2008.

Yu-Ching Lee, Hsin-Wei Hsu, Yen-Chang Chang, Ying-Lien Chen, and Ming-Chuan Chiu. Constructing a (fair) load-reduction emergency demand response program: A fee-and-rebate pricing mechanism. *International Journal of Electrical Power & Energy Systems*, 147:108831, 2023.

Chang Li, Kevin Förderer, Tobias Moser, Luigi Spatafora, and Veit Hagenmeyer. Gossen's first law in the modeling for demand side management: A first heat pump case study. In *Energy Informatics Academy Conference*, pages 111–125. Springer, 2023. doi: https://doi.org/10.1007/978-3-031-48652-4_8.

Chang Li, Gina Brecher, Jovana Kovačević, Hüseyin K Çakmak, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. Gossen's first law in the modeling for demand side management: a thorough heat pump case study with deep learning based partial time series data generation. *Energy Informatics*, 7(47), 2024a.

Chang Li, Miao Zhang, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. A comparative analysis of machine learning algorithms for aggregated electric chargepoint load forecasting. In *Proceedings of the 9th International Conference on Sustainable and Renewable Energy Engineering*, volume 545. EDP Sciences, 2024b. doi: https://doi.org/10.1051/e3sconf/202454501004.

Chang Li, Felix Langner, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. A survey on the current state of demand response implementations: Models and approaches. *Energy Informatics Review*, 5(1), 2025.

JiaHang Li, ShuXia Guo, RuLin Ma, Jia He, XiangHui Zhang, DongSheng Rui, YuSong Ding, Yu Li, LeYao Jian, Jing Cheng, et al. Comparison of the effects of imputation methods for missing data in predictive modelling of cohort study datasets. *BMC Medical Research Methodology*, 24(1):41, 2024c.

Hossein Lotfi and Reza Ghazi. Optimal participation of demand response aggregators in reconfigurable distribution system considering photovoltaic and storage units. *Journal of Ambient Intelligence and Humanized Computing*, 12:2233–2255, 2021.

Peter D. Lund, Juuso Lindgren, Jani Mikkola, and Jyri Salpakari. Review of energy system flexibility measures to enable high levels of variable renewable electricity. *Renewable and Sustainable Energy Reviews*, 45:785–807, 2015. ISSN 1364-0321.

Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

Scott M Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence*, 2(1):56–67, 2020.

Xuan Luo and Jian Pei. Applications and computation of the shapley value in databases and machine learning. In *Companion of the 2024 International Conference on Management of Data*, pages 630–635, 2024.

Jun Ma, Jack CP Cheng, Feifeng Jiang, Weiwei Chen, Mingzhu Wang, and Chong Zhai. A bi-directional missing data imputation scheme based on lstm and transfer learning for building energy data. *Energy and Buildings*, 216:109941, 2020a.

Shuaiyin Ma, Yingfeng Zhang, Yang Liu, Haidong Yang, Jingxiang Lv, and Shan Ren. Data-driven sustainable intelligent manufacturing based on demand response for energy-intensive industries. *Journal of Cleaner Production*, 274:123155, 2020b.

Hatef Madani, Joachim Claesson, and Per Lundqvist. Capacity control in ground source heat pump systems: Part i: modeling and simulation. *International Journal of Refrigeration*, 34(6): 1338–1347, 2011.

Charlotte A Malmborg, Alyssa M Willson, L_M Bradley, Meghan A Beatty, David H Klinges, Gerbrand Koren, Abigail SL Lewis, Kayode Oshinubi, and Whitney M Woelmer. Defining model complexity: An ecological perspective. *Meteorological Applications*, 31(3):e2202, 2024.

JJ McArthur. A building information management (bim) framework and supporting case study for existing building operations, maintenance and sustainability. *Procedia engineering*, 118: 1104–1111, 2015.

Jan Meyer. Aktuelle messwerte der sole-wasser wärmepumpen anlag, 2024. URL https://www.effiziente-waermepumpe.ch/messdaten/index.php.

Saeed Mohajeryami, Iman N Moghaddam, Milad Doostan, Behdad Vatani, and Peter Schwarz. A novel economic model for price-based demand response. *Electric Power Systems Research*, 135:1–9, 2016.

