Decision support tools for smart Home Energy Management Systems (HEMSs)

Vinasco, Julian Lemos

Publication date:
2022

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

Ph.D. Thesis
Doctor of Philosophy

DTU Compute
Department of Applied Mathematics and Computer Science

Decision support tools for smart Home Energy Management Systems (HEMSs)

Industrial Ph.D. project

Julián Lemos-Vinasco

Kongens Lyngby 2022

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

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

Summary

In Europe, significant efforts have been made to transition the energy sector from fossil-fuel-based to a model based on renewable energy sources. The energy transition comes with multiple challenges, with smart grids initiatives being created in response.

Recent advancements made in information and communication technologies, electricity storage systems, renewable electricity generation and smart appliances have been made at a residential level as part of smart grid initiatives. This growing trend provides the technical foundation and infrastructure for houses with smart Home Energy Management Systems (HEMSs). Smart HEMSs can be defined as systems providing energy management services to efficiently monitor and manage electricity generation, storage, and consumption in households. The main objective of this research project is to study and develop a set of decision support tools for the deployment of HEMSs in Denmark that can bring the maximum value to home owners. This includes tools for the prediction of renewable energy sources generation and users' consumption, as well as control strategies for the different household components. Moreover, this project intends to contribute with the building blocks needed for functional HEMSs that support the energy transition in Europe.

Preface

This Ph.D. thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark (DTU) in partial fulfillment of the requirements for acquiring a Ph.D. degree. The thesis consist of a summary report and three academic articles dealing the research conducted over the period of January 2019 to December 2021.

The Ph.D study was concentrated on the building blocks needed for successful implementation of Home Energy Management Systems (HEMSs). This includes models for forecasting of individual households' electrical loads, and optimization-based control strategies. The project was conducted together with a company and hence there is a strong focus on direct application of the developed algorithms, as well as real data cases.

Kongens Lyngby, March 31, 2022

Julián Lemos-Vinasco

Acknowledgements

I would like to express my gratitude to my supervisors Peder Bacher, Jan Kloppenborg Møller, and Daniela Guericke for their earnest support, insightful advice, and pleasant attitude throughout the project.

Special thanks go to my supervisor, Anders Spur Hansen, and the rest of my colleagues at Watts A/S for making the project a reality and for their constant support.

Particular thanks go to Ph.D. S. Ali Pourmousavi in the School of Electrical and Electronic Engineering at the University of Adelaide for the collaboration during the project.

I would like to especially thank Michaela Ziskova for her emotional support and affection. Special thanks also goes to Daniel Barreneche-Giraldo for his caring friendship. I am sure that without your support my whole experience throughout the project would have been different. Finally, I would like to thank my family for always encouraging my academic endeavors and everyone that in one way or another contributed to my Ph.D. experience.

This work was partially funded by Innovation Fund Denmark through the project identified with case No. 8053-00156B.

List of acronyms

HEMS home energy management system

EV electric vehicle

RES renewable energy source

ESS energy storage system

NWP numerical weather prediction

IoT internet of things

TSO transmission system operator

EL electrical load

NWP numerical weather prediction

PV photovoltaic

LF load forecasting

PLF probabilistic load forecasting

TS time series

PDF probability density function

CDF cumulative density function

AR auto-regressive

RLS recursive least squares

GHI global horizontal irradiation

PICP prediction interval coverage probability

PINAW prediction interval normalized average

CRPS continuous ranked probability score

VarS variogram score

MILP mixed-integer linear program

SoC state of charge

AC alternate current

DC direct current

SDE stochastic differential equation

ANN artificial neural network

LSTM-ANN long short-term memory artificial neural network

QRNN quantile regression neural network

SHT smart home technologies

DA day-ahead

LINMAP linear programming technique for multidimensional analysis of preference

S-LINMAP stochastic linear programming technique for multidimensional analysis of preference

Contents

Summary i

Preface iii

Acknowledgements v

List of acronyms vii

Contents ix

I Summary report 1

1 Introduction 3
1.1 Context and motivation . . . 3
1.2 Thesis objectives . . . 5
1.3 Outline of the thesis . . . 6

2 Energy markets and home energy management systems 7
2.1 The role of HEMSs in energy systems . . . 7
2.2 Energy market overview . . . 8
2.3 HEMSs architecture . . . 9

3 Multivariate probabilistic forecasting 17
3.1 Motivation . . . 17
3.2 Modelling methodology . . . 18
3.3 Application on residential probabilistic load forecasting . . . 19
3.4 Extension to multivariate temporal forecasting . . . 25
3.5 Discussion . . . 26

4 Control strategies for home energy management systems 27
4.1 Motivation . . . 27
4.2 Modelling methodology . . . 29
4.3 Application on economic evaluation of HEMSs . . . 31
4.4 Application on bi-objective optimization of HEMSs . . . 35
4.5 Application on EV charging . . . 37
4.6 Discussion . . . 40

5 Conclusion 43
5.1 Contributions . . . 43
5.2 Perspectives and opportunities for further research . . . 44

II Publications 47

A Probabilistic load forecasting considering temporal correlation: Online models for the prediction of households' electrical load 49
A.1 Introduction . . . 50
A.2 Methods and models . . . 52
A.3 Data and simulation study . . . 57
A.4 Results . . . 60
A.5 Discussion . . . 66
A.6 Conclusion . . . 68

B Economic evaluation of stochastic home energy management systems in a realistic rolling horizon 69
B.1 Introduction . . . 70
B.2 Modeling and optimization of HEMSs . . . 73
B.3 Case study . . . 75
B.4 Simulation results . . . 78
B.5 Dicussion . . . 82
B.6 Conclusion . . . 83

C Stochastic bi-objective home energy management without articulation of user preferences 85
C.1 Introduction . . . 86
C.2 Related Work . . . 87
C.3 Methodology . . . 89
C.4 Case Study . . . 93
C.5 Conclusion . . . 99

Bibliography 101

Part I

Summary report

CHAPTER 1

Introduction

1.1 Context and motivation

In Europe, significant efforts have been made to transition the energy sector from fossil-fuel-based to a model based on renewable energy sources (RESs) in order to address climate change [38]. The energy transition comes with multiple challenges, one associated to the relation between intermittent power generation coming from RESs and a traditionally inflexible demand. Smart grid initiatives were born in response to these challenges, with smart grids being defined as "energy networks that can automatically monitor energy flows and adjust to changes in energy supply and demand accordingly" [19]. Moreover, when coupled with smart metering systems, smart grids reach consumers and suppliers by providing information on real-time consumption. In this regard, electricity market operators and policy makers are currently pushing for a wider adoption of real-time tariffs that reflect the true conditions of the power system and provide cost-savings to residential consumers [92].

Home Energy Management Systems (HEMSs) can be defined as a set of computer-aided tools used to monitor, control, and optimize the use of energy at a residential level. This type of systems rely on a combination of elements such as smart meters, sensing devices, communication hardware and protocols, smart appliances, and control and optimization techniques [12]. Moreover, additional systems and elements such as energy storage systems (ESSs), home RESs, and electric vehicles (EVs) could also be integrated. HEMSs are expected to be a major component of smart grids. Furthermore, high expectations have been placed on home energy management systems (HEMSs) by many industry and governmental stakeholders given the systems' potential to provide a dynamic combination of production, storage, and flexible demand [77, 42, 49], especially, if they are coupled with real-time tariffs and other signals. Therefore, studies on this topic have emerged from a variety of disciplines over the last decade, focusing on different components of the HEMSs. Although research on HEMSs has been an active research topic, there are doubts regarding the impact, as most studies are made using simulations with assumptions that might not hold in real applications and/or use laboratory conditions [112].

One of the main activities of HEMSs is the control and optimization of the different household components. In order to do so, HEMSs rely on decision support tools such as user interfaces, prediction methods (forecast), and control strategies. Typically, information gathered via user interfaces, sensing devices, and external sources is used by prediction models. These models have as main goal to provide information about the household future needs. At the same time, the output of the prediction models works as inputs for the control strategies. Here, different techniques such as optimization-based approaches, heuristic methods and/or rule-based methods, are used to define how to operate the different household components in a predefined time horizon. The control strategies can follow different objectives such as consumers' comfort, security parameters, and cost or/and CO_2 emissions minimization. When new information is available or when a new event triggers the system, the whole control loop starts again ensuring correct use of the household assets.

The smart meters roll out and the implementation of the Danish DataHub have opened the door to

service providers to consumers' data [35]. This has motivated companies to research and develop innovative approaches to differentiate themselves from the competition by providing additional insights and information to electricity consumers, which was the main idea behind the creation of Watts A/S. Watts was born as an energy assistant, where consumer were able to consult and visualize historic electricity consumption. As part of the company's efforts, in 2016 the company co-sponsored a industrial PhD project titled "Data-driven models for energy advising leading to behavioural changes in residences". The PhD project was centered around the creation of models able to forecast long-term consumer behavior and disaggregated feedback [74]. The research and posterior implementation of the project became the foundation of the Watts smart phone App, which currently has more than 190,000 users. Today, Watts is an electricity retailer which provides its users a portfolio of services such as consumption overview and alerts, budget (consumption forecast), hourly prices, CO_2 emission information, and smart appliances control [124].

Watts has a strategy to be one of the strongest players the in decentralize and digital space, providing premium services for consumers. These services are expected to span the vertical value chain including a combination of local energy hardware (photovoltaic (PV), heat pumps, batteries and EVs), smart inverters which can provide the necessary link between the physical/digital, and digital services which provide the consumer with an easy interface to controlling energy usage, trade, and engagement with the community. This strategy is clearly aligned with HEMSs and the development of smart grids and, therefore, it was how the project was born. Moreover, the company was interested to explore the following questions: how big is the potential for HEMSs?, how can the results be transformed into engaging user experiences and new business models?, what type of technological infrastructure is needed?. Thus, the project sits at a converging point in Watts strategic development, the energy transition, and the advancement of smart grid technologies.

1.2 Thesis objectives

The present thesis attempts to combine the current Watts infrastructure and data access with research on HEMSs in order to develop a setup suitable for a commercial application. The main objective of this thesis is to study and develop a set of decision support tools for the deployment of HEMSs in Denmark that can bring the maximum value to consumers. Specifically, the thesis has the following specific objectives:

- Understand the possible role that HEMSs can play in the Danish energy system and characterize the different data sources that HEMSs have potential access to.
- Define a HEMS architecture aligned with current market rules and conditions. The architecture should use technologies currently available to consumers.
- Develop models for the different HEMSs architecture elements.
- Integrate data access and the developed models into control strategies for the operation of HEMSs.
- Evaluate the economic and environmental impacts of HEMS architecture on residential users.

1.3 Outline of the thesis

This thesis is structured in two parts. Part I is a summary report outlining the main contributions and results of the thesis. In detail, Chapter 2 describes the expected role of HEMSs within the Danish energy system, defines the HEMS architecture considered in the other chapters, and presents the data accessible to the systems. Chapter 3 presents a multivariate probabilistic forecasting methodology developed in response to the HEMSs needs. It presents an application example together with possible extensions to the methodology and a critique of the work. Chapter 4 is dedicated to control strategies suitable for our HEMS architecture. It shows how to integrate the multivariate probabilistic forecast methodology and the data accessible to the system for the decision-making process. An evaluation of the control strategies considering cost and CO_2 emissions involving several users is presented as application examples. Finally, conclusions and future perspectives are presented in Chapter 5.

Part II consists of the publications that contributed to this thesis. Three journal articles are included. Paper A presents a methodology for multivariate probabilistic forecasting, with an application to the prediction of households' electrical load. It is a journal article published in *Applied Energy*. In Paper B, an economic evaluation of stochastic home energy management systems in a realistic rolling horizon is presented. This is a journal article currently under review in *Applied Energy*. Finally, Paper C presents a stochastic bi-objective home energy management model which considers electricity costs and CO_2 emissions in the optimization process. The manuscript is being prepared as a journal article and it is currently under internal review. The paper will be submitted to *IEEE transactions on smart grids*.

CHAPTER 2

Energy markets and home energy management systems

In this chapter, a brief argumentation of the role of HEMSs in energy systems is presented in Section 2.1. Given that our HEMS setup operates in a Danish context, an overview of the energy market in Denmark is presented in Section 2.2. Finally, the different elements of a HEMS architecture tailored to the Danish market are presented in Section 2.3.

2.1 The role of HEMSs in energy systems

As one of the major smart grid technologies, HEMSs are expected to play a key role managing energy consumption at the residential level by reacting to real-time prices and/or CO_2-based signals. This coincides with efforts made by electricity market operators and policy makers to push for a wider adoption or real-time tariffs for residential consumers that reflect the true condition of the power system and provide cost-savings for consumers in Denmark [33]. High expectations have been placed on HEMSs by many industry stakeholders given their potential to provide a dynamic combination of generation, storage, and flexible demand. In this reward, current data sources and available hardware allow the possibility to explore the potential that automatic HEMSs could bring the consumers and the energy system.

Several studies exploring the possible role of HEMSs in the power system have been made as it is demonstrated by review articles on the subject [77, 12]. While the study of HEMSs has emerged across a variety of disciplines, results reveal a bias towards technical perspectives with a need to investigate more holistic solutions that allow to assess the impacts of this type of technology [79]. Furthermore, in the context of this thesis, this implies to find a balance between the research on technical tools for HEMSs under realistic assumptions about the energy market, the power system, users capabilities, and access to technology.

Thus, prior to the research and development of the decision support tools for HEMS, understanding current market conditions, the rules that consumers may obey, current data sources, among other factors, are key to focus the main body of work. This is done to focus the efforts on topics relevant to a HEMSs implementation accessible in the near future to the majority of Danish residential consumers. With this in mind, the rest of the present chapter continues with an overview of the energy market in

Denmark and an introduction to the different HEMSs elements that must be considered.

2.2 Energy market overview

The energy market in Denmark can be characterized by three main elements: making electricity through generation, transporting the electricity, and selling it to the final consumer. Energy companies can work in almost any of these areas because the energy market in Denmark is privatized. This means that private companies make sure the final consumer has the energy they need. It also means that consumers can choose which companies supply their energy. Figure 2.1 presents a graphical representation of the energy market structure in Denmark. Furthermore, a brief description of the elements that compose the energy market is presented next.

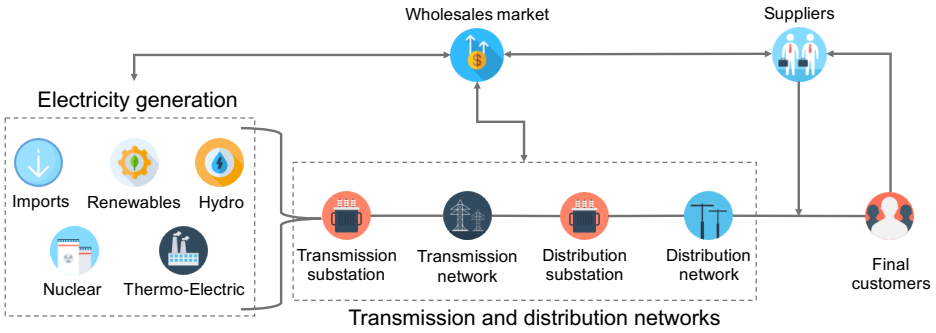

Figure 2.1: Illustration of the energy market elements in Denmark.

- **Electricity generation**

  Most electricity is generated at large power stations connected to the national transmission network. However, electricity can also be generated in smaller scale power stations which are connected to the regional distribution networks. The number and type of power stations built is the decision of each individual company based on market signals and government policy on issues such as the environment.

- **Transmission and distribution networks**

  Two main networks categories are part of the electricity system, that is transmission and distributions networks. While transmission networks deal with the mission of carrying electricity long distances around the country, distribution networks run at lower voltages and take electricity from the transmission system into homes and businesses. Moreover, the company in charge of the transmission and security of the power system in real time is the transmission system operator (TSO). The TSO also coordinates supply and demand of electricity in order to avoid fluctuations in frequency or interruptions of supply. In the Danish case, the TSO is named Energinet.

- **Energy supply**

  Suppliers (also known as retailers or utility companies) buy energy in the wholesale market and sell it to consumers. Suppliers work in a competitive market where consumers are free to choose any supplier to provide them with gas and electricity.

- **Energy regulation**

  The energy market in Denmark is regulated by the Danish Energy Agency, which is part of the Ministry of Climate, Energy and Utilities. The agency is responsible for tasks linked to energy production, supply and consumption, as well as Danish efforts to reduce carbon emissions. The agency is also responsible for supporting the economic optimization of utilities that includes (in addition to energy) water, waste, and telecommunications [23].

- **Wholesale market**

  The wholesale market refers to the place where suppliers buy the electricity they sell to the end-user. The price is influenced by a number of factors; the price of the input fuel used to produce electricity, as well as demand spikes and supply changes. Companies are allowed to buy wholesale energy weeks, months and years in advance as well as on the day of use. Although different platforms are used for trading electricity, most of the electricity in Denmark is traded in the Nord Pool power market. This power market is Europe's leading power market and offers trading, clearing, settlement and associated services in both day-ahead and intraday markets across 16 countries [89].

In the context of HEMSs, current regulation allows residential users to interact with the energy markets solely through electricity retailers for procurement or selling electricity. Recent studies are pushing new business models in the energy markets that would take advantage of the consumers capabilities, specifically, the figure of aggregators is nowadays commonly used [65]. In the Danish case, the TSO defines aggregators as "a person, business or technology which brings together (or aggregates) a set of related products, in this case upward or downward regulation capacity in the electricity market" [31]. Furthermore, the aggregator initiatives are still in early stages with pilot and market projects in current development by third parties with support from the Danish TSO [32]. Although there is limited information about commercial implementation of aggregators, the current rules and regulations indicate that a group of technologies are needed for the prosumer side.

Energy markets are complex systems composed by different elements which themselves are complex as well. It is out of the scope of this thesis to provide an in depth description of such elements. For a complete description see [95, 81, 84].

2.3 HEMSs architecture

A HEMS can be defined as a set of computer-aided tools used to monitor, control, and optimize the use of energy at a residential level. A typical HEMS relies on a combination of minimal elements such as smart meters, sensing devices, communication hardware and protocols, smart appliances, control and optimization techniques [12]. Moreover, additional systems and elements such as home size RES, ESS, and EVs could also be integrated, which increases the scale and complexity of the HEMS. Figure 2.2 presents an illustration of a complete HEMS architecture.

Depending on the architecture, HEMSs provide a unidirectional or bidirectional communication between homes and the electricity utility (or provider of the HEMS platform) to monitor, control and analyze the data that involves the consumption of electricity in the smart home [114]. The energy management controller interacts via local communication to different elements such as smart appliances, electric car chargers, ESSs, and home RESs in order to get the latest status of the different systems components. Furthermore, through an internet connection, relevant information such as electricity prices, CO_2 emissions, numerical weather predictions (NWPs), and manual user preferences are inputed to the controller for the decision-making process. Posteriorly and depending on the HEMS capabilities, different decision

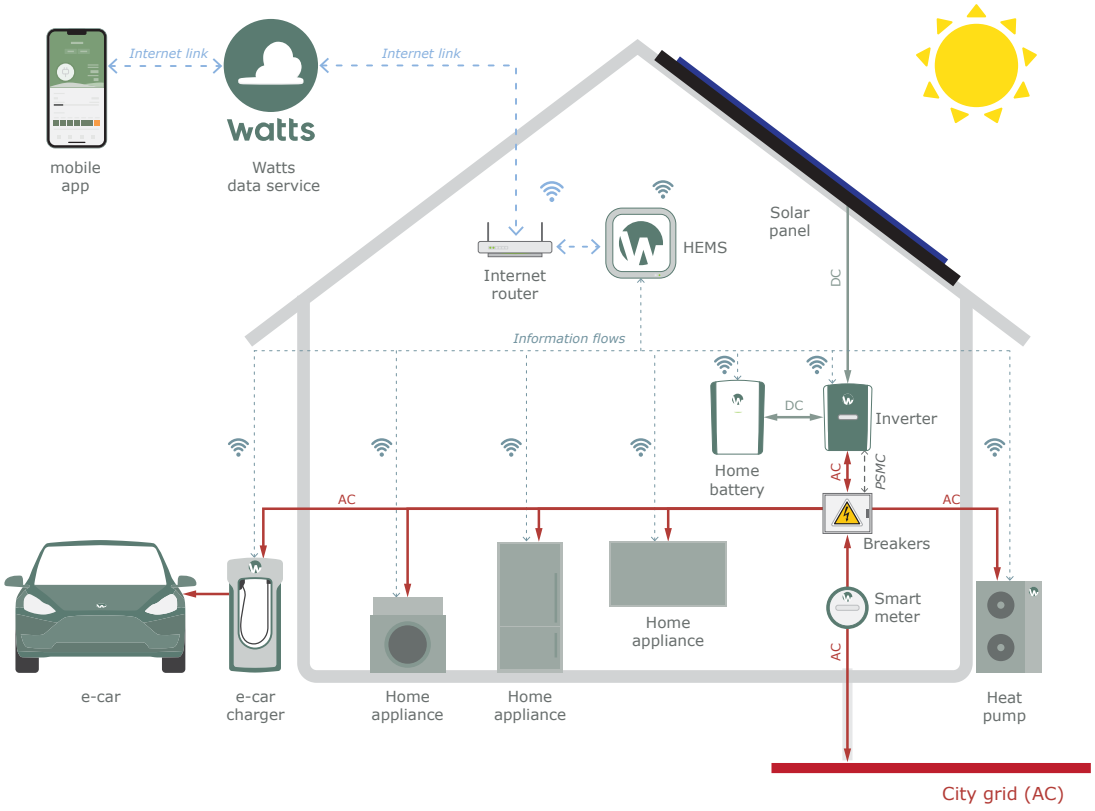

Figure 2.2: Illustration of a HEMS arquitecture.

support tools are triggered. Algorithms for different purposes are triggered for tasks such as forecast of future RES generation, demand, and prices, together with optimization and control routines that schedule smart appliances and charge/discharge of ESS and EVs. Several key elements are needed for the HEMS to work, a brief introduction of these elements is presented next:

- **Data access**

  Several data signals are needed for the HEMS functioning: historical data of the home's non-controllable load, smart appliances, charge/discharge EV sessions, and RES production. This data is combined with exogenous historical and future information such as NWPs, electricity prices, CO_2 emissions, and users' preferences to form coherent datasets needed for different purposes.

- **Mathematical and statistical models**

  Different mathematical and statistical models are needed in a HEMS. The HEMS elements rely on models to describe their behavior. Thus, different techniques are used to model elements e.g., RES production, electricity consumption, batteries behavior, user preferences, electricity prices, CO_2 emission, physical system conditions, to name a few. These models are closely interconnected and depended on datasets for their optimal parameter estimation. Finally, proper calibrated models are used for forecasting future behavior, as part of other HEMS components, and/or to provide insights on system performance.

- **Optimization and control**

  Data and models are combined and used for control of the different HEMS elements. In this regard, different optimization techniques are used in combination with other control strategies to provide set-points that the different HEMS components should follow. These elements try to optimize a given success criteria defined by the end-user, such as minimize costs or CO_2 emissions.

Please note the previous HEMS overview and key elements are presented within the scope of this thesis. This means that important elements such as electrical phases, voltage control, communication software and hardware details, internet of things (IoT) devices, electronic specifications, user interfaces among others are not included, however, they play a critical roll for the system functioning. The rest of this section will be focused on the description of the different data elements available for the project. Some of the data sources are tightly connected to the electricity market. This means that data availability plays a role in shaping the other key elements of the HEMS. Given that the mathematical and statistical models, and the control strategies compose an extensive part of the research, complete chapters are dedicated to these elements.

2.3.1 Data access

The HEMS elements require reliable access to different data coming from multiple sources. This section presents details of the different data elements together with an introductory characterization of the data.

2.3.1.1 Smart meter data

Energinet is an independent public enterprise owned by the Danish Ministry of Climate and Energy. The company owns, operates, and develops transmission systems for electricity and natural gas in Denmark. Energinet is the TSO and part of its responsibilities is to provide access to CO_2 signals, data from the electricity market and consumption and production data [30]. Moreover, consumers grant secure access to their data through Energinet's third party access-solution. The solution applies to electricity agents, energy consultants or other market participants authorized by Energinet to collect data. Once the authorization is granted, the third party has access to consumption data at an hourly resolution in kWh units. Furthermore, Watts A/S, as an authorized third party, collects electrical load (EL) (consumption data) data from its users. Figure 2.3 presents an example of the data collected for the year 2020 for one particular user.

From the figure it can be seen that the EL time series presents distinctive features such as non-stationarity and stochasticity. Although the plot's time resolution does not allowed, multi-seasonality features are expected to be found given the daily, weekly, and yearly patterns inherent to the electricity consumption. The EL data is a central element for the HEMS decision making process and it is therefore, presented and modeled in detail in Chapter 3.

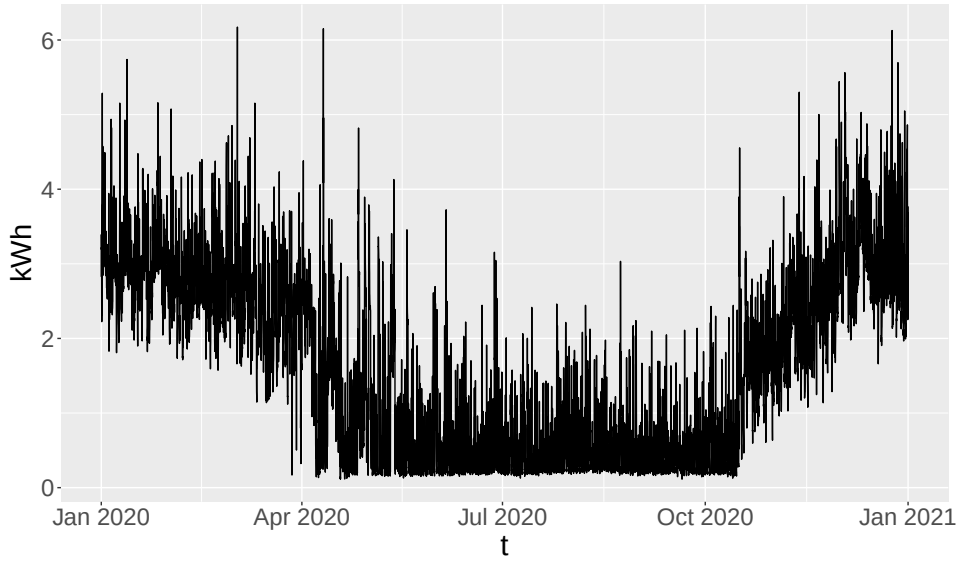

Figure 2.3: EL from a residential household. The data comes from a smart meter at an hourly resolution for the year 2020.

2.3.1.2 Weather information

It is common knowledge that weather has a significant influence on energy systems as energy production, demand, and therefore prices are affected by it. At a residential level, weather also has a direct impact on electricity consumption and production, as weather influences RES and consumers' electricity use. Thus, weather information is typically used as input signals in several of the mathematical models describing different HEMS elements. Moreover, weather information is available from different sources (free access or by paid subscriptions) in the form of NWPs. In our particular case, NWPs are provided by the OpenWeatherMap service at an hourly time resolution. The main variables included in their forecast are:

- Precipitation
- Wind direction
- Wind speed
- Wind gust
- Ambient temperature
- Pressure
- Humidity
- Cloud coverage
- Visibility

The data is accessible via an API, details of the NWPs data can be consulted in [90].

2.3.1.3 Electricity prices

Although traditional fixed tariffs are still the main way consumers are charged for their electricity consumption, the high penetration of smart meters in Denmark has opened the door for electricity retailers to offer hourly prices to residential consumers. Typically, this type of tariff is derived from the day-ahead wholesale electricity market prices, also known as ELSPOT [44]. In the ELSPOT market, different zones/regions have their unique day-ahead prices. In Denmark, two zones exist: Western Denmark (DK1) and Eastern Denmark (DK2). To obtain electricity prices for the residential consumer, retailers add taxes, levies, and fees to the day-ahead prices. In our particular case, access to the final

prices for consumers is provided through Watts A/S platform. Furthermore, the current Danish regulation allows residential consumers to sell their surplus electricity back to the grid. The feed-in-tariff is decided by retailers, with most of them offering the ELSPOT price minus associated operational fees as a feed-in-tariff to consumers. For illustrative purposes, the day-ahead and retail prices as sale and purchase prices respectively for DK2 are presented in Figure 2.4.

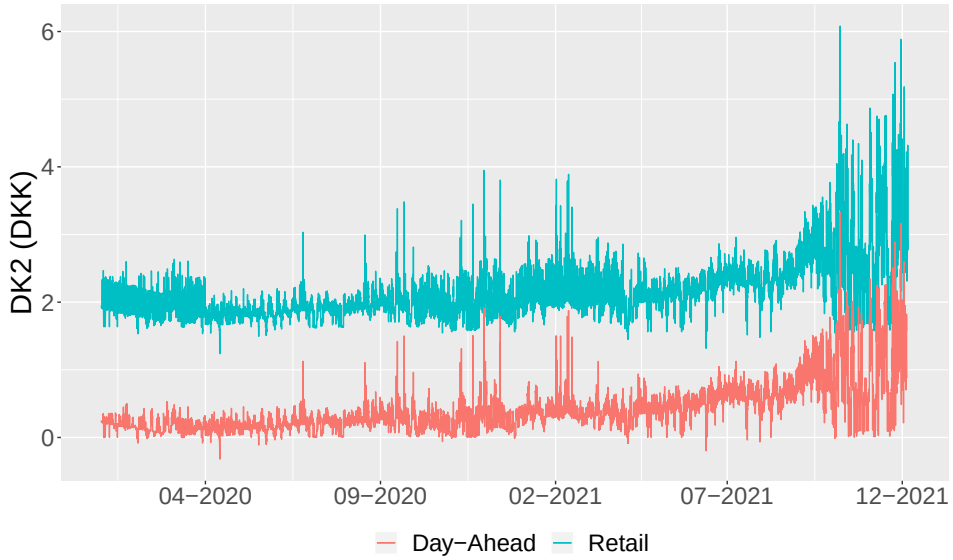

Figure 2.4: DK2 retailer and day-ahead prices for period 01-01-2020 to 20-12-2021.

Given that consumers interact with electricity markets through suppliers, it is the responsibility of the supplier to transmit price information to their clients. Thus, access to price data is expected to change according to the supplier, in our particular case, prices are accessible through the Watts A/S platform at a hourly resolution. Furthermore, electricity prices are the main signal driving the optimization and control strategies and elements such as volatility, non-stationarity, and seasonality of the prices play a main role in the HEMS decision making process. The impact of prices on HEMSs control strategies is explored in detail in Chapter 4.

2.3.1.4 CO_2 emissions

As part of the Watts A/S initiatives to address climate change, CO_2 emissions associated to the electricity generation are estimated, forecast, and shared with the consumers. This data is calculated based on information provided by the European network of transmission system operators for electricity (ENTSO-E). ENTSO-E is the association for the cooperation of the European transmission system operators, it includes 39 TSOs representing 35 countries [36]. Their transparency platform host different power generation data needed for the estimation of the CO_2 emissions (for details see [37]). For illustration purposes, Figure 2.5 presents the estimated CO_2 emissions for the period 01-01-2019 to 15-03-2021.

Although it might be difficult to see in the plot, seasonal behavior similar to the one seen in electricity prices is expected for CO_2 emissions, with higher emissions during colder months in comparison to warmer months.

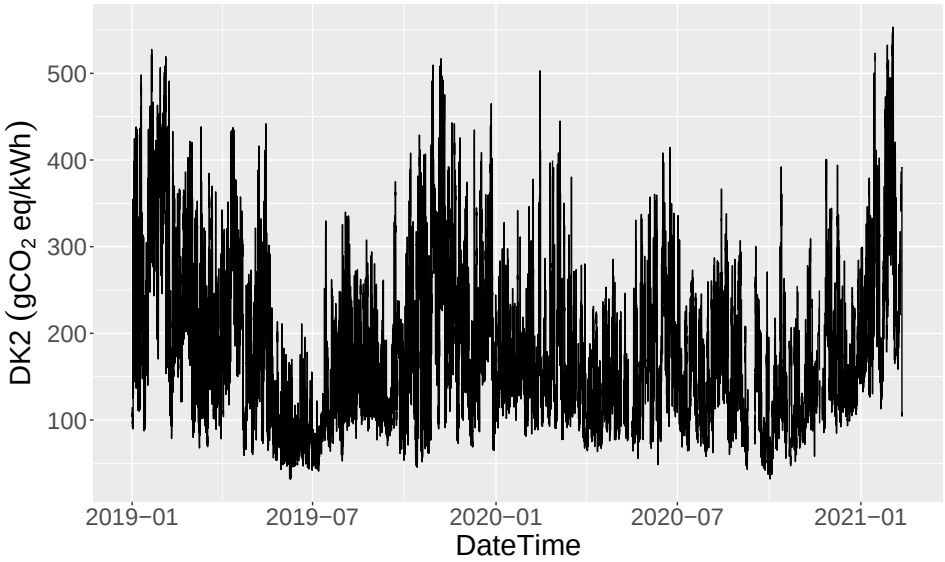

Figure 2.5: CO_2 equivalent emissions associated with electricity generation.

2.3.1.5 Local devices

Although sensing devices and smart appliances are available to residential users, platforms that collect and produce coherent datasets are not easily available. In our particular case, consumers data was mainly accessible through smart meters. This implies that all behavior happening behind the meter was not accessible. Having detailed information about PV generation, smart appliances consumption, and EVs charging sessions was not available at the beginning of the project. However, the efforts made by Watts A/S to build a HEMS-ready platform provided integration to EVs charge sessions towards the end of the project.

2.3.2 Simplification of the HEMS architecture

It is evident that managing every single element present in a complete HEMS architecture is a complex task. Furthermore, even if time would allow it, a possible real-time implementation is limited by current market rules, technologies, and data access. For this reason, and considering that this is a project with an industrial application in mind, a substantial amount of time was dedicated to study and select the HEMS elements to be part of the research.

In the first instance, it was decided that a minimal control approach was desired. This meant that HEMSs elements which require manual intervention with the end user, such as smart appliances were avoided. This was motivated by two main reasons. First, although smart appliances are being adapted more in households, at the moment, there are not widely adapted communication protocols. This can be problematic if we consider the software development of HEMS. Supporting different smart appliances protocols may be complicated and prone to errors. The second reason is associated with the actual potential these technologies may bring. Several studies and pilot projects have questioned the real benefits that consumers will be able to perceive, with results showing that electricity savings are minimal and in some cases these technologies cause energy intensification [50, 86, 85].

On the other hand, elements such as PV systems, batteries and heat pumps have been proven to bring significant savings for users [77, 42, 49]. EVs are also a key element that must be considered given that

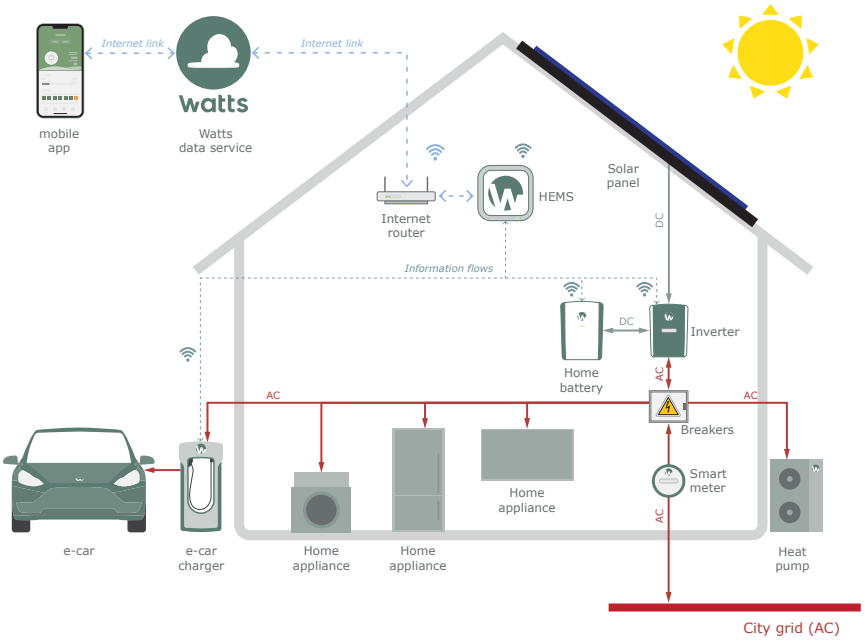

Figure 2.6: Illustration of a simplified HEMS architecture.

it implies significant electricity consumption for households, although their adoption continues to be low in Denmark [68]. Thus, originally presented HEMS architecture was reduced to the above mentioned elements and is presented in Figure 2.6.

CHAPTER 3

Multivariate probabilistic forecasting

This chapter focuses on the development of a multivariate probabilistic forecast methodology suitable for HEMSs. In particular, a brief argumentation for the development of this methodology is presented in Section 3.1. A generic description of the modelling methodology is presented in Section 3.2. The methodology is applied to the prediction of residential electrical loads, with the results being presented in Section 3.3. Possible applications to different domains and extensions of the modelling methodology are presented in Section 3.4, with a discussion of this research being introduced in Section 3.5.

3.1 Motivation

The implementation of HEMSs has several technical challenges such as the prediction of the home electricity demand, i.e., electrical load forecasting (LF). Forecasting is essential for the optimal planning and procurement of electricity in homes, in order to maximize the economic and/or environmental benefit without compromising users' comfort [54]. At a home level, LF presents several challenges given its multi-seasonality, non-stationarity, and stochastic characteristics. Additionally, multi-step predictions are often needed for the HEMS decision making process, which implies that a multivariate setup is needed.

A literature study included in Appendix A.1 showed that there is extensive literature on LF, specially point forecasting methods, i.e. estimate the expected value of the process and make a general assumption about the process distribution. Given the complexity of the EL, having a point forecast is often sub-optimal and more information about the multivariate distribution is needed. Thus, recent studies have applied probabilistic forecasting models. These studies have proven to yield accurate results for individual forecast horizons. However, most of the studies do not address the temporal correlation between the forecast horizons.

With the above in mind and considering our HEMS setup, the hourly data resolution of the EL, and that no details about the physical system are provided, a modeling methodology for probabilistic load forecasting (PLF) suitable to our needs was developed. The modeling methodology is introduced in the following section.

3.2 Modelling methodology

The following model description is based on the methodology presented in Appendix A.2. Let us start by introducing the notation to be used throughout this chapter. Let $\boldsymbol{X}_t = (X_{t,1}, X_{t,2}, \ldots, X_{t,K})^\intercal$ denote a K-dimensional random variable at time t for future values $t+1, t+2, \ldots, t+k$. Upper case letters are used for random variables while lowercase letters denote the corresponding observations. Moreover, vectors and matrices are emphasized using bold font. Thus, $\boldsymbol{x}_t = (x_{t,1}, x_{t,2}, \ldots, x_{t,K})^\intercal$ is used for the realizations of the random vector $\boldsymbol{X}_t$.

Let Y_t be a univariate time series (TS). The forecast of Y_t can be expressed as a multivariate variable at each time t for future lead times $t + k$ until the maximum future horizon K as

$$\boldsymbol{Y}_t = (Y_{t,1}, Y_{t,2}, \ldots, Y_{t,K})^\intercal \tag{3.1}$$

where we use the previously introduced notation s.t.

$$Y_{t,k} = Y_{t+k} \quad \forall\, k \in \{1, 2, \ldots, K\} \tag{3.2}$$

Our interest is to forecast the probability density function (PDF) for $\boldsymbol{Y}_t$ denoted by $f_t(\boldsymbol{y})$. Proposing a functional form for f_t implies a simultaneous description of both the marginal densities (for each horizon k) as well as the temporal dependency [111]. Two approaches have been proposed to estimate f_t. The first approach uses recursive least squares (RLS) models for the estimation of the expected value $\widehat{\boldsymbol{Y}}_t$ and then analyses the models' residuals in order to obtain an estimation of the marginal densities as well as the temporal correlation of $\boldsymbol{Y}_t$. The second approach uses quantile regression models for the prediction of the marginal distributions of $\boldsymbol{Y}_t$, and a Gaussian copula normalizes the process. In both approaches, two methods were considered to model the correlation structure. The first method estimates the correlation structure from the data with a full covariance matrix model, while the second method uses an auto-regressive (AR) model – in which the covariance matrix is parameterized with only a few parameters (order of the AR process + 1). Figure 3.1 presents a graphical summary of the modelling methodology process. A complete mathematical description of all of the elements presented in the modelling methodology can be seen in Appendix A.2.

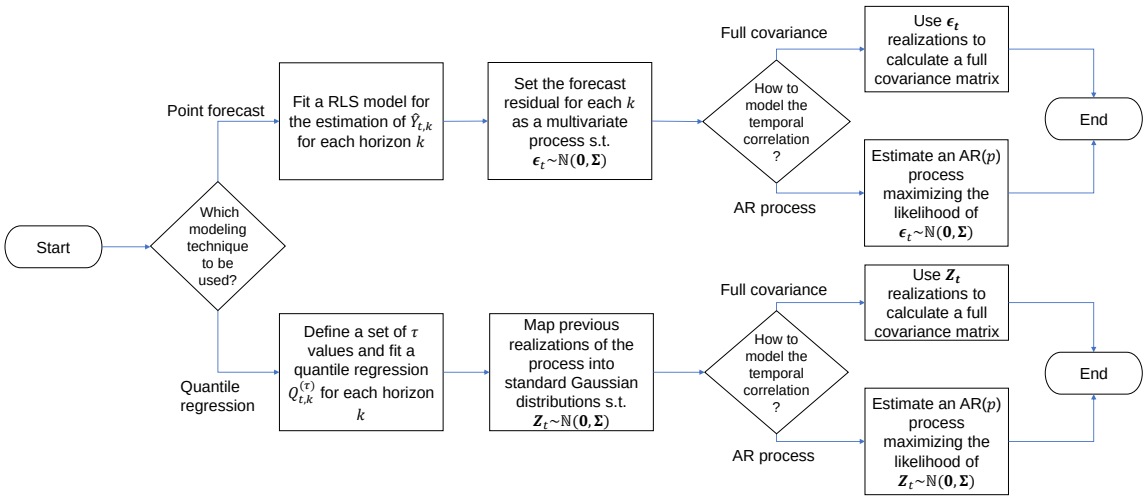

Figure 3.1: Process graph summarizing the modelling methodology.

3.3 Application on residential probabilistic load forecasting

3.3.1 Data description

As a proof of concept, the presented modeling methodology was applied to forecast the EL of a residential home. The data available for the test corresponds to the period 2017-06-01 to 2018-05-31 in hourly resolution. Little is known about the home besides the number of inhabitants, and that the home uses a heat pump during winter and it is not cooled during summer. The TS plot and data distribution over the mentioned time period can be seen in Figure 3.2. The EL distribution shows a natural lower bound of zero with some high values, which clearly makes the distribution right skewed. Additionally, given the data comes from a real inhabited residential home, time periods of near-zero load are seen. These low consumption periods can be attributed to absence of the house inhabitant due to holidays or similar events.