Robert W Morton, Kevin T Murphy, Sean R McKellar, Brad J Schoenfeld, Menno Henselmans, Eric Helms, Alan A Aragon, Michaela C Devries, Laura Banfield, James W Krieger, et al. A systematic review, meta-analysis and meta-regression of the effect of protein supplementation on resistance training-induced gains in muscle mass and strength in healthy adults. *British journal of sports medicine*, 52(6):376–384, 2018.

NIBE. Installateurhandbuch nibe f1155, 2023. URL `https://www.nibe.eu/assets/documents/24306/331341-5.pdf`.

Pavlos Nikolaidis and Andreas Poullikkas. Sustainable services to enhance flexibility in the upcoming smart grids. *Sustaining resources for tomorrow*, pages 245–274, 2020.

Gilles Notton and Cyril Voyant. Forecasting of intermittent solar energy resource. In *Advances in renewable energies and power technologies*, pages 77–114. Elsevier, 2018.

OpenADR. Case Studies & White Papers. `https://www.openadr.org/case-studies`, 2011. Accessed: 2024-09-18.

James Owens, Tim Schittekatte, and Emre Gençer. Commercial electric vehicle fleets in us ancillary services markets: A stochastic analysis to inform utility rate design. *Utilities Policy*, 90:101790, 2024.

Emre Ozkop. A survey on direct load control technologies in the smart grid. *IEEE Access*, 2024.

Peter Palensky and Dietmar Dietrich. Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE transactions on industrial informatics*, 7(3):381–388, 2011.

Subhasis Panda, Sarthak Mohanty, Pravat Kumar Rout, Binod Kumar Sahu, Mohit Bajaj, Hossam M Zawbaa, and Salah Kamel. Residential demand side management model, optimization and future perspective: A review. *Energy Reports*, 8:3727–3766, 2022.

Bryony Parrish, Phil Heptonstall, Rob Gross, and Benjamin K Sovacool. A systematic review of motivations, enablers and barriers for consumer engagement with residential demand response. *Energy Policy*, 138:111221, 2020.

Thibault Q Péan, Jaume Salom, and Ramon Costa-Castelló. Review of control strategies for improving the energy flexibility provided by heat pump systems in buildings. *Journal of Process Control*, 74:35–49, 2019.

Frederic Pelletier and Ahmad Faruqui. Does dynamic pricing work in a winter-peaking climate? a case study of hydro quebec. *The Electricity Journal*, 35(2):107080, 2022.

Grace CY Peng, Mark Alber, Adrian Buganza Tepole, William R Cannon, Suvranu De, Savador Dura-Bernal, Krishna Garikipati, George Karniadakis, William W Lytton, Paris Perdikaris, et al. Multiscale modeling meets machine learning: What can we learn? *Archives of Computational Methods in Engineering*, 28:1017–1037, 2021.

Dario Peralta, Claudio A Cañizares, and Kankar Bhattacharya. Ground source heat pump modeling, operation, and participation in electricity markets. *IEEE Transactions on Smart Grid*, 13 (2):1126–1138, 2021.

Mette K Petersen, Kristian Edlund, Lars Henrik Hansen, Jan Bendtsen, and Jakob Stoustrup. A taxonomy for modeling flexibility and a computationally efficient algorithm for dispatch in smart grids. In *2013 American control conference*, pages 1150–1156. IEEE, 2013.

Damien Picard, Maarten Sourbron, Filip Jorissen, Ji Cigler, Lukás Ferkl, Lieve Helsen, et al. Comparison of model predictive control performance using grey-box and white box controller models. In *4th International High Performance Buildings Conference at Purdue*, 2016.

Hajir Pourbabak, Jingwei Luo, Tao Chen, and Wencong Su. A novel consensus-based distributed algorithm for economic dispatch based on local estimation of power mismatch. *IEEE Transactions on Smart Grid*, 9(6):5930–5942, 2017.

Kaiying Qiu, Junlu Yang, Zhi Gao, and Fusuo Xu. A review of modelica language in building and energy: Development, applications, and future prospect. *Energy and Buildings*, 308:113998, 2024.

Abidur Rahman, Tareq Aziz, Shohana Rahman Deeba, et al. A time of use tariff scheme for demand side management of residential energy consumers in bangladesh. *Energy Reports*, 7: 3189–3198, 2021.

Matteo Ranaboldo, Mònica Aragüés-Peñalba, Emrah Arica, A Bade, Eduard Bullich-Massagué, Alessandro Burgio, Chiara Caccamo, A Caprara, Domenico Cimmino, Bruno Domenech, et al. A comprehensive overview of industrial demand response status in europe. *Renewable and Sustainable Energy Reviews*, 203:114797, 2024.