Independent variables are defined from NWPs from the provider mentioned in Section 2.3.1.2. Furthermore, as described in Appendix A.3.1, ambient temperature and solar radiation are proven to have a significant influence on the EL, so they were included as independent variables. Moreover, the solar irradiation signal was derived as a combination of the global horizontal irradiation (GHI) and the percentage of cloud cover provided by the NWPs. A selected short period of the TS is presented in Figure 3.3. A closer look at the EL plot allow us to identify the presence of intra-day patterns. This is due to normal inhabitants' activities such as cooking for breakfast and dinner, and evening activities. The patterns are typically different during workdays and weekends.

In order to account for the intra-day patterns seen in the data, Fourier series with n harmonics were included as independent variables. The procedure describing how to calculate Fourier series based on the time of day for week and weekend is fully described in A.3.1. The resulted Fourier signals are referenced as

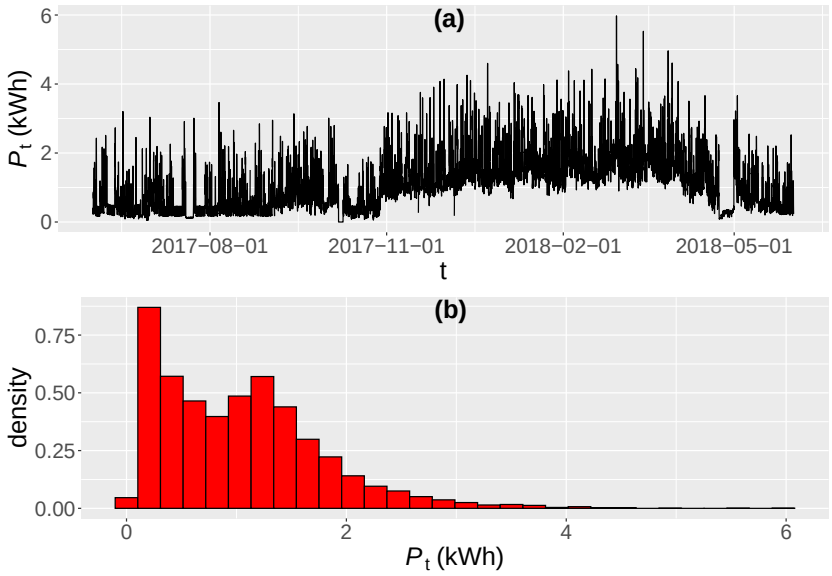

Figure 3.2: Electrical load P_t time series (a) and empirical distribution (b). The data corresponds to the period 2017-06-01 to 2018-05-31.

$$\begin{aligned}
\boldsymbol{F}_{t,k}^{(\sin,\, week)} &= (F_{t,k,1}^{(\sin,\, week)}, \ldots, F_{t,k,n}^{(\sin,\, week)}) \\
\boldsymbol{F}_{t,k}^{(\cos,\, week)} &= (F_{t,k,1}^{(\cos,\, week)}, \ldots, F_{t,k,n}^{(\cos,\, week)}) \\
\boldsymbol{F}_{t,k}^{(\sin,\, weekend)} &= (F_{t,k,1}^{(\sin,\, weekend)}, \ldots, F_{t,k,n}^{(\sin,\, weekend)}) \\
\boldsymbol{F}_{t,k}^{(\cos,\, weekend)} &= (F_{t,k,1}^{(\cos,\, weekend)}, \ldots, F_{t,k,n}^{(\cos,\, weekend)})
\end{aligned} \tag{3.3}$$

Moreover, it is natural for the EL to present a significant auto-correlation. Thus, with $Y_{t,k}$ as our dependent variable of interest, we can include the lag value y_t as an independent variable to account for this effect. An entry in the design matrix will contain ambient temperature, solar irradiation, a lag value, and $2n$ pair harmonics as independent variables for a future time $t + k$ given t, s.t.

$$\begin{aligned}
\boldsymbol{x}_{t,k} = (T_{t,k}^{(\mathrm{a})}, G_{t,k}, y_t, \boldsymbol{F}_{t,k}^{(\sin,\, week)}, \boldsymbol{F}_{t,k}^{(\cos,\, week)}, \\
\boldsymbol{F}_{t,k}^{(\sin,\, weekend)}, \boldsymbol{F}_{t,k}^{(\cos,\, weekend)})^{\intercal}
\end{aligned} \tag{3.4}$$

3.3.2 Simulation setup and performance metrics

A simulation study was designed to resemble a real-time application. The aim was to produce a probability forecast of the EL at time t that would cover the following 24 hours in an hourly resolution. Considering the description presented in Section 3.3.1, a logarithmic transformation was done in order to ensure the zero-bound feature of the EL. Thus, the dependent variable of interest Y_t was defined as

$$Y_t = \ln(P_t) \tag{3.5}$$

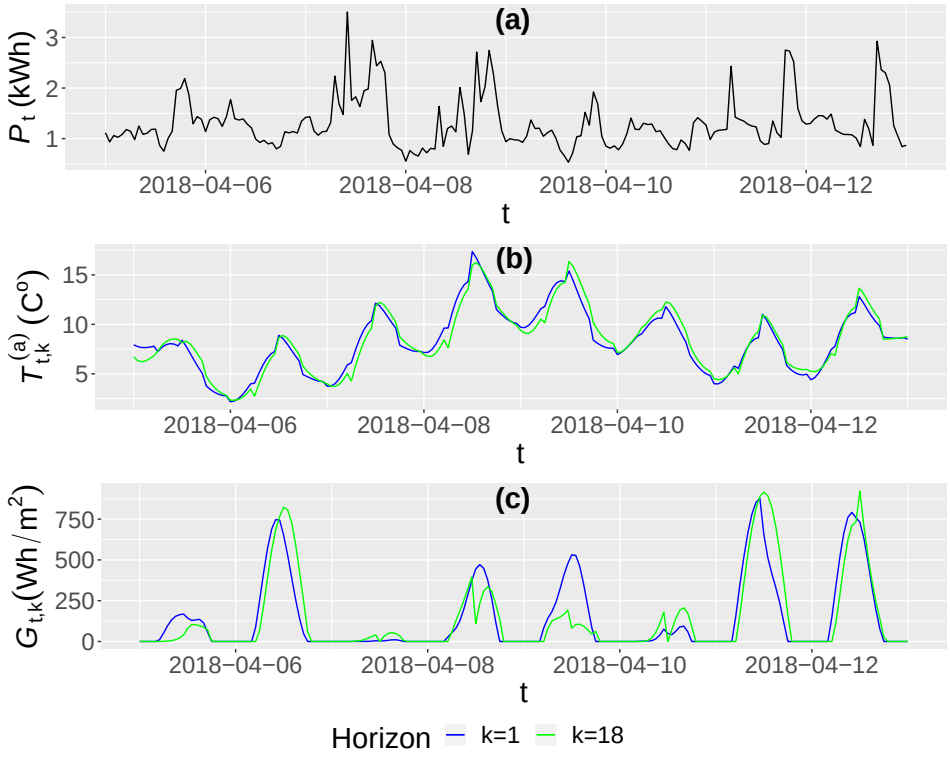

Figure 3.3: (a) Electrical load P_t. (b) Ambient temperature $T^{(a)}_{t,k}$. Solar irradiation $G_{t,k}$. The data corresponds to the period 2018-04-05 - 2018-04-20.

The forecast is expected to be updated every hour, considering the independent variables' latest available information $\boldsymbol{X}_t = [\boldsymbol{x}_{t,1}, \ldots, \boldsymbol{x}_{t,K}]$. The online setting implies updating the models' parameters at each time t. While this is inherent to the RLS, the quantile-copula and the models used for covariance estimation assume a time-invariant framework. Thus, a sliding time window approach was considered in order to deal with the time dependency. This approach will re-estimate the models' parameters at time t using the latest N values. In this way, the non-stationarity characteristics of the EL are modeled. Figure B.5 summarizes the simulation setting.

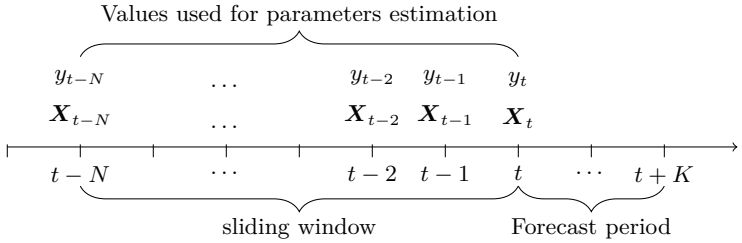

Figure 3.4: Graphical representation of the simulation setting for time t.

The simulation setting described previously was used to validate the performance of the proposed models. The period 2018-01-17 to 2018-02-16 was considered as test period. The selection aims to validate

the results during winter time, where the volatility of the EL is higher. The time period 2017-06-01 to 2018-01-16 was used for the RLS calibration. Moreover, five models were implemented. The first model is called "RLS-free", which indicates the use of the RLS method with the correlation structure being estimated directly from the empirical covariance matrix. The second model is called "RLS-AR", which indicates the use of the RLS method with the correlation structure estimated using the maximum likelihood procedure presented in Section A.2.2.3. The third and fourth models are called "Copula-free" and "Copula-AR", which indicates the use of the quantile-Copula method with the same correlation structure connotations. The fifth model is called "RLS", indicating the use of the RLS method without correlation structure, it is used as a reference model. Furthermore, the density function is estimated numerically by using 500 scenarios of $\boldsymbol{Y}_t$ at each time t for all models.

Different performance metrics are used to evaluate the reliability, sharpness, log-score, and the representation of the temporal correlation structure of the predicted values. These score metrics are: prediction interval coverage probability (PICP), prediction interval normalized average (PINAW), pinball score, continuous ranked probability score (CRPS), and the variogram score (VarS). The mathematical details of the different scores can be seen the Appendix A.2.

3.3.3 Results

The results presented in this section intend to present a summary of the complete results presented in Appendix A.4. Details related to the validation of the forecast models and estimation of the temporal correlation structures are omitted here but can be seen in the previously referenced appendix.

Simulations using the different models were carried out. In Figure 3.5 the probabilistic forecast, the EL measurements, and 5 random scenarios using the different models are shown. The forecast period starting at t = 2018-02-14 18:00 was selected as an illustrative example. Note that the results are presented in the EL original domain, which implies the use of the exponential function on the different models' results. From the RLS models we see an almost symmetric distribution around the mean, which is expected due to the normality assumption made for the residuals. This naturally differs from the distributions coming from copula models, where we forecast the whole distribution without assuming symmetry at all. Furthermore, the effect of the temporal correlation models can be seen in the smoothness of the scenarios in comparison with the "RLS" model, which presents a more erratic behavior.

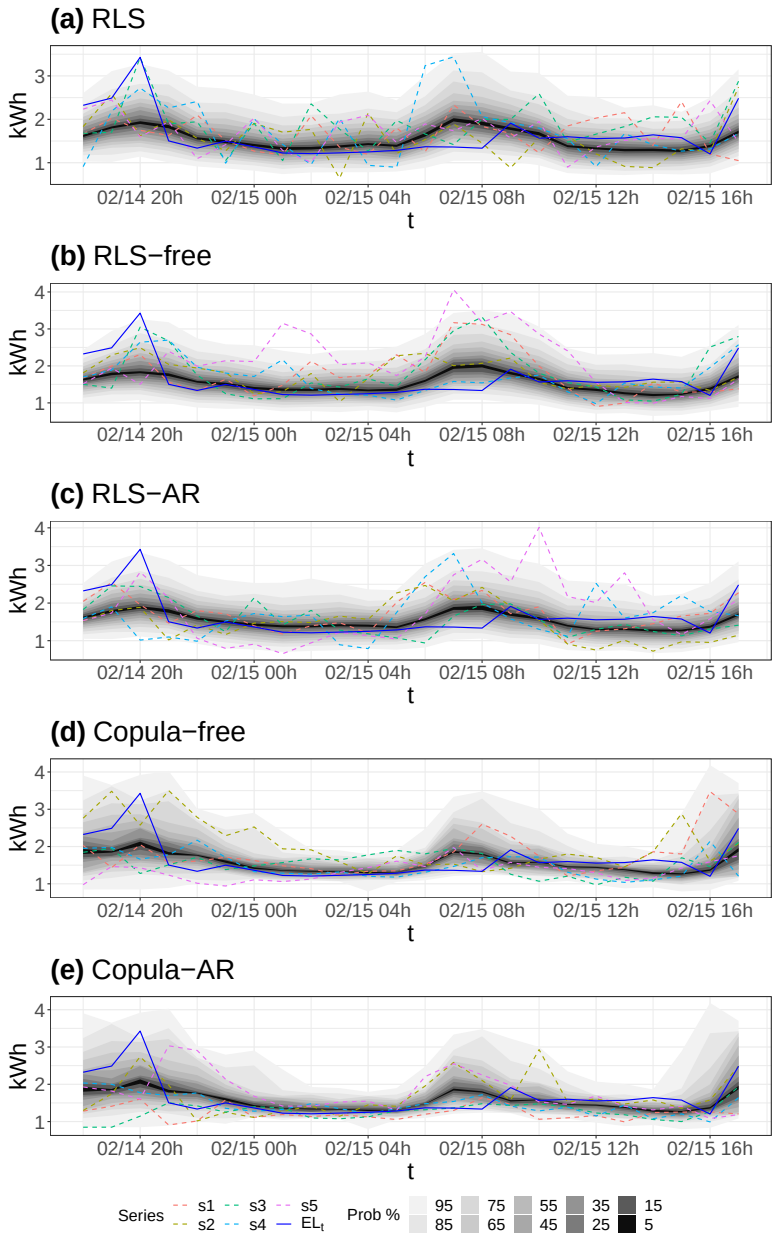

Figure 3.5: Probabilistic forecast, EL realisation and 5 scenarios using the RLS (a), the RLS-free (b), the RLS-AR (c), the Copula-free (d), and the Copula-AR (e) models. The forecast period corresponds to 2018/02/14 18:00 - 2018/02/15 18:00.

In order to properly score the performance of the models, a quantitative analysis must be applied. Thus, the different metrics presented in Section 3.3.2 were calculated for the different scenarios generated with each model. The mean CRPS, total pinball loss sum, PICP, and PINAW per horizon k are presented in Figure 3.6. The plots show a similar performance between RLS models and a similar performance between copula models. When comparing the performance between the different methods, it can be seen that the quantile-copula presents better performance than the RLS for the metrics that evaluate the probability distribution as a whole (CRPS and pinball loss). However, we can see that for $k = 1$ the RLS presents better performance. This is explained by the effect of the lag value included as input

in the models, with the lag value effect degrading with longer prediction horizons.

The PICP and PINAW scores were used to evaluate the reliability and sharpness of the forecast. From the PICP results, we can see that both methods present a good reliability based on a 90% nominal coverage interval, with the RLS outperforming the quantile-copula. However, the higher reliability of the RLS models comes at the expense of a considerable bigger prediction interval, as can be seen in the PINAW results. Furthermore, from the PINAW results one might argue that prediction intervals are on average large (up to 50% of the test period data range for the RLS for longer horizons). However, one should consider the level of random variation of the process at hand. Predicting with a high degree of sharpness the consumption of a single house is a complex task if we consider the data characteristics presented in Section 3.3.1 and the longer time horizons.

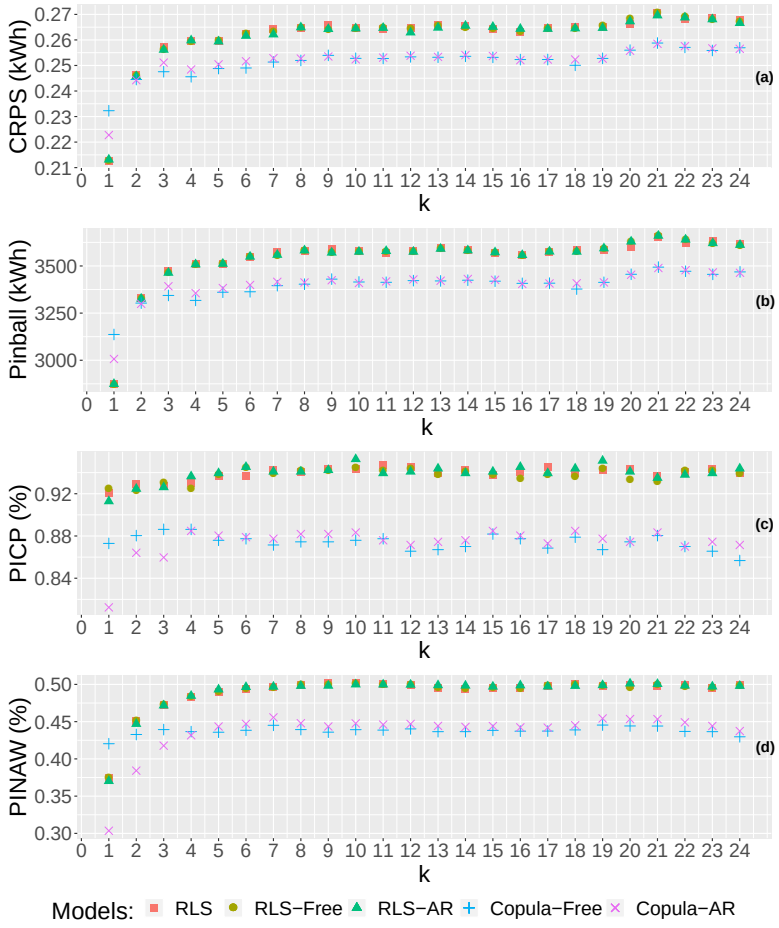

Figure 3.6: Mean CRPS (a), pinball loss sum (b), mean PICP (c), and mean PINAW (d) per horizon k for the test period 2018-01-17 - 2018-02-17. Note that the PICP and PINAW are based on a 90% prediction interval.

The VarS was also calculated for the test period, Table 3.1 presents a summary with the mean score and the percentage of change in comparison to the reference model "RLS". As expected, the reference model presents the worst performance given the lack of a description of the temporal correlation in the model. The proposed methods present a significant performance improvement in comparison with the reference – with the quantile-copula models presenting the best results.

	Mean VarS $(\text{kWh})^2$	VarS % of change
RLS	9.52	-
RLS-free	7.74	-18.68%
RLS-AR	7.79	-18.13%
Copula-free	6.79	-28.76%
Copula-AR	6.82	-28.35%

Table 3.1: VarS result for the proposed methods and percentage of change relative to the reference model (RLS). Calculations were made for the test period 2018-01-17 - 2018-02-17.

VarS results were also discriminated by difference $j-i$ of the vector entries (see Appendix A.4.3). Results show that the quantile-copula-based method presents a better performance for all $j-i$ differences compared to the RLS-based method. This supports the aggregated results introduced previously.

3.4 Extension to multivariate temporal forecasting

Although this chapter is entirely dedicated to the prediction of residential EL, the modelling methodology is presented in a general form which allows its application in other areas. In fact, the research work presented in Appendix A was done simultaneously with the development of a software package called `onlineforecast` in the programming language `R`. The package is based on the experience and knowledge from the team at DTU Compute: Dynamical Systems (DYNSYS) [7]. The development of the package was led by associate professor Peder Bacher, who is also one of the supervisors of this PhD project.

As it is presented in the package documentation, a similar modelling approach to the one presented in Section 3.2 has been applied for short-term heat load forecasting for single family houses [10], short-term solar power forecasting [8], and load forecasting of supermarket refrigeration [101]. The main difference between the package and the research presented in Appendix A is the probabilistic framework where the temporal dependency is addressed. This is something that is not present in the package at the moment. Moreover, the presented methodology addresses the temporal correlation once the data is Gaussian (Gaussian residuals in the case of the point forecast approach or Gaussian transformation under the quantile-copula). This opens the possibility to address the cross-correlation present when forecasting several multivariate TS. This idea is easy to explain and generalized and it is introduced next.

Let $\boldsymbol{Y}_t^{(1)}$ and $\boldsymbol{Y}_t^{(2)}$ two different multivariate variables as described in Section 3.2. Let us assume that these TS besides having inner temporal dependency, they also present dependency between each other. The presented modelling methodology could be expanded such that both TS are considered simultaneously. Thus, addressing the temporal and cross correlation properties of $\boldsymbol{Y}_t^{(1)}$ and $\boldsymbol{Y}_t^{(2)}$ at the same time. Mathematically that is

$$\boldsymbol{Y}_t = (\boldsymbol{Y}_t^{(1)}, \boldsymbol{Y}_t^{(2)}) \tag{3.6}$$

then, the process continues as described in 3.1. Please note that given that in this particular case two different multivariate TSs are treated as a whole vector, only the full covariance method is applicable. The AR approach as described in Appendix A.2.2.3 is designed for one multivariate variable

at a time. Modelling a covariance structure from a multivariate AR process is a topic of future research.

3.5 Discussion

Two different forecasting methods based on RLS and quantile-copula modelling techniques were fully analyzed. Furthermore, two different ways of modelling the temporal correlation structure were investigated as part of the implemented models. The results indicate that modelling the temporal correlation has a significant impact on the performance of the forecasting models. Moreover, additional results presented in Appendix A.4 show robust behavior when checking the different assumptions of the models. While the results shown in this methodology rely on RLS and quantile-copula models, other modelling techniques could be used. Moreover, modelling the temporal correlation by the dynamics seen in the residuals opens the possibility to use a wide range of modelling techniques for the marginal distributions knowing that the time dynamics can be addressed in a different modelling stage.

Although overall results indicate a good performance of the modelling methodology, there are critique points that need to be considered such as:

- While the presented methodology is designed to adapt by the use of a recursive method or a sliding window, sudden regime changes in the household behavior, e.g., inhabitants going on holidays, may present a challenge. Thus, research on expanding the current methodology to a multi regime framework should be explored.
- The modelling methodology presented relies on behind the meter readings, which implies that the EL measured is the sum of controllable and uncontrollable loads. These loads could be modeled separately e.g., specific models for elements such as heat pumps and hot water storage units.
- Disaggregated load models will increase the complexity of the HEMS. Thus, modelling complexity should be evaluated not only in a traditional way (validating the model's assumption) but considering the end goal in mind (in light of the objective of the control strategy) as it is presented in Sections 4.3 and 4.4.
- The presented modelling methodology provides a good picture of the EL dynamics at a given discrete time $t + k$, however, the dynamics between the time steps is continuous and is not modeled. Knowing the dynamics between time steps could be critical for the close to real-time controllers that would run as part of HEMSs.

CHAPTER 4

Control strategies for home energy management systems

This chapter focuses on the implementation and results analysis of different control strategies for HEMSs. The motivation behind the selection of a specific control framework is presented in Section 4.1. A generic description of the modelling methodology is presented in Section 4.2. An initial application where the economic potential of HEMSs is simulated is presented in Section 4.3. This work is extended to a bi-objective case, where electricity costs and CO_2 emissions are considered together as outlined in Section 4.4. Application on EV charging is presented in Section 4.5. Finally, a discussion of this body of research is presented in Section 4.6.

4.1 Motivation

A central element of HEMSs is the control decision support tool. This control element is expected to be in charge of the operation of the components part of the HEMS presented in Section 2.3. In this regard, several studies have proposed sophisticated technical solutions. These solutions have typically assumed direct control of components such as heating systems, smart appliances, RESs, ESS, and EVs, with some parameters being defined by the consumer [103, 127]. Furthermore, it is important for the control strategy to consider complex system features such as the multi-seasonality, non-stationarity, and stochasticity of RESs and consumers' EL [26]. As highlighted in Appendix B.1, different studies have included several of these features by using different control techniques such as model predictive control, heuristics, and optimization-based control strategies. Furthermore, the reviewed literature also showed that there is a limited number of studies that comprehensively assess the potential of HEMSs for the end user under real market conditions and which take account of consumers' current capabilities. Not considering consumers' current capabilities could lead to misleading results as the assumptions would not hold in real applications, but only in controlled laboratory conditions and/or simulations.

In this regard, the simplified HEMS architecture presented in Section 2.3.2 was considered when researching possible control strategies. Let us recall that in our HEMS setup we focus on the control of a PV and battery system. Although a simpler HEMS setup may lead to a simpler control problem, market rules, price structure, and hardware capabilities make it a challenging problem. As it was introduced in Section 2.3.1.3, consumers have access to different prices for selling and buying electricity from the grid, which implies a need to differentiate between in/out power flows. Moreover, the generated

PV electricity can be used for charging the battery, to be returned to the grid, and/or directly for the home consumption. This implies that the battery could be charged from different sources e.g., the PV and/or the grid, and also it could discharged to the grid or to the household. At the same time, all of the power flows are subject to the type of electrical connection (AC or DC connection) typical on residential consumers, which implies that electricity losses must be considering in the control problem (see Figure 2.6 for a graphical representation).

Moreover, given that prices come in an hourly resolution, the control schedule is expected to operate accordingly. This means that a decision related to charge of discharge of the battery must be made. This decision should also be made considering not only current system information but also future information (future prices, expected EL, PV, and NWPs). Thus, the decision-making process carries hidden complexity given the different power flows in the system and the fact that charge and discharge events are mutually exclusive processes. In other words, given the characteristics of our system, the prices, and data resolution, binary decisions are involved.

Another key element to be considered is CO_2 emissions inherent to electricity generation. Typically the HEMS decision-making process focuses on cost minimization, however, analyzing the relation between prices and expected CO_2 emissions indicated that these signals are not aligned. Figure 4.1 shows the correlation between emissions and prices, the plot shows that in general the two signals present a weak correlation. This may imply that optimizing for cost minimization will yield a solution that may be far from being environmentally friendly. Thus, consumers could be interested in optimizing their electricity use according to their expected CO_2 footprint.

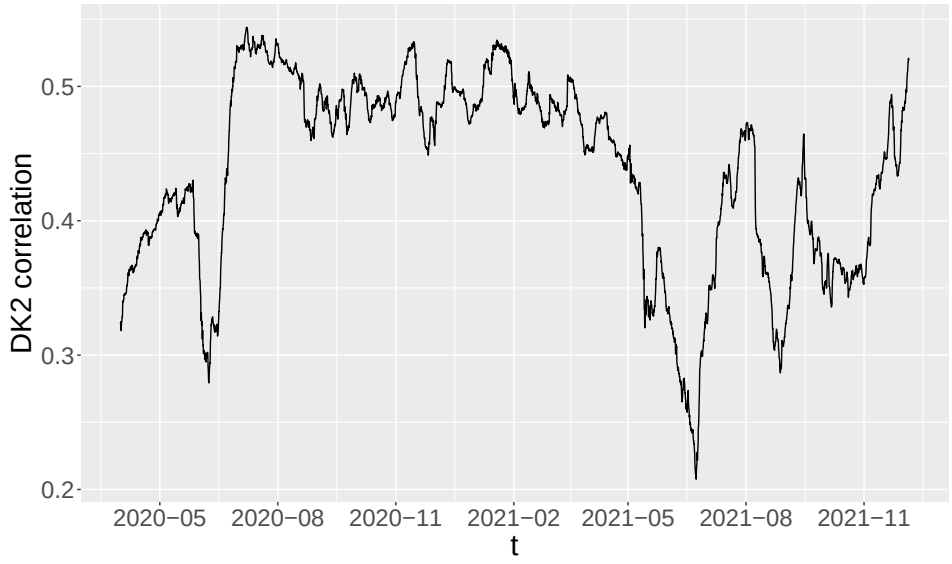

Figure 4.1: Price and CO_2 emission correlation. The correlation is calculated on using a rolling window with three month of data at a hourly resolution.

All of the above elements were considered when formulating the mathematical models of the HEMS and the corresponding details are presented in the following subsections.

4.2 Modelling methodology

The previously introduced arguments where considered when researching possible control strategies for the HEMS. Our simplified HEMS setup reduced the complexity of our problem to such degree that optimization-based control strategies were considered. If we recall, our HEMS setup relies only on the control of the battery with the dynamics of the household consumption and PV generation being independent of the control variables. This implies that this information could enter linearly into the optimization-based control strategy. Although battery dynamics are known to have a nonlinear behavior [61], linearity was assumed when modelling charge and discharge processes and the battery state of charge (SoC). This was found to be a common practice in the literature when working at an hourly time resolution [93, 28].

With the HEMS components entering linearly to the optimization-based control and binary decisions being involved in the decision-making process, a mixed-integer linear program (MILP) framework was selected as modelling technique. Moreover, MILP frameworks couple nicely with the ϵ-constraint method. This method was considered to address the multi-objective characteristic of our problem. A brief introduction to MILP problems and the ϵ-constraint method are presented next.

4.2.1 Mixed-integer linear programming (MILP)

Optimization problems in which some or all of the decision variables are restricted to integer values are known as mixed integer problems. For illustration purposes, the mathematical description of a MILP problem presented in [20] is introduced next.

$$\begin{aligned} \max_{x,y} \; & c^\intercal x + h^\intercal y \\ \text{subject to} \quad & Ax + Gy \leq b \\ & x \geq 0 \\ & y \geq 0 \end{aligned} \tag{4.1}$$

Where the data are row vectors c, h, an $m \times n$ matrix A, an $m \times p$ matrix G and a column vector b. Moreover, it is assumed that all entries c, h, A, G, b are rational and the column vectors x and y contain the variables to be optimized. The set of feasible solutions to the problem described in Equation (4.1) is given by

$$S := \{(x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^p : Ax + Gy \leq b\} \tag{4.2}$$

A mixed 0, 1 set is a set of the form presented in Equation (4.2) in which the integer variables are restricted to take the value 0 or 1 s.t.

$$S := \{(x, y) \in \{0, 1\} \times \mathbb{R}_+^p : Ax + Gy \leq b\} \tag{4.3}$$

Solving integer programs is a difficult task in general but well studied topic in the literature with multiple advanced software implementations available for use. In our particular case, it was out of the scope

to dig into MILP solution strategies, interested readers can find details in [20].

4.2.2 Stochastic programming

In real applications, many decision making problems are subject to uncertainty. Stochastic programming provides a framework in which the uncertainty of the parameter values and outcomes can be considered within an optimization.

In stochastic programming, each uncertain parameter is modeled as a random variable. Each uncertain parameter follows a probability distribution, which can be represented by a finite set of realizations or scenarios. Each scenario represents a possible outcome of the uncertain parameter. Moreover, we typically have a set of decisions to be taken without full information on some random events. These decisions are called first-stage decisions. Later, information is received on the realization of some random vector $\boldsymbol{\xi}$. Then, second-stage or corrective actions are taken. A general mathematical formulation of a two-stage MILP is (the description is based on [13])

$$\begin{aligned} \min_{x,y} \; & c^\intercal x + h^\intercal y + \ell(x,y) \\ \text{subject to} \quad & Ax + Gy \leq b \\ & x \geq 0 \\ & y \geq 0 \end{aligned} \tag{4.4}$$

where $\ell(x,y) = \mathbb{E}_{\boldsymbol{\xi}}\, \boldsymbol{Q}(x,y,\boldsymbol{\xi})$ is the value function or recourse function, $\boldsymbol{\xi}$ is a random vector, and $\mathbb{E}_{\boldsymbol{\xi}}$ denotes the mathematical expectation with respect to $\boldsymbol{\xi}$. Please note that in the above notation boldface denote random vectors.

4.2.3 Bi-objective optimization and the ϵ-constraint method

In single-objective optimization problems, as described in Equation (4.4), the aim is the minimization of a scalar objective function over decision variables x and y. If we denote the objective function f, in a multi-objective optimization, the assumption of f being one-dimensional is dropped, such that it is a vector of objectives $\boldsymbol{f} \in \mathbb{R}^M$ with M being the cardinality $|M|$ of the set of objectives $m \in M$. Moreover, the optimization problem consists of finding solutions x^* and y^* that are Pareto-optimal. A solution is considered Pareto-optimal if it is not dominated by any other solution that yields a lower objective value [25]. Furthermore, important terminology in multi-objective optimization includes the Pareto-front (the set of all Pareto-optimal solutions [18]), Nadir point (the vector of upper bounds on each objective in the set of Pareto-optimal objective space) and the utopian point (the vector of lower bounds on each objective in the set of Pareto-optimal objective space) [25]. In our particular application, only two objectives were considered (minimizing operational costs and carbon emissions).

There exits an extensive literature on methods to determine a set of Pareto-optimal solutions [82]. However, the ϵ-constraint method was considered because it is easy to implement. This method is an iterative approach where in each iteration i, one objective is optimized while the other objective is bound by a parameter. A general formulation of the problem can be written as

$$\begin{aligned} \min_{x,y} \; & f_1(x,y) \cdot \alpha + f_2(x,y) \cdot (1-\alpha) \\ \text{subject to} \quad & Ax + Gy \leq b \\ & f_2 \leq \frac{1}{1-i} f_2^- \\ & x \geq 0 \\ & y \geq 0 \end{aligned} \tag{4.5}$$

where $\alpha \in \{0, 1\}$ is a binary parameter used by the ϵ-constraint algorithm. The application of the ϵ-constraint method proceeds as follows: Iteratively, the MILP formulated is solved with decreasing upper bounds on the objective f_2. These bounds represent the ϵ-constraints and are linearly decreasing within a predefined interval. This interval ranges from the highest (f_2^+) and lowest (f_2^-) objective values. For a number of partitions I, a set of Pareto-front solutions will be given by $\boldsymbol{x} = \{x_1, x_2, \ldots, x_I\}$ and $\boldsymbol{y} = \{y_1, y_2, \ldots, y_I\}$. A pseudo-code description of the ϵ-constraint algorithm is presented Algorithm 1.

Algorithm 1 ϵ-constraint implementation

1: Solve (4.5) for $\alpha = 1$ and save objective values f_1^+ and f_2^-
2: Solve (4.5) for $\alpha = 0$ and save objective values f_1^- and f_2^+
3: **for** i = 0, ..., I-1 **do**
4: Solve (4.5) with upper-bound secondary objective $f_2 \leq \frac{1}{1-i} f_2^-$
5: **end for**

4.3 Application on economic evaluation of HEMSs

This section presents the simulation results of the HEMS setup described in Section B.2. It combines the data elements presented in Section 2.3.1, the EL forecast methods presented in Section 3, and the optimization framework introduced in Section 4.2. Moreover, all of the elements presented in the following subsections are included with further detail in Appendix B.

4.3.1 Data and simulation setup

Different data sources have to be combined for the HEMS decision making process. In this application, NWPs and historical data collected from smart meters are used as inputs for the PLF methods as it was described in Section 3.3. Furthermore, a simulation approach was used to approximate the PV generation. Given that the consumers have the possibility to sell their electricity back to the grid, different prices for selling and buying electricity were used. A detail description of all the data needed for the simulation can be found in Appendix B.3.

The simulation study was designed to resemble a real-time application. The aim of the simulation is to optimize the battery operational setpoints for the next hour when considering a 24-hour horizon. Given that new information is received every hour, a rolling horizon approach is used. This means that the PLF, PV simulation, and HEMS optimization will be updated in order to determine the new operation schedule. A graphical representation of the rolling horizon simulation setting at time t is presented in Figure 4.2.

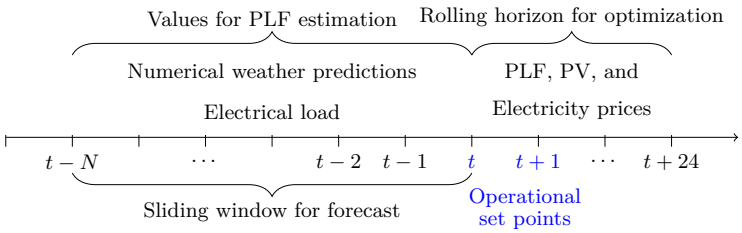

Figure 4.2: Graphical representation of the simulation setting for time t.

The PLF methods presented in Section 3.2 allow different combinations of forecasting and optimization methods. This was considered when simulating the results. The analyzed combinations are presented next for future reference:

- **PI-RH**: perfect information (PI) in a rolling horizon, i.e., using the proposed HEMS optimization, assuming that the consumer's load is known. This method is not applicable in practice, since the PI-RH method assumes a perfect knowledge of the future demand. But it can be used to give performance bounds on the optimization in the other settings.
- **RLS-SP**: the proposed HEMS optimization using 100 scenarios generated by the RLS forecasting method.
- **RLS-E**: the proposed HEMS optimization using the expected value of the 100 scenarios generated by the RLS forecasting method.
- **Copula-SP**: the proposed HEMS optimization using 100 scenarios generated by the Copula forecasting method.
- **Copula-E**: the proposed HEMS optimization using the expected value of the 100 scenarios made by the Copula forecasting method.

Please note that given that the results in Section A.4 indicated that there is no significant difference between modeling the covariance structure using a full covariance model or an AR process, the full covariance model was used because of its implementation simplicity. Finally, January, April, July, and March were selected as the months to be simulated in order to have a comprehensive view of the possible HEMS impact across seasons.

4.3.2 Results

The results section was designed to compare consumers' cost savings when using different control strategies. Two main strategies were considered: a naive controller and an optimization-based controller. A naive controller refers to a consumer with a PV and battery system without a HEMS. This controller maximizes self-consumption by only selling electricity to the grid when the battery is fully charged. The naive controller uses neither forecasting nor optimization methods and it is usually the default controller in the considered HEMS setup. The optimization-based controllers refer to a consumer using a HEMS as described in Appendix B.2.

4.3.2.1 Comparison of different optimization-based control strategies

Simulation was carried out for the optimization-based methods previously introduced. The total cost for all simulated months for all consumers is presented in Table 4.1. The results indicate that for all users the methods which consider the stochasticity of the EL present better performance presenting a smaller difference to the reference method (the perfect information case PI-RH). Furthermore, results on a monthly level were also implemented, showing a similar behavior for each simulated month. Details of the monthly results can be seen in Section B.4.

Consumer	PI-RH		Copula-E		Copula-SP		RLS-E		RLS-SP	
no.	DKK	%	DKK	%	DKK	%	DKK	%	DKK	%
1	5798.61	-	6220.66	7.28	5976.54	3.07	6143.84	5.95	5895.99	1.68
2	12142.80	-	12887.78	6.14	12553.15	3.38	12745.82	4.97	12438.41	2.43
3	6689.17	-	7259.67	8.53	6977.89	4.32	7398.95	10.61	6984.19	4.41
4	2071.31	-	2438.35	17.72	2210.56	6.72	2514.62	21.40	2183.16	5.40
5	1247.45	-	1638.04	31.31	1445.67	15.89	1695.82	35.94	1402.20	12.41
6	7435.47	-	8322.06	11.92	7887.49	6.08	8254.65	11.02	7824.03	5.23
7	9294.46	-	9945.68	7.01	9610.14	3.40	9903.44	6.55	9557.67	2.83
8	3157.25	-	3547.44	12.36	3312.33	4.91	3639.82	15.28	3316.83	5.05
9	6616.65	-	6970.68	5.35	6752.09	2.05	6954.68	5.11	6715.13	1.49

Table 4.1: Total cost of the optimization and forecasting methods. All percentages are calculated relative to the PI-RH method. Please note that smaller the difference to the PI-RH the better the performance of the control strategy.

4.3.2.2 Comparison of naive and optimal control strategies

The results presented in Section 4.3.2.1 indicated the RLS-SP was the best out the optimization-based methods analyzed. This is explained given that RLS methods are more precise on one-step predictions (see Section 3.3.3), which are the predictions that influence the optimization routine the most. The RLS-SP method was selected and compared in detail with a simple naive control strategy. Figure 4.3 shows the RLS-SP cost saving with the cost of the naive strategy as a baseline. Results indicate that in colder months (January and October), the optimization-based controller presents additional savings in comparison with naive controller. However, in warmer months (April and July) the RLS-SP controller performs worst than the simpler naive controller. This indicates that the seasonality of the consumers consumption and prices presents a very significant impact on the performance of the control strategies.

In particular, in winter months higher electricity prices are seen in combination with low PV generation and high EL. Thus, optimization pursues profit mainly by load shifting. While in summer the high PV generation and low EL makes self-consumption a better strategy. A comprehensive analysis of this phenomena is part of the results presented in Appendix B.4.

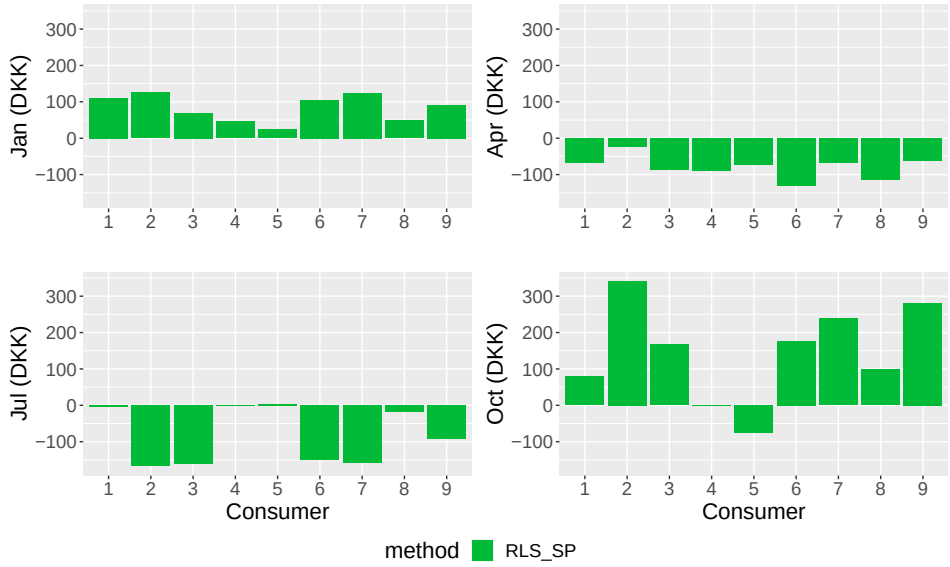

Figure 4.3: Monthly cost saving of the naive controller and RLS-SP. The cost savings results are presented with the naive cost as a baseline. Please note that a positive value indicates that the RLS-SP performs better than the naive strategy.

It was concluded that a combination of a naive and optimization-based (naive+RLS-SP) controller was an overall optimal strategy. Thus, a cost comparison between a consumer without any home hardware, a user using a naive controller, and a combined controller (naive+RLS-SP) was made. The intention is to see the potential economic benefits that consumers may achieve by installing hardware with default behavior, and using a more sophisticated control strategy. Table 4.2 shows the total cost of the simulated months, where the results show that significant cost savings can be achieved in the naive and the proposed control strategy in comparison with a passive consumer. In particular, the results show that consumers with higher EL benefit most from installing the hardware and controllers. Moreover, the differences between the two control strategies show that the combined controller (naive+RLS-SP) provides on average 8.05% additional savings for consumers with higher load (excluding consumers 4 and 5 with the lowest load of all consumers) in comparison with the naive controller, as can be seen in detail in Table 4.3. Please note that although the presented results obey particular conditions (2021 prices and particular users behavior), similar results are expected in the presence of a HEMS setup and similar price structures.

Consumer No.	Passive DKK	%	Naive DKK	%	Naive+RLS-SP DKK	%
1	9219.20	-	6015.12	34.75	5825.17	36.81
2	16734.67	-	12715.96	24.01	12248.38	26.81
3	10799.02	-	6972.54	35.43	6735.15	37.63
4	5353.58	-	2140.78	60.01	2094.73	60.87
5	3901.01	-	1281.02	67.16	1332.69	65.84
6	11208.06	-	7827.33	30.16	7544.91	32.68
7	13370.53	-	9695.89	27.48	9333.01	30.20
8	6455.34	-	3331.51	48.39	3184.64	50.67
9	10806.51	-	6935.86	35.82	6563.42	39.26

Table 4.2: Total cost savings of the four simulated months for passive consumers, naive, and optimal (naive+RLS-SP) control strategies. All percentage values are calculated relative to passive consumers' costs. Note that the higher the cost savings the better the control strategy.