Stewart Robinson. Exploring the relationship between simulation model accuracy and complexity. *Journal of the Operational Research Society*, 74(9):1992–2011, 2023.

Félix Ruiz-Calvo, M De Rosa, José Acuña, JM Corberán, and C Montagud. Experimental validation of a short-term borehole-to-ground (b2g) dynamic model. *Applied Energy*, 140: 210–223, 2015.

SA Saleh, R Ahshan, M Haj-Ahmed, Julian L Cardenas-Barrera, J Meng, and E Castillo-Guerra. Energy not-served-based method for assessing smart grid functions in residential loads. *IEEE Transactions on Industry Applications*, 58(2):1720–1729, 2022.

Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, Marco Ratto, et al. *Sensitivity analysis in practice: a guide to assessing scientific models*, volume 1. Wiley Online Library, 2004.

Ioan Sarbu and Calin Sebarchievici. General review of ground-source heat pump systems for heating and cooling of buildings. *Energy and buildings*, 70:441–454, 2014.

Shivam Saxena, Hany EZ Farag, Leigh St Hilaire, and Aidan Brookson. A techno-social approach to unlocking vehicle to everything (v2x) integration: A real-world demonstration. *IEEE Access*, 11:17085–17095, 2023.

Leandra Scharnhorst, Daniel Sloot, Nico Lehmann, Armin Ardone, and Wolf Fichtner. Barriers to demand response in the commercial and industrial sectors–an empirical investigation. *Renewable and Sustainable Energy Reviews*, 190:114067, 2024.

Michiel Schotten, Wim JN Meester, Susanne Steiginga, Cameron A Ross, et al. A brief history of scopus: The world's largest abstract and citation database of scientific literature. In *Research analytics*, pages 31–58. Auerbach Publications, 2017.

Ahmed Shaban, Mohamed Salhen, Mohamed A Shalaby, and Tamer F Abdelmaguid. Optimal household appliances scheduling for smart energy management considering inclining block rate tariff and net-metering system. *Computers & Industrial Engineering*, 190:110073, 2024.

Lloyd S Shapley. Notes on the n-person game—ii: The value of an n-person game. *RAND memorandum*, 1951.

Maryam H Shoreh, Pierluigi Siano, Miadreza Shafie-khah, Vincenzo Loia, and João PS Catalão. A survey of industrial applications of demand response. *Electric Power Systems Research*, 141: 31–49, 2016.

Tomasz Śliwa and Andrzej Gonet. Theoretical model of borehole heat exchanger. *Journal of Energy Resources Technology*, 2005.

Statistisches Bundesamt. Energy consumption of private households. `https://www.destatis.de/EN/Themes/Society-Environment/Environment/Environmental-Economic-Accounting/private-households/Tables/energy-consumption-households.html`, 2023. Accessed: 2025-09-10.

Stefan Strömer, Ksenia Poplavskaya, and Niclas Damsgaard. Model-based analysis of the effect of pricing rules and new entrants on the balancing market efficiency. case study of the swedish market. *Energy Reports*, 8:12628–12642, 2022.

The Joint Research Centre: EU Science Hub. Sensitivity analysis: Samo. https://joint-research-centre.ec.europa.eu/sensitivity-analysis-samo_en, 2025. Accessed: 2025-07-01.

John Meurig Thomas, Peter P Edwards, Peter J Dobson, and Gari P Owen. Decarbonising energy: The developing international activity in hydrogen technologies and fuel cells. *Journal of Energy Chemistry*, 51:405–415, 2020.

Simon Thorsteinsson, Alex Arash Sand Kalaee, Pierre Vogler-Finck, Henrik Lund Stærmose, Ivan Katic, and Jan Dimon Bendtsen. Long-term experimental study of price responsive predictive control in a real occupied single-family house with heat pump. *Applied Energy*, 347:121398, 2023.

Yan Tian, Kaili Zhang, Jianyuan Li, Xianxuan Lin, and Bailin Yang. Lstm-based traffic flow prediction with missing data. *Neurocomputing*, 318:297–305, 2018.

Valentina Tomat, Marika Vellei, Alfonso P Ramallo-González, Aurora González-Vidal, Jérôme Le Dréau, and Antonio Skarmeta-Gómez. Understanding patterns of thermostat overrides after demand response events. *Energy and Buildings*, 271:112312, 2022.