Consumer No.	Naive	Naive+RLS-SP	Difference	%
1	3204.09	3394.03	189.95	5.93
2	4018.70	4486.29	467.59	11.64
3	3826.48	4063.87	237.39	6.20
4	3212.80	3258.85	46.05	1.43
5	2619.99	2568.32	-51.67	-1.97
6	3380.73	3663.15	282.42	8.35
7	3674.64	4037.51	362.87	9.88
8	3123.83	3270.70	146.88	4.70
9	3870.65	4243.09	372.44	9.62

Table 4.3: Naive and optimal control strategies (naive+RLS-SP) total cost savings, and their difference in value and percentage. Note that the higher the cost savings, the better the control strategy.

4.4 Application on bi-objective optimization of HEMSs

In this section, the HEMS single objective problem presented in Section 4.3 is extended to a bi-objective problem, where not only the electricity cost is considered but also the associated CO_2 emissions. The bi-objective formulation follows the general description presented in Section 4.2. Although the explicit formulation was omitted as part of this section, details can be found in Appendix C.3. Furthermore, as described in Section 4.2.3, using the ϵ-constraint method implies having a set of Pareto-front solutions. In real-word applications, one solution out of the Pareto-front set of solutions must be selected to be used as setpoints for the control strategy. One way to select a compromise solution among the Pareto-front set of solutions is the linear programming technique for multidimensional analysis of preference (LINMAP). This method selects the solution from the Pareto-front that yields the lowest normalized distance to the Utopian point, which can be written as

$$\sqrt{\sum_{n=1,\dots,M}\left(\frac{f_n(x_i,y_i)-f_n(x^U,y^U)}{f_n(x^U,y^U)}\right)^2} \leq \sqrt{\sum_{n=1,\dots,M}\left(\frac{f_n(x_j,y_i)-f_n(x^U,y^U)}{f_n(x^U,y^U)}\right)^2} \quad \forall j=1,\dots,I \tag{4.6}$$

where x^U and y^U are the utopian point solutions. The LINMAP metric can be extended to stochastic programming such that the selected solution minimizes the distance between the set of objective values corresponding to the selected solution (x_i, y_i) in scenario $s \in S$ and the discrete distribution of objective values in the Utopian point, i.e. select that solution (x_i, y_i) for which

$$\sqrt{\sum_{n=1,...,M,\, s\in S} \left(\frac{f_{n,s}(x_i, y_i) - f_{n,s}(x^U, y^U)}{f_n(x^U)} \right)^2} \leq \sqrt{\sum_{n=1,...,M,\, s\in S} \left(\frac{f_{n,s}(x_j, y_j) - f_{n,s}(x^U, y^U)}{f_{n,s}(x^U)} \right)^2} \quad \forall j = 1, ..., I \tag{4.7}$$

The stochastic solution selection based on Equation C.5 is referenced as stochastic linear programming technique for multidimensional analysis of preference (S-LINMAP) in the results.

4.4.1 Results

Simulations was carried out for four control configurations: cost minimization, emissions minimization, and the LINMAP and S-LINMAP methods. In order to ease comparability across users and objectives, both levelized costs of electricity (LCOE) and CO_2 emissions have been computed per MWh. We compare the results to cost/emission minimization under perfect information (PI), meaning perfect foresight of the entire optimization horizon. These are theoretical lower bounds on the respective objectives and we refer to them as PI cost and PI emissions.

On average, in Table 4.4 we can see that the PI cost amounts to 165.7 EUR/MWh. The cost minimizing controller exceeds this value on average by 10.9 EUR/MWh ranging from 121.9 EUR/MWh for User 5 to 198.3 EUR/MWh for User 2. When minimizing emissions, the difference to the PI costs increases to 19.6 EUR/MWh on average. As expected, the costs under the bi-objective optimization models are in between the single objective results at 179.1 EUR/MWh with no significant difference between the two solution selection methods LINMAP and S-LINMAP.

User	PI	Bi-objective		Single-objective	
	Cost min.	LINMAP	S-LINMAP	Cost min.	CO2 min.
	[EUR/MWh]	[EUR/MWh]			
1	170.3	181.9	181.8	179.2	188.4
2	190.9	200.6	200.6	198.3	206.0
3	163.0	176.7	176.8	174.3	182.0
4	108.1	127.6	127.5	124.8	134.3
5	96.2	125.0	124.9	121.9	134.4
6	177.2	192.3	192.3	190.4	198.4
7	184.8	196.2	196.1	193.8	201.4
8	135.6	153.8	153.7	150.0	161.4
9	161.2	171.7	171.6	168.7	178.1
Mean	165.7	179.1	179.1	176.6	185.3

Table 4.4: Average LCOE in EUR/MWh across users under perfect (Cost min) and across controller setups (total cost divided by total demand across the four simulated months).

Total emissions results are shown in Table 4.5. While average emissions are highest under a cost-minimizing setup (112.5 kg CO_2-eq/MWh), CO_2 levels are lowest when optimizing for emissions at 107.3 kg CO_2-eq/MWh. Under either of the bi-objective controllers, emissions are slightly higher at 108.0 kg CO_2-eq/MWh. In general, the results show that optimizing solely towards CO_2 minimization implies on average 8.7 EUR/MWh additional cost for a reduction in emissions of 5.2 kg CO_2-eq/MWh, while the bi-objective strategies can reduce emissions by 4.5 kg CO_2-eq/MWh with only an increase in cost of 2.5 EUR/MWh.

User	PI	Bi-objective		Single-objective	
	CO_2 min.	LINMAP	S-LINMAP	Cost min.	CO2 min.
	[kg CO_2-eq/MWh]	[kg CO_2-eq/MWh]			
1	107.5	114.7	114.7	119.5	114.3
2	111.8	118.4	118.3	123.1	117.5
3	95.5	105.1	105.1	109.3	104.3
4	66.2	79.7	79.7	83.6	79.3
5	64.5	83.7	83.6	87.1	84.2
6	105.5	116.0	116.1	120.0	115.3
7	110.3	118.1	118.0	122.7	117.2
8	80.4	93.3	93.3	97.5	92.7
9	92.8	100.0	99.9	105.5	99.0
Mean	98.8	108.0	108.0	112.5	107.3

Table 4.5: Average emissions in kg CO_2-eq/MWh across users under perfect information (CO_2 min) and across controllers setups (total emissions divided by total demand across the four simulated months).

4.5 Application on EV charging

In practical terms, EVs can be characterized as an additional battery in the HEMS setup presented in Section 4.1. Depending on the charger and EV capabilities, the EV can be integrated in such a way that power can flow from the house to the vehicle (grid-2-vehicle) or from the EV to the house grid (vehicle-2-grid).

If we consider that in most residential houses vehicle-2-grid capabilities are not available, and that current market rules do not incentivize actively participation of residential consumers (see Section B.5),

the problem can be reduced to decide when to charge the EV in order to minimize cost or CO_2 emissions. This can be easily done by adding two additional considerations to our HEMS setup i.e., disallow the vehicle-2-grid flow and ensure a specific SoC value within the planning horizon.

While adding an EV can be done in an easy way, the source of complexity may arise from the uncertainty associated to the behavior of the EV user. This can be tackled by using forecast models aiming to predict the EV availability and SoC of the vehicle or by using an interactive approach. In the latter, the use of an app solution such that the user can register his/her availability is required. Unfortunately, EV charge/discharge sessions data was not available during most of the project time span, so this was not part of the simulation results presented in Sections 4.3 and 4.4. Nevertheless, studies have shown the potential that EV may bring [1, 80].

Although the integration of EVs into HEMS was not studied, a way to tackle the problem was explored. A stand-alone prototype was developed to help EV users charge their cars in an optimal way. The prototype will get as inputs from the user the SoC of the vehicle (in a manual or automatic way), the time range where the car will be available to be charged, and a decision parameter that priorities between cost or CO_2 emissions. Subsequently, the algorithm defines the charge schedule for the vehicle relaying on the EV charger to follow it. The algorithm pseudo-code can be seen in Algorithm 2.

Algorithm 2 EV bi-objective charge schedule

Define SoC target
Define time index $t \in \{1, 2, \ldots, T\}$
Rescale electricity cost $c_t \in [0, 1]$ and CO_2 emissions $e_t \in [0, 1]$
Create combined cost value $x_t = \alpha e_t + (1 - \alpha)c_t$
Sort the x_t vector in ascending order and get time stamps t in a new time index ordered set I
Initialize $y_t = 0 \, \forall \, t \in T$
while $i \in I$ or $\sum y_i < SoC$ **do**
    if $SoC - \sum y_i \leq$ max charge **then**
        $y_i = SoC - \sum y_i$
    end if
    if $SoC - \sum y_i >$ max charge **then**
        $y_i =$ max charge
    end if
end while

where $\alpha \in [0, 1]$ is a priority parameter defined by the user and "max charge" is the charger's physical limit. Please note that although the above algorithm is not framed in a classical optimization format, the solution is expected to be optimal. Moreover, the presented algorithm can also be easily integrated into the HEMS setup as an additional known load that has to be satisfied (y_t values). In this way, the EV schedule is considered without increasing the HEMS complexity.

4.5.1 Results

As a proof of concept, an EV charge session from a Watts A/S user is presented in detail in Table 4.6. The table shows a charging session set by the user where his/her availability to charge the car was defined from 2021-12-02 16:00 - 2021-12-03 08:00 (typical office worker availability in Denmark). In this particular case, the user has an EV charger with a physical limit of 11 kWh. Moreover, the electricity price, emissions intensity, and four different charge schedules are presented. The schedules description is:

- **Cost optimal**
  Schedule found using the optimization algorithm with $\alpha = 0$.

- **CO_2 optimal**
  Schedule found using the optimization algorithm with $\alpha = 1$.
- **Balance**
  Schedule found using the optimization algorithm with $\alpha = 0.5$
- **Naive**
  Corresponds to the user connecting the EV when he/she arrives home. This schedule is assumed as the default behavior users have and is implemented for comparison purposes.

As expected, the schedules shown in 4.6 differ significantly between the cost and CO_2 optimal schedules, with the balance schedule presenting a similar solution than the CO_2 optimal one.

Time	Price [DKK/kWh]	Emissions [g CO_2 eq/kWh]	Naive Shcedule [kWh]	Cost opt ($\alpha = 0$) Shcedule [kWh]	CO_2 opt ($\alpha = 1$) Shcedule [kWh]	Balance ($\alpha = 0.5$) Shcedule [kWh]
16.00 – 17.00	5.182	280.630	11	0	0	0
17.00 – 18.00	4.736	240.421	11	0	0	0
18.00 – 19.00	3.996	235.437	8	0	0	0
19.00 – 20.00	3.164	209.344	0	0	11	11
20.00 – 21.00	3.016	233.309	0	0	8	0
21.00 – 22.00	2.925	226.590	0	0	11	11
22.00 – 23.00	2.745	246.609	0	0	0	8
23.00 – 00.00	2.877	255.076	0	0	0	0
00.00 – 01.00	2.726	311.290	0	0	0	0
01.00 – 02.00	2.661	342.607	0	8	0	0
02.00 – 03.00	2.599	328.785	0	11	0	0
03.00 – 04.00	2.633	288.928	0	11	0	0
04.00 – 05.00	2.779	271.554	0	0	0	0
05.00 – 06.00	3.144	276.295	0	0	0	0
06.00 – 07.00	3.871	257.056	0	0	0	0
07.00 – 08.00	3.949	281.581	0	0	0	0

Table 4.6: EV charge schedules found using Algorithm 2 for three different α values.

Moreover, the total cost and CO_2 emissions given the different schedules are presented in Table 4.7. We can clearly see that the cost-optimal schedule presents significant differences in comparison with all other schedules. Optimization for cost comes with a higher CO_2 cost, with the cost-optimal polluting 43.14% more than the emissions optimal. Moreover, the CO_2 optimal schedule is 15.55% more expensive than the cost optimal, with the balance schedule presenting a good compromise. The balance schedule is 12.81% more expensive than the cost optimal but with only 1.60% more emissions. Finally, the naive schedule presents poor performance in both objectives, polluting 14.31% more and being 78.93% more expensive than its optimal counterparts.

Schedule type	Total cost [DKK]	Δ Cost opt [DKK]	Δ Cost opt [%]	Emissions [g CO_2 eq]	Δ CO_2 opt [g CO_2 eq]	Δ CO_2 opt [%]
Naive	141.07	62.23	78.93	7615.06	953.31	14.31
Cost opt	78.84	0.00	0.00	9535.70	2873.95	43.14
CO_2 opt	91.10	12.26	15.55	6661.75	0.00	0.00
Balance	88.94	10.10	12.81	6768.15	106.40	1.60

Table 4.7: Total cost and CO_2 emissions for the different strategies.

4.6 Discussion

All the data elements and forecast methods introduced in previous chapters were successfully combined and integrated into the HEMS control strategy and simulations presented in the application sections of this chapter. Key findings in the results are summarized as:

- The economic evaluation presented in Section 4.3 validates the performance of the probabilistic forecasting modelling methodology. In detail, cost savings depend on the quality of the forecast, with results showing a small gap between the perfect information cases and the optimization-based strategies using the forecast framework for all users.
- Considering the stochasticity of the EL has a significant impact on the control strategies performance. This is clear from the results where in all cases the stochastic optimization models outperform the expected value ones.
- Price and market rules shape the possibilities of control strategies. Results indicate that in the presence of high PV generation and low EL, there is no need for sophisticated control strategies, with self-consumption strategies being effective. This is a consequence of the gap between sale and purchase prices that residential users have access to given current market conditions. Moreover, one could argue that research oriented to new business models, and regulatory and price frameworks is needed in a higher degree than technical solutions.
- Under current market rules and HEMS setup, consumers can find value in each step of the way. Results indicated that there is significant economic incentive for the user to install a PV and battery system. Furthermore, adding the decision support tools presented increases the benefits for the consumers, with more benefits expected in energy systems with higher share of wind and solar.
- There are significant differences between cost and CO_2 emission optimization. This implies that full information about emissions is not completely contained in the electricity price signal, which opens the door for bi-objective frameworks or highlights a need for regulatory change if climate change goals are expected to be made.

Although overall results validate the developed technical tools and support their future implementation as part of HEMSs, there are critique points that we must consider:

- Other control frameworks need to be explored. There are control frameworks such as Model Predictive Control and dynamic programming, used in the literature that could be implemented.
- The stochasticity associated with PV generation must be included in the optimization-based control strategy. As seen with the EL, it is vital for the decision-making process to consider uncertainty source. In this case, access to data limited the possibility to build adequate PV forecast models, which allows to have more insights on the PV probability distribution. This could be tackled by applying the modelling methodology presented in Chapter 3. PV scenarios can be generated and included in MILP formulations in a similar fashion to the EL ones.
- While it is expected that including EVs under the HEMS will improve consumers' economic and environmental benefits, realistic simulation results were not made. This is a clear step in future research.
- Extensions to the control strategies are needed for real-time implementations. The presented control strategies yield, as results, target set points at hourly time resolution. In a physical implementation, deviations from these set points might occur at an intra-hour time resolution, which would require real-time actions. This is something not included in the current research.

- Given that the research was based on simulation setups, there is a need for field test implementations that validate the results. Additionally, scaling up the studies to many consumers is needed to be able to generalize the effect.
- Clustering methods in combination with the presented simulation framework in order to identify which consumer could benefit the most and other insights.

CHAPTER 5

Conclusion

5.1 Contributions

In the course of the research described in this thesis, the expected role of HEMSs in the Danish energy system was described. The description allows us to have a clear context for future implementations. Key elements such as data access, time resolution, price structure, and market rules, shaped the HEMS architecture followed during the research. Although the followed HEMS setup was simplified in comparison to a full architecture, important technical challenges were identified. These challenges were then used as an initial point for technical research. While this part of the research may look trivial, it is key to ensure that developed algorithms respond to real needs, thus, by contributing to a good contextualization of the system, the probability of a future implementation increases.

The need for reliable information about consumers future needs sparked forecast research. The research was designed for the prediction of the EL needs at a residential level, which was proven to be a complex task given the EL characteristics. The presented multivariate probabilistic forecasting methodology allows to combine the information present in exogenous variables and the TS itself to produce a good estimation of the marginal distributions that compose a multivariate forecast. Furthermore, by modelling the temporal dependency in a different modelling stage, we gain in modularity. This means that different modules could be improved separately, which could help future development. Moreover, results (see Section 3.3.3) validated the performance of the modelling technique, indicating that modelling the temporal dependency had a significant impact on the out-of-sample performance metrics. Finally, although the results come from a particular application, this research contributes with a generic formulation that facilitates its application in other domains.

The knowledge gained by the contextualization of HEMSs in the Danish market and the forecast methodology was combined and used in the research of control strategies suitable for HEMSs. Results indicated that there are significant incentives for users to install a HEMS setup and the presented control strategies. In particular, including the stochasticity inherent to the EL has a significant impact on the performance of the control strategies, with the stochastic optimization methods outperforming their deterministic counterparts for every simulated user. Moreover, it was found that seasonality and the prices available to residential consumers have a profound impact on the performance of the control strategies, with self-consumption strategies outperforming the optimization-based controllers in high PV generation and low price volatility conditions. Although the simulation study presents a small sample size, the research contributes insights on how this type of systems might perform on a real implementation and that more sophisticated control strategies might be limited by the current market rules. This also means that more complex HEMS architectures should be evaluated in light of the more simple approaches, as higher complexity is not a warranty of higher benefits for the consumers. However, in a future scenario with a high share of RESs, prices are foreseen to be much more volatile and, therefore, the proposed advanced forecasting and optimization are likely to be more important than in the studied current conditions.

An analysis made on the relation of electricity prices and CO_2 emissions showed that electricity prices do not contain complete information on CO_2 emissions. This means that consuming electricity when it is cheaper does not necessarily imply lower CO_2 emissions. The fact that this behavior was present in the data, opened the door to bi-objective stochastic control strategies. The research on this topic showed that considering both objectives yielded solutions that have a significant impact on the CO_2 footprint at a relatively small increase in costs. Thus, this research contributes by highlighting a potential deficiency in how electricity prices are created (at least from an environmental perspective) while presenting a possible way to address the problem.

It is common knowledge the potential that EVs can bring to the energy system as well to the final users. And although data access was a limiting factor in this project, the proof of concept presented in Section 4.5 shows the potential of these types of assets. Moreover, while it can be considered easy to add an EV to the presented HEMS setup, given that it is battery, a stand alone solution for the vehicle charge process was needed. The results show that significant cost savings could be achieve just by controlling the charge schedule, with very significant CO_2 reductions when emissions are also considered. Thus, this research contributes with a prototype solution suitable for users without the full HEMS technology stack, which is important if we considered that not all consumers have the means and/or the motivation to invest in HEMSs and EVs at the same time.

5.2 Perspectives and opportunities for further research

Many future research possibilities can be derived from the current work. One important aspect is the user interface and the impact that additional information might have on consumers. As it was shown in [74], there is a positive response from the consumer side when they are exposed to information feedback via the Watts App or, in that particular case, a game. User interfaces are expected to play a very significant role as more elements such as smart appliances continue their roll-out. Specially, if we consider that their inclusion in the HEMS decision-making process increases the system's complexity.

In terms of the forecast capabilities of the system, access to sensing devices at the home side will bring additional information about the EL. This might imply the disaggregation of the load data. Disaggregated loads open the possibility to tailor individual models for elements such as heat pumps, heaters, controllable loads as well as the uncontrollable loads. In this regard, a balance between a higher technical approach (more models) and their benefits must be kept. Regime changes are also a key element which was not tackled in this research. Additional sensing devices and user interfaces should bring enough information to detect anomalies in the household behavior such as poor heat insulation, energy leaks or the absence of the household inhabitants. This should trigger changes in the forecast algorithms in order to rapidly adapt to such changes and alert customers.

Although the results of the stochastic control strategies indicate a good performance of the algorithms, a large scale test covering a wider range of users it still needed, and a clear path for future research. In this regard, clustering techniques could be applied to identify users whom a HEMS will bring more value along with additional insights that could be useful for Watts. An interesting application of such insights is aggregation. Watts could harvest data coming from a large simulation setting to experiment with new business models including aggregators, as the individual behavior could be governed by the independent HEMSs at the users' side.

The research presented in this thesis is mainly derived from historic data hence conditions in the historic period. Results must be updated as the energy system changes. This is expected to support the arguments made in favor of HEMSs and similar technologies. However, the full development of a real-world application is currently underway at Watts. The research made has helped to define the design of the back end and data services that are expected to be accessible to customers in late 2022.

Moreover, the presented tools are currently being developed as software libraries and partnerships with hardware providers are also underway in order to develop custom products able to articulate with the software solutions. It is clear that field tests and extensive trials are needed before the roll out of a final product. However, this research provides the basic building blocks of a smart HEMS product able to help with European energy transition as well as reducing CO_2 emissions.

Part II

Publications

PAPER A

Probabilistic load forecasting considering temporal correlation: Online models for the prediction of households' electrical load

Authors:
Julian Lemos-Vinasco, Peder Bacher, Jan K. Møller

Published in:
Applied Energy, Volume 303, 1 December 2021, 117594

Doi:
https://doi.org/10.1016/j.apenergy.2021.117594

Probabilistic load forecasting considering temporal correlation: online models for the prediction of households' electrical load

Julian Lemos-Vinasco[1] [2], Peder Bacher[1], Jan K. Møller[1]

Abstract

HEMSs are expected to become an inevitable part of the future smart grid technologies. To work effectively, HEMSs require reliable and accurate load forecasts. In this paper, two new modelling methods are presented. They are both suited for producing multivariate probabilistic forecasts, which consider the temporal correlation between forecast horizons. The first method employs point forecasts generated with RLS models and subsequently analyses the forecasts' residuals to estimate the marginal distributions and temporal correlation. The second method is based on quantile regression to estimate marginal distributions, and a Gaussian copula for linking them together. Furthermore, the application of two modelling approaches for the temporal correlation estimation are investigated for each of the two modelling methods. As a case study, a numerical experiment is designed to emulate an online HEMS operation using data from an inhabited home located in Denmark. Simulation results show a robust performance for the proposed models, with the quantile-copula ensemble outperforming the RLS-based models in predicting the marginal distributions and capturing the temporal correlation.

A.1 Introduction

With the aim of helping the integration of RESs, improving energy markets, and allowing consumers to better regulate their energy consumption, smart grids are being developed especially in Europe [19]. As part of the smart grid technologies, HEMSs are expected to play a key role activating price and/or CO_2 demand response. The implementation of smart HEMSs has several technical challenges such as the prediction of the home electricity demand, i.e. electrical LF. Forecasting is essential for the optimal planning and procurement of electricity in homes, in order to maximize the economical and/or environmental benefit without compromising users' comfort [54].

At a home level, LF presents several challenges given its multi-seasonality, non-stationarity, and stochastic characteristics, which makes it an interesting research topic [27]. Additionally, HEMSs require load forecast values for several future time points, ranging from hours to days ahead in rather high time resolution. Many different statistical and artificial intelligence techniques have been applied for LF in recent times. Modelling techniques such as multiple linear regression [11, 102, 117], TS [39, 17], artificial neural networks (ANNs) [100, 4], and stochastic differential equations (SDEs) [132, 87] are suitable for the task. These modelling techniques are often applied to an aggregated load level, use exogenous variables, e.g. NWPs, estimate the expected value of the process i.e point forecast, and make a general assumption about the process distribution.

Given the complexity of the EL, having a point forecast is often not optimal and more details about the multivariate distribution are needed for the optimal HEMS decision making process – especially, at an individual residential level. Recent studies have therefore applied probabilistic forecasting models. Modelling techniques such as quantile regression [63, 75], kernel density estimation [126, 130], long

[1]Technical University of Denmark, Department of Applied Mathematics and Computer Science, Lyngby, Denmark
[2]Watts S/A, Section of Research and Development, 36 Main Street, Svinninge, Denmark

short-term memory artificial neural networks (LSTM-ANNs) [120], density-estimating ANNs [118], and quantile regression neural networks (QRNNs) [51, 130] have been used. The above-mentioned studies are proven to yield accurate results for individual forecast horizons. However, they do not address the temporal correlation between the forecast horizons.

From the above studies, the need for PLF models, which consider the correlation in a multivariate setting for individual households, was identified. Moreover, while the idea of modelling the temporal correlation in PLF for individual households seems to be rarely explored, it does not mean that this problem has not been studied in other fields, e.g. wind forecasting and weather forecasting. In these fields, different techniques have been used to produce multivariate probabilistic models considering temporal correlation. Copula-based modelling techniques such as ensemble Bayesian Model Averaging (BMA) and copula [83], quantile regression and copula [97, 111], and kernel-based support vector quantile regression and copula [52] have been applied. Other techniques suitable for the task such as maximum likelihood estimation for conditional AR models [67], and multivariate conditional parametric models [110] have also shown positive results. Among the existing techniques, the copula-based methods stand out due to their modularity. In these approaches, the probabilistic forecast for individual time horizons is separated from the modelling of the temporal correlation between distributions. This allows the use of robust modelling techniques for the estimation of marginal distributions which are then linked, using a copula, to a posterior modelling stage where the temporal correlation is addressed.

This paper contributes by introducing and analysing different approaches to temporal forecasting, i.e. forecasting where the temporal correlation is taken into account. Thus, these approaches allow the generation of reliable scenarios forecasts (ensembles), which are needed for implementing stochastic control algorithms; in the present case, applied to forecasting of electricity load of a single residential household. The suggested approaches are formed by two stages: in the first stage, a multivariate PLF model is applied, and in a second stage, the temporal correlation is modelled. Two different multivariate methods are applied for the first stage: the RLS and the quantile-copula. The RLS is inspired by the load forecast methods by [11, 102, 117] and adapted in the present work to model electricity load instead of heat load. The quantile-copula approach is a combination of methods presented for wind power forecasting in [16] and [110]. In the second, stage the temporal correlation is modelled using the two different methods: the "free" where the correlations are modelled by calculating the full cross-correlation matrix between all horizons (it is free of a model specification). And the "AR" where the correlations are modelled by the multivariate k-step ahead predictive covariance of auto-regressive (AR) models derived for the particular cases at hand. All four combinations of the two methods in each stage are applied and compared to a simple reference model.

Thus, a total of five complete models are presented. The first model is referenced as "RLS-free" and uses RLS models for the prediction of the load expected value with the temporal correlation being estimated directly from a full covariance matrix model. The second model is referenced as "RLS-AR" indicating the use of RLS models with an AR model for the temporal correlation. The third and fourth models are referenced as "Copula-free" and "Copula-AR" and indicate the use of quantile regression, a Gaussian copula. The reference model is without any correlation structure method and is named "RLS" was implemented for comparison purposes. The different models are tested by forecasting the EL of an individual household located in Denmark in an online simulation setting i.e. rolling forecast. In order to evaluate the performance of the different models, the PICP, PINAW, pinball loss and CRPS were used as adequate metrics to evaluate the performance of the individual PLF horizons. Moreover, the VarS was used as the metric to evaluate the performance of temporal correlation modelling. The results indicate a significant performance improvement in comparison to the reference model using the models which consider temporal correlation. Furthermore, the results also indicate that the models from quantile-copula method produce superior results.

The suggested approaches can be applied in a wide range of forecast settings, where temporal correlation must be taken into account and a quite generic description is included in the text – we hope this

will make it rather easy for others to implement and use for other forecasting applications.

The paper starts by presenting the mathematical details of the implemented models in Section A.2. Next, a description of the data and the simulation setting used for testing the models is presented in Section A.3. The Results are presented in Section A.4, which include a verification of the models' assumptions, details on the temporal correlation structures, and metric comparisons. Finally, a discussion of the findings and perspectives for future work is outlined in Section A.5 with the paper's conclusion being presented in Section A.6.

A.2 Methods and models

A.2.1 Notation

First, the general notation to be used throughout the text is introduced. Notation follows [76] with the following modifications. Let $\boldsymbol{X}_t = (X_{t,1}, X_{t,2}, \ldots, X_{t,K})^\intercal$ denote a k-dimensional random variable at time t for future values $t+1, t+2, \ldots, t+k$. Moreover, upper case letters are used for random variables while lowercase letters denote the corresponding observations. Furthermore, vectors and matrices are emphasised using bold font. Thus, $\boldsymbol{x}_t = (x_{t,1}, x_{t,2}, \ldots, x_{t,K})^\intercal$ is used for the realisations of the random vector $\boldsymbol{X}_t$. Random variables are assigned to letters from the last part of the alphabet, while deterministic terms are assigned to letters from the first part of the alphabet.

A.2.2 Methods and Models

Let Y_t be a univariate TS. The forecast of Y_t can be expressed as multivariate random variable at each time t for future lead times $t+k$ as

$$\boldsymbol{Y}_t = (Y_{t,1}, Y_{t,2}, \ldots, Y_{t,K})^\intercal \tag{A.1}$$

where we use the previously introduced notation s.t.

$$Y_{t,k} = Y_{t+k} \quad \forall\, k \in \{1, 2, \ldots, K\} \tag{A.2}$$

Our interest is to forecast the PDF for $\boldsymbol{Y}_t$ denoted by $f_t(\boldsymbol{y})$. Proposing a functional form for f_t implies a simultaneous description of both the marginal densities (for each horizon k) as well as the temporal correlation [111]. In this study, two approaches are proposed to estimate f_t. The first approach uses RLS models for the estimation of the expected value $\widehat{\boldsymbol{Y}}_t$ and then analyses the models' residuals in order to obtain an estimation of the marginal densities as well as the temporal correlation of $\boldsymbol{Y}_t$. The second approach uses quantile regression models for the prediction of the marginal distributions of $\boldsymbol{Y}_t$, and a Gaussian copula to address the temporal correlation. In both approaches, two methods were considered to model the correlation structure. The first method estimates the correlation structure from the data with a full covariance matrix model, while the second method uses an AR model – in which the covariance matrix is parametrised with only a few parameters (order of the AR process + 1). Figure A.1 presents a graphical summary of the modelling methodology process.

Different score metrics were considered to evaluate the performance of the proposed models. While the PICP, PINAW, pinball loss, and CRPS evaluate the accuracy of the estimated marginal distributions, the VarS measures the accuracy of the temporal correlation. The mathematical details of all of the implemented methods and models are presented in the following subsections.

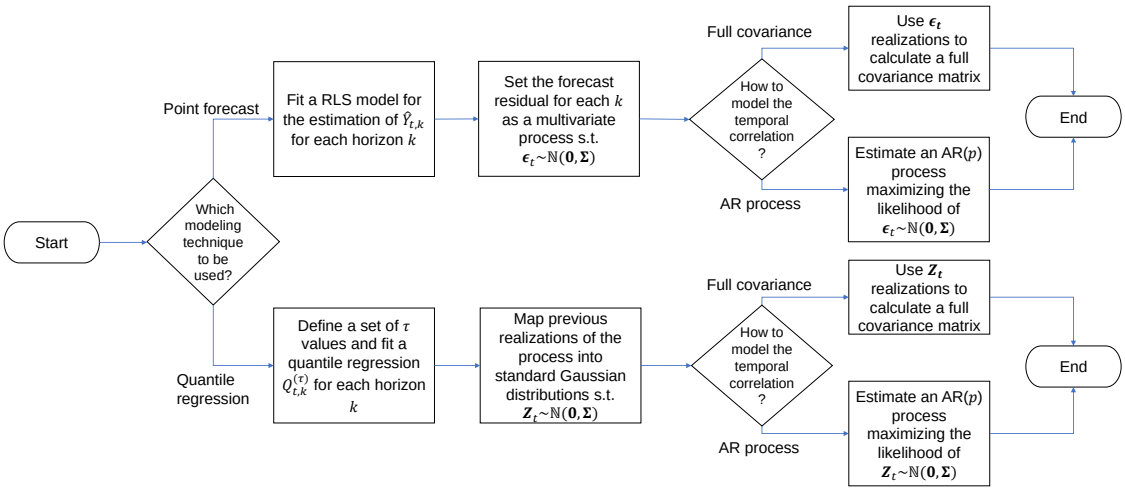

Figure A.1: Process graph summarising the modelling methodology.

A.2.2.1 Recursive Least Squares (RLS)

Let us assume that the elements of $\boldsymbol{Y}_t$ are independent, i.e. independence between the forecast horizons. We can then use the following parametric models to describe each of them individually

$$Y_{t,k} = \boldsymbol{x}_{t,k}^{\intercal} \boldsymbol{\theta}_{t,k} + \varepsilon_{t,k} \tag{A.3}$$

$$\widehat{Y}_{t,k} = \mathbb{E}\left\{Y_{t,k} | \boldsymbol{x}_{t,k}\right\} = \boldsymbol{x}_{t,k}^{\intercal} \boldsymbol{\theta}_{t,k} \tag{A.4}$$

where $\boldsymbol{x}_{t,k}$ is a known vector of independent variables. In order to consider possible non-stationary characteristics of the TS, the RLS model with a forgetting factor can be used for the estimation of the parameters $\widehat{\boldsymbol{\theta}}_{t,k}$. For further details on the RLS method see [76]. For the errors $\boldsymbol{\varepsilon}_t = (\varepsilon_{t,1}, \ldots, \varepsilon_{t,k})^{\intercal}$, we may assume that $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ if the errors are well-behaved. Thus, the PDF of $\boldsymbol{Y}_t$ will be given by

$$\boldsymbol{Y}_t \sim \mathcal{N}(\widehat{\boldsymbol{Y}}_t, \mathbf{\Sigma}) \tag{A.5}$$

The problem is then reduced to the estimation of the covariance matrix $\mathbf{\Sigma}$. This can be done empirically from data (using previous realisations of the errors) or exploiting the fact that $\boldsymbol{\varepsilon}_t$ is a TS itself by using models suitable for the task. Section A.2.2.3 presents the estimation of $\mathbf{\Sigma}$ using an AR model for $\boldsymbol{\varepsilon}_t$. Although in the above formulation the temporal correlation is given by the time-invariant matrix $\mathbf{\Sigma}$, a sliding window approach is considered in the implementation as presented in Section A.3.2.

A.2.2.2 Quantile regression with a Gaussian copula (quantile-copula)

Let $F_{t,k}(y) = P(Y_{t,k} \leq y)$ be the cumulative density function (CDF) of the random variable $Y_{t,k}$. When conditioned on $\boldsymbol{x}_{t,k}$ it can be described by a quantile regression model

$$Q^{(\tau)}_{t,k} = F^{-1}_{t,k}(\tau) = \inf\{y_k : F_{t,k}(y) \geq \tau\} \tag{A.6}$$

$$\{Q^{(\tau)}_{t,k} | \boldsymbol{x}_{t,k}\} = \boldsymbol{x}_{t,k}^{\intercal} \boldsymbol{\beta}_k^{(\tau)} \tag{A.7}$$

where $\tau \in (0, 1)$. Now let $\rho_\tau(r)$ be the loss function

$$\rho_\tau(r) = r(\tau - \mathbf{1}_{r<0}(r)) \tag{A.8}$$

such that the estimation of the parameters $\widehat{\boldsymbol{\beta}}_k^{(\tau)}$ can be achieved by solving the minimisation problem

$$\widehat{\boldsymbol{\beta}}_k^{(\tau)} = \arg\min_{\boldsymbol{\beta}_k^{(\tau)}} \sum_{n=1}^{N} \rho_\tau(y_{t-n,k} - \boldsymbol{x}_{t-n,k}^\intercal \boldsymbol{\beta}_k^{(\tau)}) \tag{A.9}$$

A quantile forecast $\widehat{Q}_{t,k}^{(\tau)}$ with nominal proportion τ is an estimate of $Q_{t,k}^{(\tau)}$ calculated at time t for a future time $t+k$. More details on quantile regression may be found in [66]. Finally, a forecast of the PDF of the variable of interest can be produced by gathering a set of m quantile forecasts [97]

$$\widehat{\boldsymbol{f}}_{t,k} = \{\widehat{Q}_{t,k}^{(\tau_i)} | \boldsymbol{x}_{t,k},\, 0 \leq \tau_1 < \cdots < \tau_i < \cdots < \tau_m \leq 1\} \tag{A.10}$$

The above approach allows to have a model for the CDF of each random variable $Y_{t,k}$. Now let F_t be a multivariate CDF describing the distribution of the random vector $\boldsymbol{Y}_t$ given by

$$F_t(\boldsymbol{y}) = P(Y_{t,1} \leq y_1, Y_{t,2} \leq y_2, \ldots, Y_{t,K} \leq y_K) \tag{A.11}$$

In most cases there is not an obvious distribution F_t, so instead a copula approach is applied. The copula allows to decompose the problem of estimating F_t into two parts. First, marginal predictive cumulated densities $F_{t,k} = P(Y_{t,k} \leq y)$ for each horizon are obtained using quantile regression, thus describing the random variables $Y_{t,k}$. Then, the marginal distributions are linked together to obtain F_t using a copula function. The mathematical foundation of copulas is given by Sklar's [109] theorem

$$F_t(\boldsymbol{y}) = \boldsymbol{C}(F_{t,1}(y_1), F_{t,2}(y_2), \ldots, F_{t,K}(y_K)) \tag{A.12}$$

Now let $\boldsymbol{C}$ be a function that maps $\boldsymbol{y}$ into a multivariate Gaussian distribution with zero mean, unit marginal variances and covariance matrix $\boldsymbol{\Sigma}$, and let Φ^{-1} be the inverse of a univariate standard normal CDF. A Gaussian copula is given by

$$\boldsymbol{Z}_t = \left[\Phi^{-1}(F_{t,1}(y_{t,1})), \Phi^{-1}(F_{t,2}(y_{t,2})), \ldots, \Phi^{-1}(F_{t,K}(y_{t,K}))\right]^\intercal \tag{A.13}$$

$$\boldsymbol{Z}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{A.14}$$

Note that if the marginal distributions are properly calibrated, then the random variable $\Phi^{-1}(F_{t,k}(y)) \sim \mathcal{N}(0, 1^2)$. The above formulation implies that the joint multivariate predictive density for $\boldsymbol{Y}_t$ can be represented by a latent multivariate Gaussian process. Furthermore, note that even though $F_{t,k}$ as well as F_t are time-dependent, the underlying dependence structure is characterised by the time-invariant correlation matrix $\boldsymbol{\Sigma}$. It is out of the scope of this article to study time-dependent correlation structures. At this point, the problem is reduced to the estimation of the covariance matrix $\boldsymbol{\Sigma}$. This can be done empirically using previous realisations of $\boldsymbol{Z}_t$ or using the method described in Section A.2.2.3. Finally, possible realisations of the random vector $\boldsymbol{Y}_t = \boldsymbol{y}_t$ can be drawn easily by calculating the inverse of Equation (A.13) given a realisation of the latent Gaussian process $\boldsymbol{Z}_t$ thus

$$\boldsymbol{y}_t = \left[F_{t,1}^{-1}(\Phi(z_1)), F_{t,2}^{-1}(\Phi(z_2)), \ldots, F_{t,K}^{-1}(\Phi(z_K))\right]^\intercal \tag{A.15}$$

A.2.2.3 Covariance matrix from an Autoregressive AR(p) process

Let $\boldsymbol{Z}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ be a multivariate random variable with its elements given by an AR(p) process of the form

$$Z_{t,k} = \sum_{i=1}^{p} z_{t,k-i}\alpha_i + \lambda_{t,k} \tag{A.16}$$

$$\lambda_{t,k} \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d} \tag{A.17}$$

The above formulation allows us to express entries of $\boldsymbol{\Sigma}$ as a function of σ^2 and α_i parameters of the AR process. Equations (A.18) to (A.21) show how to calculate the variance and covariance elements of $\boldsymbol{\Sigma}$

$$\begin{aligned}\mathrm{Var}\{Z_{t,k+l}|Z_{t,k}\} = \sigma^2 + \sum_{i=1}^{l-1} \alpha_i^2 \,\mathrm{Var}\{Z_{t,k+l-i}\} + \\ 2\sum_{i=1}^{l-2}\sum_{j=i+1}^{l-1} \alpha_i\alpha_j \,\mathrm{Cov}\{Z_{t,k+l-i}, Z_{t,k+l-j}\} \quad \forall l \leq p+1\end{aligned} \tag{A.18}$$

$$\begin{aligned}\mathrm{Var}\{Z_{t,k+l}|Z_{t,k}\} = \sigma^2 + \sum_{i=1}^{p} \alpha_i^2 \,\mathrm{Var}\{Z_{t,k+p+1-i}\} + \\ 2\sum_{i=1}^{p-1}\sum_{j=i+1}^{p} \alpha_i\alpha_j \,\mathrm{Cov}\{Z_{t,k+p+1-i}, Z_{t,k+p+1-j}\} \quad \forall l > p+1\end{aligned} \tag{A.19}$$

$$\mathrm{Cov}\{Z_{t,k+j}, Z_{t,k}\} = \sum_{i=1}^{j} \alpha_i \,\mathrm{Cov}\{Z_{t,k+j-i}, Z_{t,k}\} \quad \forall j \leq p \tag{A.20}$$

$$\mathrm{Cov}\{Z_{t,k+j}, Z_{t,k}\} = \sum_{i=1}^{p} \alpha_i \,\mathrm{Cov}\{Z_{t,k+j-i}, Z_{t,k}\} \quad \forall j > p \tag{A.21}$$

With the above framework in place, σ and α_i can be estimated by maximising the log-likelihood function of $\boldsymbol{Z}_t$ given by

$$\underset{\sigma,\,\alpha_i}{\arg\max} \sum_{n=1}^{N} -\frac{1}{2}\left[\ln(|\boldsymbol{\Sigma}|) + \boldsymbol{z}_{t-n}^{\intercal}\boldsymbol{\Sigma}^{-1}\boldsymbol{z}_{t-n} + k\ln(2\pi)\right] \tag{A.22}$$

A.2.3 Performance metrics

A.2.3.1 Prediction Interval Coverage Probability score (PICP)

The reliability of a probabilistic forecast can be measured using the PICP score [107] given by

$$\mathrm{PICP}(y_{t,k}, PI_{t,k}^{(\alpha)}) = \frac{1}{N}\sum_{t=1}^{N} c_{t,k} \tag{A.23}$$

where

$$c_{t,k} = \begin{cases} 1, & y_{t,k} \in PI_{t,k}^{(\alpha)} \\ 0, & y_{t,k} \notin PI_{t,k}^{(\alpha)} \end{cases} \tag{A.24}$$

and $PI_{t,k}^{(\alpha)}$ is a confidence interval with significance level α. The outcome of the score indicates the probability of a realisation $y_{t,k}$ to fall in a predicted interval.

A.2.3.2 Prediction Interval Normalized Average Width score (PINAW)

A useful score to evaluate the sharpness of a probabilistic forecast is the PINAW score given by

$$\text{PINAW}(PI_{t,k}^{(\alpha)}, A) = \frac{1}{NA}\sum_{t=1}^{N} PI_{t,k}^{(\alpha)} \tag{A.25}$$

where $P_{t,k}I^{(\alpha)}$ is a prediction interval with α probability and A is the range of the target $Y_{t,k}$. The score indicates the average proportion of the prediction interval width with respect to the target variable's range.