Konstantin Turitsyn, Scott Backhaus, Maxim Ananyev, and Michael Chertkov. Smart finite state devices: A modeling framework for demand response technologies. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 7–14. IEEE, 2011.

Umwelt Bundesamt. Climate emissions fall by 10.1 per cent in 2023 – biggest decline since 1990. https://www.umweltbundesamt.de/en/press/pressinformation/climate-emissions-fall-101-per-cent-in-2023-biggest, 2024. Accessed: 2024-11-01.

Ivo AM Varenhorst, Johann Hurink, and Marco ET Gerards. Quantifying device flexibility with shapley values in demand side management. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems*, pages 147–157, 2024.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Dan Wang, Wanfu Zheng, Siqi Li, Dongdong Li, Shen Li, Bin Li, and Zhe Wang. Field demonstration of priority stack-based controls in an office building for demand response. *Journal of Building Engineering*, page 109715, 2024.

Fei Wang, Xinxin Ge, Peng Yang, Kangping Li, Zengqiang Mi, Pierluigi Siano, and Neven Duić. Day-ahead optimal bidding and scheduling strategies for der aggregator considering responsive uncertainty under real-time pricing. *Energy*, 213:118765, 2020.

Weijun Wang, Yicen Han, Meng Wang, and Yan He. Research on fair residential critical peak price: Based on a price penalty mechanism for high-electricity consumers. *Applied Energy*, 351:121892, 2023.

Moritz Weber, Marian Turowski, Hüseyin K Çakmak, Ralf Mikut, Uwe Kühnapfel, and Veit Hagenmeyer. Data-driven copy-paste imputation for energy time series. *IEEE Transactions on Smart Grid*, 12(6):5409–5419, 2021.

Moritz Weber, Hüseyin K Çakmak, Uwe Kühnapfel, and Veit Hagenmeyer. A novel method for topology preserving static network reduction. In *2023 IEEE 11th International Conference on Smart Energy Grid Engineering (SEGE)*, pages 58–63. IEEE, 2023.

Qingsong Wen, Tian Zhou, Chaoli Zhang, Weiqi Chen, Ziqing Ma, Junchi Yan, and Liang Sun. Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*, 2022.

Eyal Winter. The shapley value. *Handbook of game theory with economic applications*, 3:2025–2054, 2002.

Chengyan Xiong, Qinglong Meng, Huilong Luo, Yu Lei, Jiao Liu, Xiuying Yan, et al. A demand response method for an active thermal energy storage air-conditioning system using improved transactive control: On-site experiments. *Applied Energy*, 339:120935, 2023.

Xing Yan, Yusuf Ozturk, Zechun Hu, and Yonghua Song. A review on price-driven residential demand response. *Renewable and Sustainable Energy Reviews*, 96:411–419, 2018.

Kun Zhang, Anand Prakash, Lazlo Paul, David Blum, Peter Alstone, James Zoellick, Richard Brown, and Marco Pritoni. Model predictive control for demand flexibility: Real-world operation of a commercial building with photovoltaic and battery systems. *Advances in Applied Energy*, 7:100099, 2022.

Peng Zhang, Kejun Qian, Chengke Zhou, Brian G Stewart, and Donald M Hepburn. A methodology for optimization of power systems demand due to electric vehicle charging load. *IEEE Transactions on Power Systems*, 27(3):1628–1636, 2012.

Zhongheng Zhang. Missing data imputation: focusing on single imputation. *Annals of translational medicine*, 4(1):9, 2016.

Nan Zheng, Xingqi Ding, Yufei Zhou, Hanfei Zhang, Liqiang Duan, and Maotong Zhang. Dynamic modeling and response characteristics of a solar-driven polygeneration system coupled with

hydrogen and thermal energy storage. *Energy Conversion and Management*, 293:117496, 2023.

Shunlin Zheng, Yi Sun, Bin Li, Bing Qi, Kun Shi, Yuanfei Li, and Xiazhe Tu. Incentive-based integrated demand response for multiple energy carriers considering behavioral coupling effect of consumers. *IEEE Transactions on Smart Grid*, 11(4):3231–3245, 2020.

Juncheng Zhu, Zhile Yang, Yuanjun Guo, Jiankang Zhang, and Huikun Yang. Short-term load forecasting for electric vehicle charging stations based on deep learning approaches. *Applied sciences*, 9(9):1723, 2019.