A.2.3.3 Pinball score

A very common way of evaluating a probabilistic forecast is using the pinball loss function given by

$$L(y_{t,k}, \hat{Q}_{t,k}^{(\tau)}, \tau) = \begin{cases} (y_{t,k} - \hat{Q}_{t,k}^{(\tau)})\tau & y_{t,k} \geq \hat{Q}_{t,k}^{(\tau)} \\ (\hat{Q}_{t,k}^{(\tau)} - y_{t,k})(1-\tau) & y_{t,k} < \hat{Q}_{t,k}^{(\tau)} \end{cases} \tag{A.26}$$

where $\hat{Q}_{t,k}^{(\tau)}$ is the quantile value associated to a level $\tau \in (0, 1)$.

A.2.3.4 Continuous Ranked Probability Score (CRPS)

Given a probabilistic forecast taking the form of the CDF $F_{t,k}$, we can apply the CRPS [78] given by

$$\text{CRPS}(F_{t,k}, y_{t,k}) = \int_{-\infty}^{\infty} (F_{t,k}(u) - \mathbf{1}_{u \geq y_{t,k}}(y_{t,k}))^2 du \tag{A.27}$$

This score measures the likelihood of an observation $y_{t,k}$ belonging to $F_{t,k}$. A drawback of all the previous presented scores, see [104], is that the CRPS does not include any dependencies between horizons k, hence it cannot be used to measure the temporal correlation structure.

A.2.3.5 Variogram score (VarS)

To effectively measure the representation of the temporal correlation structure in the predicted values, we can use VarS of order p

$$\text{VarS}_p(f_t, \boldsymbol{y}_t) = \sum_{i=1}^{K-1}\sum_{j=i+1}^{K} w_{i,j}(|y_{t,i} - y_{t,j}|^p - \mathbb{E}\{|Y_{t,i} - Y_{t,j}|^p\})^2 \tag{A.28}$$

with

$$w_{i,j} = \frac{1}{j-i} \tag{A.29}$$

The selection of the order $p = 0.5$, and the weight term $w_{i,j}$ was done as recommended in [104].

A.3 Data and simulation study

A.3.1 Data description

The dependent variable of interest is the average EL (kWh) of a residential home from the period of 2017-06-01 to 2018-05-31 in hourly time resolution. The data comes from a smart meter and there was no access to further details besides the home's heating type and number of inhabitants. Moreover, only one year of data was considered due to the inhabitants moving out of the home from 2018-06-15 and the use of a non-electrical heating type previous to 2017-06-01. The TS plot and data distribution over the mentioned time period can be seen in Figure A.2. From the plot we can identify characteristics of a non-stationary TS. It is known that the home is heated up during winter with a heat pump, which explains higher EL during the winter season in comparison with the summer season. Moreover, it is also known that the home is not cooled during summer. The EL distribution shows a natural lower bound of zero with some high values, which clearly makes the distribution right skewed. Additionally, given that the data comes from a real inhabited residential home, time periods of near-zero load are seen. These low consumption periods can be attributed to absence of the house's inhabitants due to holidays or similar events.

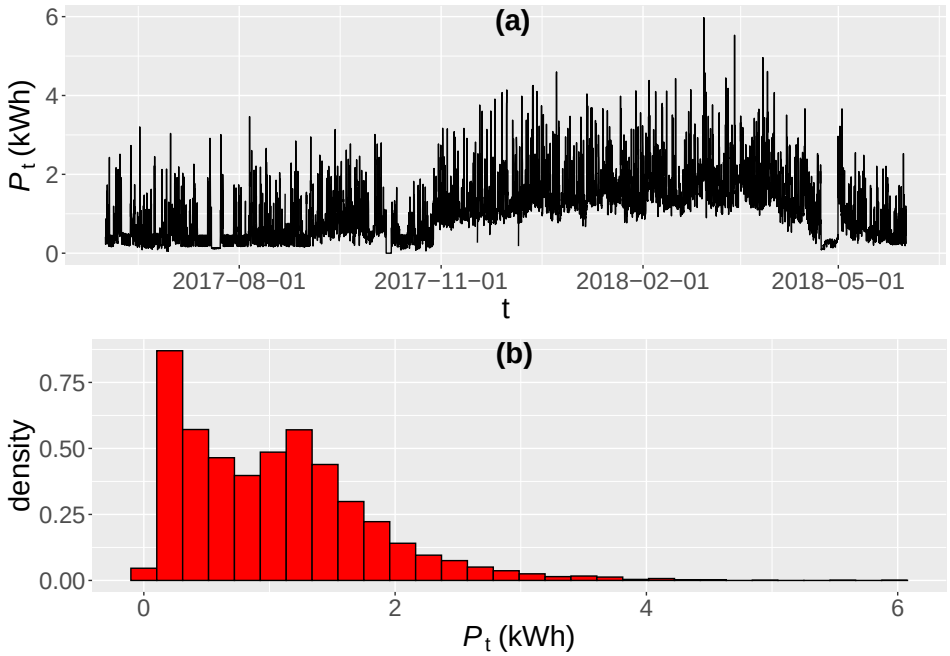

Figure A.2: Electricity Load P_t time series (a) and empirical distribution (b). The data corresponds to the period 2017-06-01 to 2018-05-31.

Note that EL empirical distribution may indicate the presence of a bimodal distribution, however, this is not the case. The EL follows a unimodal distribution with the mean value changing during the year because of the seasonality. Furthermore, independent variables are defined from NWPs available at time t. In this case, the weather forecast provider is the OpenWeatherMap service (details of the weather forecast information can be seen in [90]). Previous studies have shown that ambient temperature and solar radiation have a significant influence on the EL[11], so they were included as independent variables. It is important to mention that the solar radiation signal was derived as a combination of the global radiation (solar radiation on a horizontal plane) [96] and the percentage of cloud cover provided by the weather forecast. The objective of using this independent variable is to have a more accurate

approximation of the energy gained by the home that could affect the EL needs, especially during winter. Figure A.3 shows the TS plots for the ambient temperature and solar radiation for two different forecast horizons ($k = 1$ and $k = 18$). The TS plots present the expected seasonal behaviour.

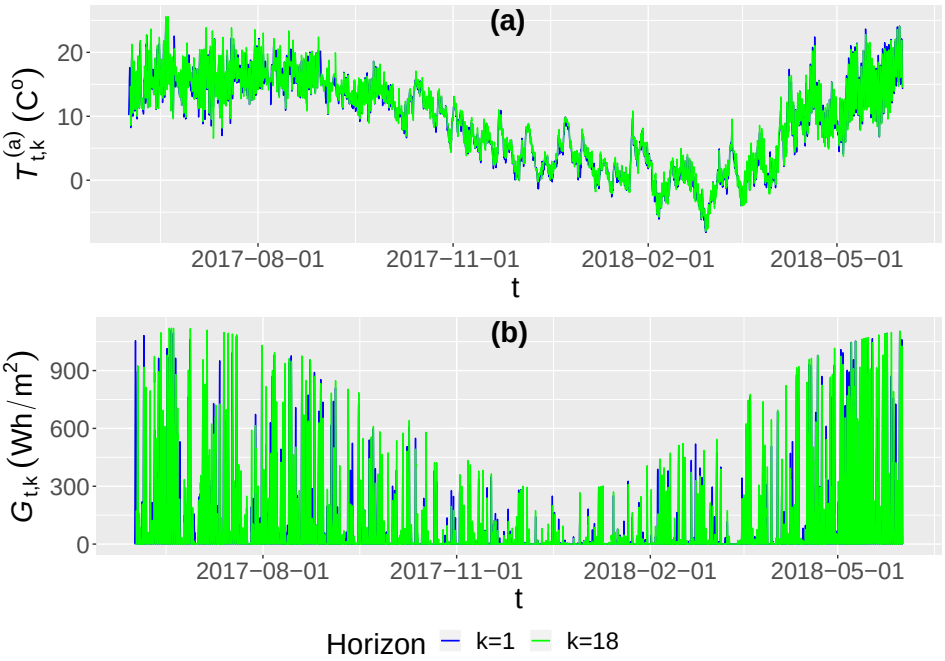

Figure A.3: (a) Ambient temperature $T^{(a)}_{t,k}$. (b) Solar radiation $G_{t,k}$ for horizons $k = 1$ and $k = 18$. The data corresponds to the period 2017-06-01 - 2018-05-31.

Figure A.4 presents a selected short period of the TS. Note that the weather forecast seems to be quite consistent with minor changes between the one-step information $k = 1$ and the eighteen-step information $k = 18$.

From the EL plot in Figure A.4, intra-day patterns can be identified. This is due to normal inhabitants' activities such as breakfast, dinner, and evening TV sessions. The patterns are typically different during workdays and weekends. Thus, in order to include signals related to this intra-day EL patterns for workdays and weekends, Fourier series with $4n$ harmonics were included as independent variables

$$\begin{aligned}
F^{(\sin,\, week)}_{t,k,i} &= \sin\left(2\pi i \frac{t_{\text{day}}}{24}\right) \mathbf{1}_A(\text{day}) \quad \forall i \in [1, 2, \ldots, n] \\
F^{(\cos,\, week)}_{t,k,i} &= \cos\left(2\pi i \frac{t_{\text{day}}}{24}\right) \mathbf{1}_A(\text{day}) \quad \forall i \in [1, 2, \ldots, n] \\
F^{(\sin,\, weekend)}_{t,k,i} &= \sin\left(2\pi i \frac{t_{\text{day}}}{24}\right) \mathbf{1}_B(\text{day}) \quad \forall i \in [1, 2, \ldots, n] \\
F^{(\cos,\, weekend)}_{t,k,i} &= \cos\left(2\pi i \frac{t_{\text{day}}}{24}\right) \mathbf{1}_B(\text{day}) \quad \forall i \in [1, 2, \ldots, n]
\end{aligned} \tag{A.30}$$

where $t_{\text{day}} \in [0, 1, \ldots, 23]$ indicates the time of the day for the t, k period, and $\mathbf{1}_A$ and $\mathbf{1}_B$ are given by

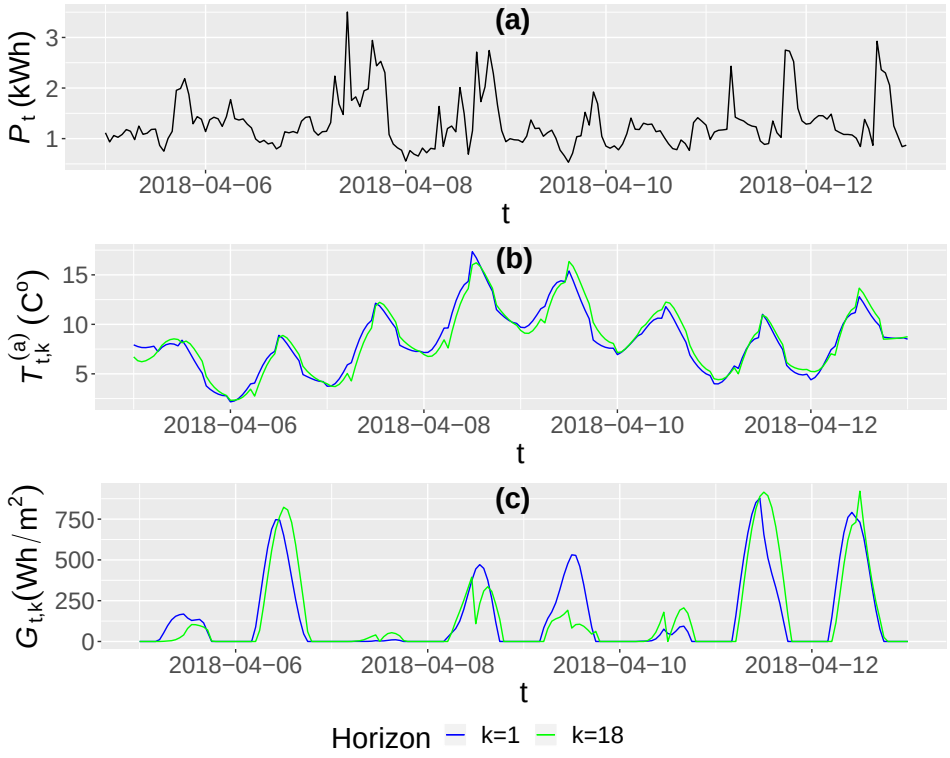

Figure A.4: (a) Electricity Load P_t. (b) Ambient temperature $T^{(a)}_{t,k}$. (c) Solar radiation $G_{t,k}$. The data corresponds to the period 2018-04-05 - 2018-04-20.

$$\begin{aligned} &\mathbf{1}_A(\text{day}) := \begin{cases} 1 \text{ if day} \in A \\ 0 \text{ if day} \notin A \end{cases} \quad \mathbf{1}_B(\text{day}) := \begin{cases} 1 \text{ if day} \in B \\ 0 \text{ if day} \notin B \end{cases} \\ &A = \{\text{Monday, Tuesday, Wednesday, Thursday, Friday}\} \\ &B = \{\text{Saturday, Sunday}\} \end{aligned} \tag{A.31}$$

The harmonics can be expressed in vector form as follows

$$\begin{aligned} \boldsymbol{F}^{(\sin,\, week)}_{t,k} &= (F^{(\sin,\, week)}_{t,k,1}, \ldots, F^{(\sin,\, week)}_{t,k,n}) \\ \boldsymbol{F}^{(\cos,\, week)}_{t,k} &= (F^{(\cos,\, week)}_{t,k,1}, \ldots, F^{(\cos,\, week)}_{t,k,n}) \\ \boldsymbol{F}^{(\sin,\, weekend)}_{t,k} &= (F^{(\sin,\, weekend)}_{t,k,1}, \ldots, F^{(\sin,\, weekend)}_{t,k,n}) \\ \boldsymbol{F}^{(\cos,\, weekend)}_{t,k} &= (F^{(\cos,\, weekend)}_{t,k,1}, \ldots, F^{(\cos,\, weekend)}_{t,k,n}) \end{aligned} \tag{A.32}$$

Moreover, [27] have shown that the EL presents a significant auto-correlation. Thus, with $Y_{t,k}$ as our dependent variable of interest, we can include the lag value y_t as independent variable to account for the auto-correlation. Hence, an entry in the design matrix will contain ambient temperature, solar radiation, a lag value, and $2n$ pair harmonics as independent variables for a future time $t + k$ given t, s.t.

$$\begin{aligned} \boldsymbol{x}_{t,k} = (&T^{(\text{a})}_{t,k}, G_{t,k}, y_t, \boldsymbol{F}^{(\sin,\, week)}_{t,k}, \boldsymbol{F}^{(\cos,\, week)}_{t,k}, \\ &\boldsymbol{F}^{(\sin,\, weekend)}_{t,k}, \boldsymbol{F}^{(\cos,\, weekend)}_{t,k})^\intercal \end{aligned} \tag{A.33}$$

A.3.2 Simulation study

The simulation study was designed to resemble an online application. The aim was to produce a probability forecast of the EL at time t that would cover the following 24 hours in an hourly resolution. Considering the description presented in Section A.3.1, a logarithmic transformation was done in order to ensure the zero bound property of the EL. Thus, the dependent variable of interest Y_t was defined as

$$Y_t = \ln(P_t) \tag{A.34}$$

Y_t can be expressed as a multivariate random variable as described in Equation (A.1) with $k \in \{1, 2, \ldots, 24\}$. Furthermore, the forecast is expected to be updated every hour, considering the independent variables' latest information available $\boldsymbol{X}_t = [\boldsymbol{x}_{t,1}, \ldots, \boldsymbol{x}_{t,K}]$. The online setting implies updating the models' parameters at each time t. While this is inherent to the RLS, the quantile-copula and the models used for the covariance estimation assume a time-invariant framework. Thus, a sliding time window (sliding window) approach was considered in order to deal with the time dependency. This approach will re-estimate the models' parameters at time t using the latest N values. In this way, the non-stationarity characteristic of the EL are modelled. Figure A.5 summarizes the simulation setting.

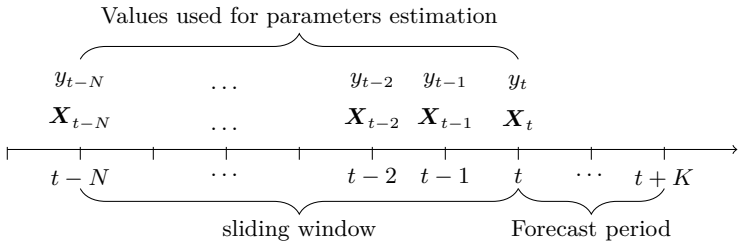

Figure A.5: Graphical representation of the simulation setting for time t.

The simulation setting described previously was used to validate the performance of the proposed models. The period starting from 2018-01-17 to 2018-02-16 was considered as test period. The selection aims to validate the results during a winter time, where the volatility of the EL is higher. The time period starting from the 2017-06-01 to 2018-01-16 was used for the RLS calibration. Moreover, as described in the introduction, five models were implemented. The first model is called "RLS-free", which indicates the use of the RLS method with the correlation structure being estimated directly from the empirical covariance matrix. The second model is called "RLS-AR", which indicates the use of the RLS method with the correlation structure estimated using the maximum likelihood procedure presented in Section A.2.2.3. The third and fourth models are called "Copula-free" and "Copula-AR", which indicates the use of the quantile-Copula method with the same correlation structure connotations. The fifth model is called "RLS", indicating the use of the RLS method without correlation structure, it is used as a reference model. Furthermore, the density function is estimated numerically by using 500 scenarios of $\boldsymbol{Y}_t$ at each time t for all models.

A.4 Results

The results section is divided into three parts. The first part validates the underlying assumptions made for the RLS and quantile-copula. The second part explores the temporal correlation estimation. Finally, a comparison of the different probability forecasts and their performance is made in the third part.

A.4.1 Assumptions check

A.4.1.1 RLS

Figure A.6.a shows EL_t, and the expected values $\exp(\widehat{Y}_{t,1})$ and $\exp(\widehat{Y}_{t,12})$ for the test period in order to evaluate the goodness of the fit. From the plot, we can observe good model fits, however, Figure A.6.b shows that the accuracy of the RLS models degrades with the longer prediction horizons. This is expected due to the less significant effect of the lagged value, the increasing uncertainty in the EL and the weather forecast for longer lead times. Finally, Figure A.6.c presents the histograms of the residuals $\widehat{\varepsilon}_{t,1}$, $\widehat{\varepsilon}_{t,12}$, and $\widehat{\varepsilon}_{t,24}$. While one could argue that the residuals are heavy-tailed, they were considered to be well-behaved. Note that the information presented in Figure A.6 does not show results for all horizons. However, the results for the remaining ks are similar. Please note that the above analysis is made in order to validate the RLS residuals behaviour, a comprehensive analysis of the models' performance is presented in Section A.4.3.

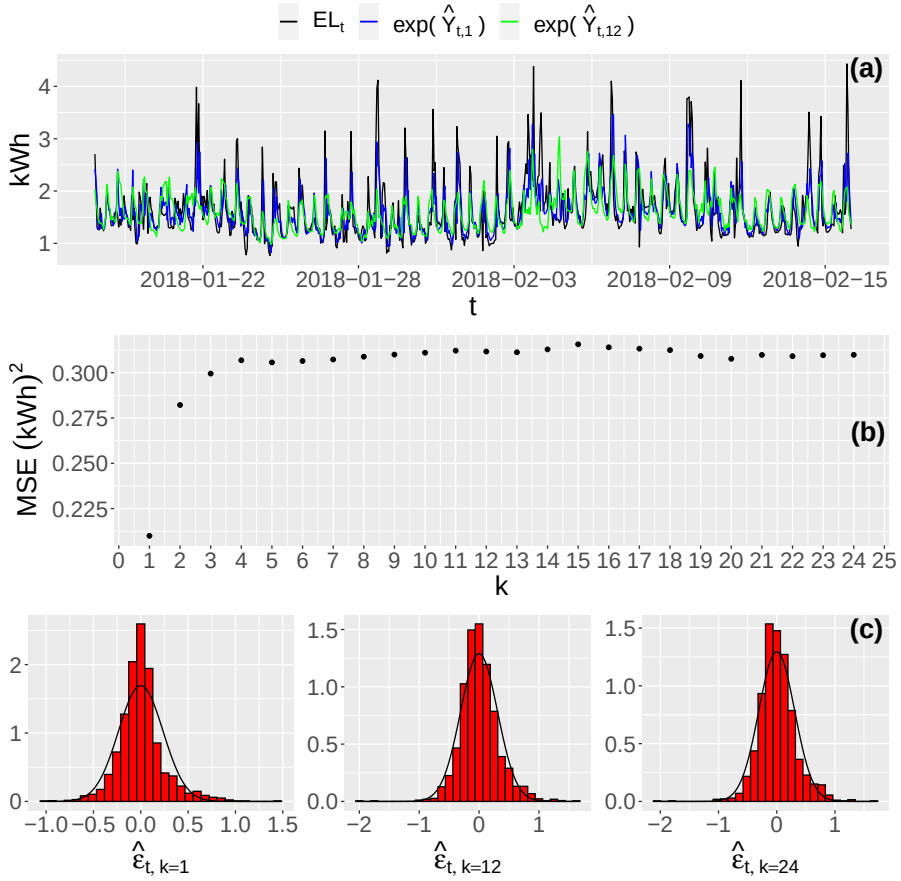

Figure A.6: (a) EL_t, expected values $\exp(\widehat{Y}_{t,1})$ and $\exp(\widehat{Y}_{t,12})$. (b) Mean Square Errors (MSE) per k horizon in the EL domain. (c) Distribution of the $\widehat{\varepsilon}_{t,1}$, $\widehat{\varepsilon}_{t,12}$, and $\widehat{\varepsilon}_{t,24}$ residuals under the logarithmic transformation.

A.4.1.2 Quantile-copula

The quantile regression was made using a uniform partition of the interval $[0.025, 0.975]$ with steps $h = 0.025$ for the τ_i values s.t.

$$0.025 = \tau_1 < \cdots < \tau_i < \cdots < \tau_m = 0.975 \text{ with } \tau_i = ih \tag{A.35}$$

One strong assumption in the quantile-copula method is that $\boldsymbol{Y}_t$ temporal correlation can be captured by a latent multivariate Gaussian process $\boldsymbol{Z}_t$. In order to do this, previous realisations of the process have to be mapped into standard Gaussian distributions as described in Equation (A.13). Thus, Figure A.7 shows QQ-plots of EL realisations under the Gaussian mapping for different time horizons k. The realisations shown in the plots correspond to the sliding window 2017-11-16 - 2018-02-15. The plots do not show enough evidence to reject the normality assumption.

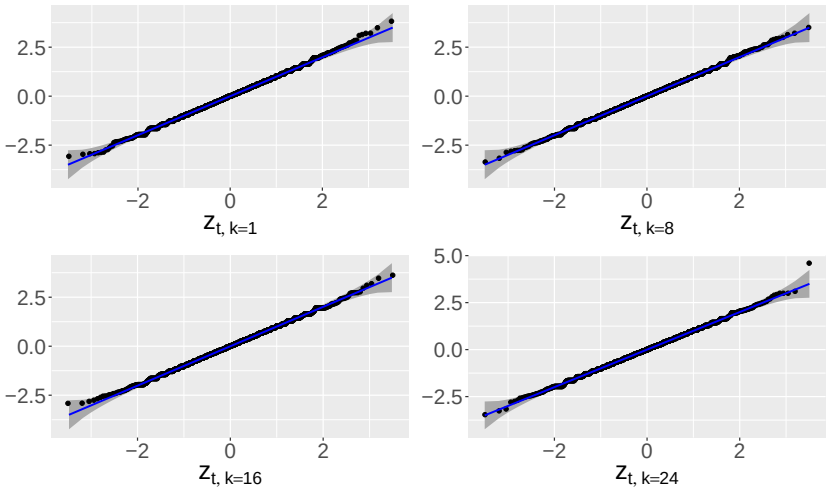

Figure A.7: Normality check of the $z_{t,1}$, $z_{t,8}$, $z_{t,16}$, and $z_{t,24}$ realisations via QQ-plots for the sliding window 2017-11-16 - 2018-02-15.

Please note that since the implemented quantile regression is a function of discrete τ values, an approximation method was needed. A linear interpolation was done in order to obtain the corresponding realisations under the Gaussian mapping for the nominal values not contained in the definition of τ given in Equation (A.35).

A.4.2 Estimation of the temporal correlation structure

A.4.2.1 RLS

Let us start with the estimation of the temporal correlation associated with the RLS method. Figure A.8 shows the $\boldsymbol{\varepsilon}_t$ correlation structure estimated from the full covariance matrix model, and using an AR model for the sliding window period 2017-11-17 - 2018-01-17. At this particular time t the optimal AR process corresponds to an AR(3) with $\boldsymbol{\alpha} = (0.64, -0.05, 0.05)^\intercal$ and $\lambda_{t,k} \sim \mathcal{N}(0, 0.24^2)$. The correlation structures present similar characteristics using the two different models – as seen by the similar around-diagonal patterns and slightly different off-diagonal values.

A.4.2.2 Quantile-copula

The temporal correlation structure estimated with the quantile-copula, for the sliding window period 2017-11-17 - 2018-01-17, is presented in Figure A.9. The correlation structures of the latent Gaussian process $\boldsymbol{Z}_t$ estimated from the full covariance matrix model and using an AR process are presented

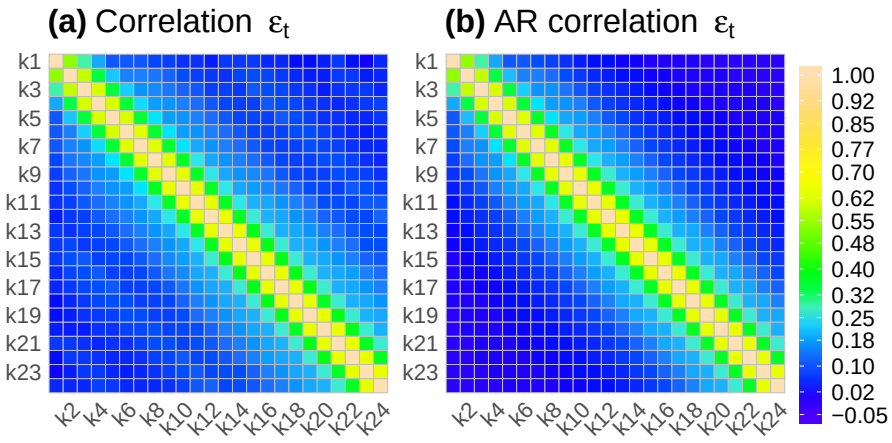

Figure A.8: Residuals $\hat{\varepsilon}_t$ correlation estimated a full covariance matrix model (a) and using an AR(3) process (b). Estimations were made using the sliding window period 2017-11-17 - 2018-01-17.

in the plots. Results indicate a similar estimated correlation structure under the Gaussian domain for both models. The estimated correlation structures seem to catch the around-diagonal elements with smaller correlation present in the off-diagonal elements.

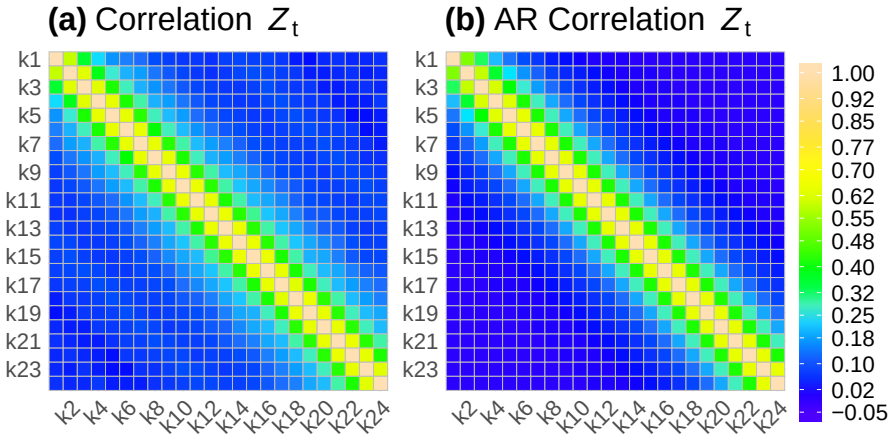

Figure A.9: $\boldsymbol{Z}_t$ correlation estimated using a full covariance matrix model (a) and using realisations of an AR(3) process (b). Estimations were made using the sliding window 2017-11-17 - 2018-02-17.

A.4.3 Forecast performance comparison

Simulations using the different models were carried out. In Figure A.10 the probabilistic forecast, the EL measurements, and 5 random scenarios using the different models are shown. The forecast period starting at t = 2018-02-14 18:00 was selected as an illustrative example. Note that the results are

presented in the EL original domain, which implies the use of the exponential function on the different models' results. From the RLS models we see an almost symmetric distribution around the mean, which is expected due the normality assumption made for the residuals. This naturally differs from the distributions coming from copula models, where we forecast the whole distribution without assuming symmetry at all. Furthermore, the effect of the temporal correlation models can be seen in the smoothness of the scenarios in comparison with the "RLS" model which presents a more erratic behaviour.

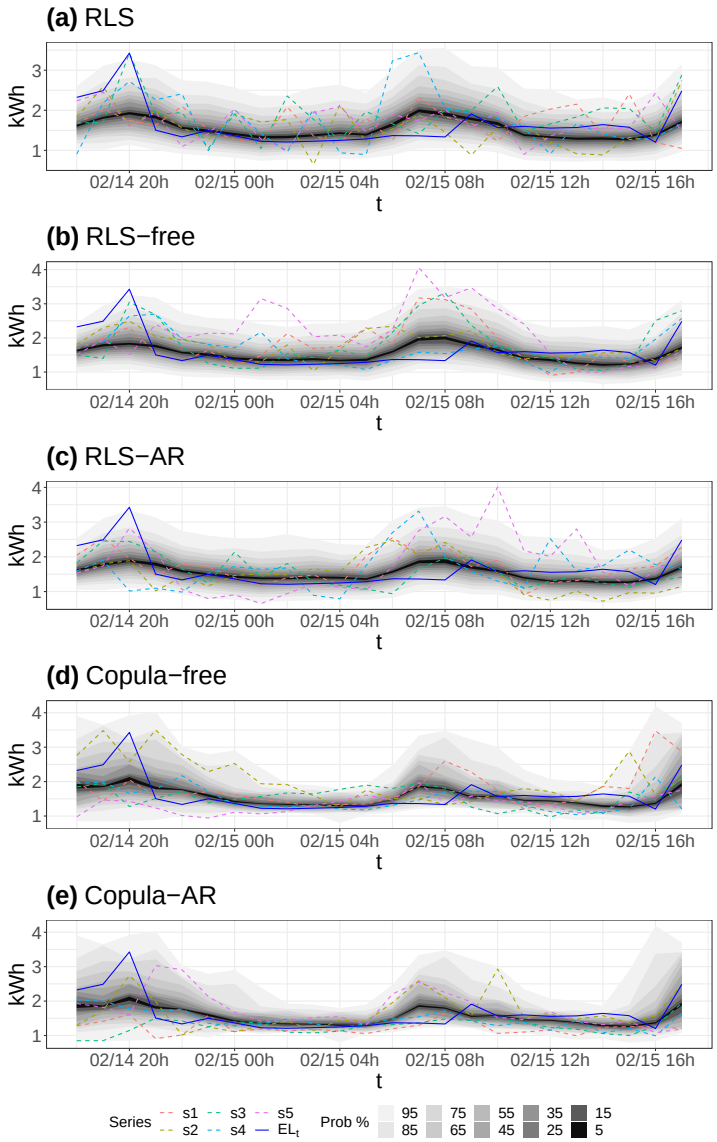

Figure A.10: Probabilistic forecast, EL realisation and 5 scenarios using the RLS (a), the RLS-free (b), the RLS-AR (c), the Copula-free (d), and the Copula-AR (e) models. The forecast period corresponds to 2018/02/14 18:00 - 2018/02/15 18:00.

In order to properly score the performance of the models, a quantitative analysis must be applied. Thus, the different metrics presented in Section A.2.3 were calculated for the different scenarios generated with each model. The mean CRPS, total pinball loss sum, PICP, and PINAW per horizon k are presented in Figure A.11. The plots show a similar performance between RLS models and a similar performance

between copula models. When comparing the performance between the different methods, it can be seen that the quantile-copula presents better performance than the RLS for the metrics that evaluate the probability distribution as a whole (CRPS and pinball loss). However, we can see that for $k = 1$ the RLS presents better performance. This is explained by the effect of the lag value included as input in the models, with the lag value effect degrading with longer prediction horizons. The PICP and PINAW scores were used to evaluate the reliability and sharpness of the forecast. From the PICP results, we can see that both methods present a good reliability based on a 90% significance level prediction interval with the RLS outperforming the quantile-copula. However, the higher reliability of the RLS models comes at the expense of a considerable bigger prediction interval as can be seen in the PINAW results. Furthermore, from the PINAW results one might argue that prediction intervals are in average large (up to 50 of the test period data range for the RLS for longer horizons). However, one should consider the level of random variation of the process at hand. Predicting with a high degree of sharpness the consumption of a single house is a complex task if we consider the data characteristics presented in Section A.3.1 and the longer time horizons.

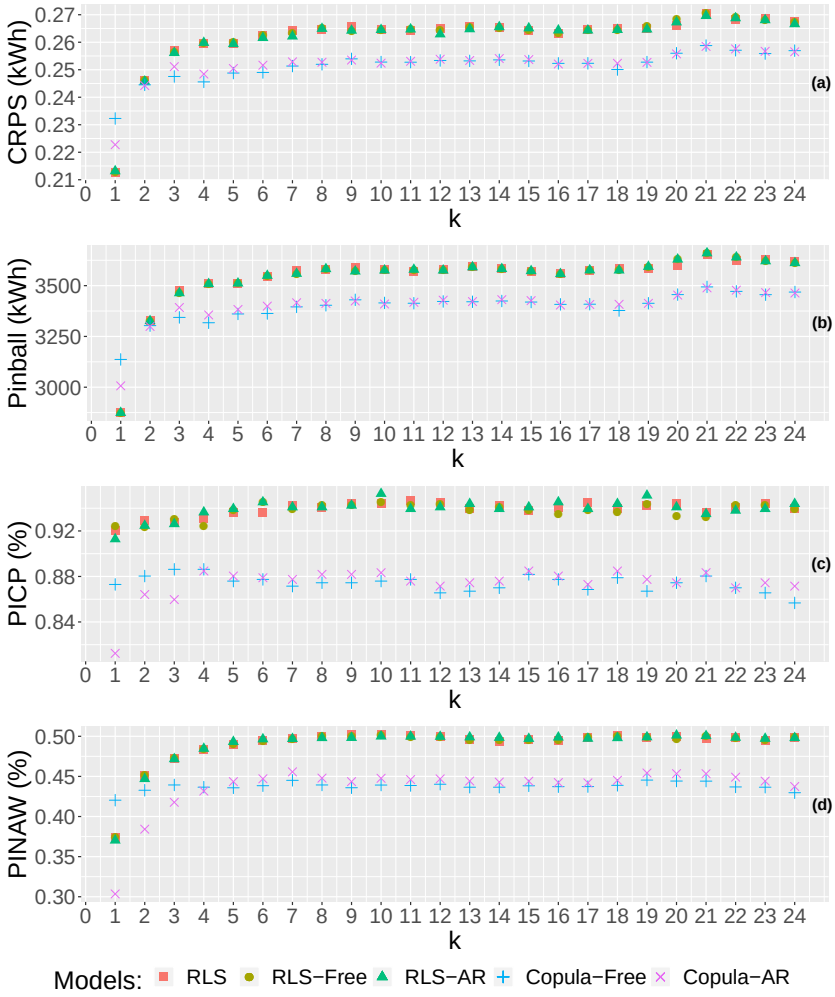

Figure A.11: Mean CRPS (a), pinball loss sum (b), mean PICP (c), and mean PINAW (d) per horizon k for the test period 2018-01-17 - 2018-02-17. Note that the PICP and PINAW are based on a 90% prediction interval.

The VarS was also calculated for the test period, Tab. A.1 presents a summary with the mean score and the percentage of change in comparison to the reference model "RLS". As expected, the reference model presents the worst performance given the lack of a description of the temporal correlation in the model. The proposed methods present a significant performance improvement in comparison with the reference – with the quantile-copula models presenting the best results.

	Mean VarS $(\text{kWh})^2$	VarS % of change
RLS	9.52	-
RLS-free	7.74	-18.68%
RLS-AR	7.79	-18.13%
Copula-free	6.79	-28.76%
Copula-AR	6.82	-28.35%

Table A.1: VarS result for the proposed methods and percentage of change relative to the reference model (RLS). Calculations were made for the test period 2018-01-17 - 2018-02-17.

With the intention of having a better results interpretation, the VarS was discriminated by the difference $j-i$ of the vector entries. Figure A.12 presents the VarS results as relative improvement with respect to the reference model (RLS). Results show that the quantile-copula based method present a better performance for all $j-i$ differences compared to the RLS based methods. Furthermore, little difference is seen between models using a similar modelling technique with different correlation models in addition to the number of parameters used. While in the "free" approach 576 parameters were estimated, in the "AR" approach only 4 parameters were needed.

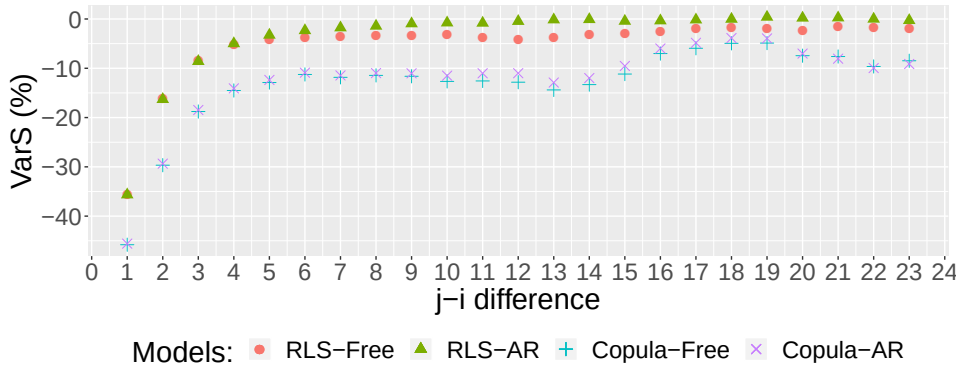

Figure A.12: VarS percentage of change relative to the reference model (RLS) discriminated by difference $j-i$ of the vector entries. Calculations were made for the test period 2018-01-17 - 2018-02-17.

A.5 Discussion

While multivariate probabilistic forecasting considering temporal correlation is a well-studied subject in different domains, we identified a lack of studies applying this type of forecast to households' PLF problem. In the presented study, two different forecasting methods based on RLS and quantile-copula modelling techniques were fully analyzed. Furthermore, two different ways of modelling the temporal correlation structure were investigated as part of the implemented models. The results indicate that

modelling the temporal correlation has a significant impact on the performance of the forecasting models.

In detail, from Section A.4.1, we see that although in this case the RLS residuals were considered well-behaved, one could argue that $\varepsilon_{t,k=1}$ seems to have more density around the mean than in the shoulders of the distribution. This may indicate that the RLS is quite precise, but it may be considered to violate the normality assumption. When analysing the quantile-copula assumptions, the mapping of the nominal values into standard Gaussian distributions presents a robust behaviour. Thus, it is found that it is a superior approach for modelling the particular single household EL case presented, and it is likely that this result generalises to most similar cases – however that should be further studied in order to reach a conclusion.

Checking the marginal distributions of the different models, we can see that the quantile-copula has a significant impact as indicated by the different metric results. The quantile regression has a linear model for each τ value (in this case 39) of each marginal distribution, which allows for a better description of the obviously non-Gaussian distribution of the EL. This is seen in the copula results presented in Figure A.10, where the distribution adapts according to the input data. In contrast, the RLS models use one model per marginal distribution and assumes symmetry around the mean. Furthermore, the results of the performance metrics presented in Figure A.11 indicate a robust behaviour of the presented methodology. However, the interpretation of the PINAW results could be open to a debate. Although one could argue that better sharpness could be expected, one have to considered the data characteristics and the result of other techniques applied to the same and/or similar data. We expect the presented improvements to carry over to examples with better predictive performance. Eg. data sets with less randomness as the aggregated EL of several households. Thus, future studies could focus on the applicability and performance of the presented methodology on different phenomena as well as comparing the methodology to other state-of-the-art modelling techniques on the same and/or similar data sets.

When addressing the estimation of the temporal correlation, the VarS results indicate a significant impact of modelling the correlation structure, even though the estimation of the correlation was performed using a sliding window approach. Furthermore, the results indicate that there is no significant difference between modelling the correlation using an AR(p) process and estimating it from the full covariance matrix model. However, one may argue that the models using the AR(p) process are more robust given the use of significantly less parameters. Finally, it is noted, creating models that consider time-dependent correlation structures may represent a significant improvement – an obvious subject for future studies.

Considering the results, possible future research could also focus on different modelling techniques for the marginal distributions. Please note that the presented modelling methodology is expected to remain the same regardless of the modelling techniques used for the marginal distributions. Along with better models for the marginal distributions, time-dependent models for the temporal correlation are worth studying. Given that the main focus of this study was to present the modelling methodology and the temporal correlation modelling impact, a more comprehensive input model selection and an analysis of the uncertainty produced by the home inhabitants is needed. Future research could involve the impact of changing the sliding window size, adapting the Fourier series, e.g. to public and non-public holidays. Moreover, large-scale tests (a test involving data from hundreds of users) of the presented methods will be needed to confirm the robustness of the models. The large-scale tests should include a more detailed analysis of the independent variables' impact especially during different seasons. Even though the need for a large-scale test is latent, the modelling method presents a significant step towards models ready for a production setting where the generated ensemble forecasts are used as input for optimising the EL household needs.

A.6 Conclusion

In this paper, it was shown how to build multivariate probabilistic models, which consider temporal correlation to accurately generate ensemble forecasts of the EL of single households. The models use as inputs a combination of NWPs, and Fourier series harmonics to describe intra-day patterns. Two main modelling methods were applied: A method based on RLS models for the estimation of the EL expected value in combination with models for the RLS's errors to have a full picture of the process' multivariate distribution. A second method based on linear quantile regression models and a Gaussian copula to link the different marginal distribution, thus estimating the process' multivariate distribution. Results show that the quantile-copula models present better performance describing the individual marginal distributions and the temporal correlation compared to the RLS based models.

As part of the developed models, two different models for the estimation of the temporal correlation were presented. In the first model, the temporal correlation is estimated using a full covariance matrix model. The other uses a likelihood method to estimate an AR(p) process which describes the temporal correlation. Results show that there is not a significant difference between the two modelling approaches. This implies that the estimation of the temporal correlation could be done using a full covariance model, avoiding the extra computational time inherent to the likelihood calculation. However, in cases where the robustness of the results is a greater concern, the approach using a AR model would be more suitable. Finally, the simulation study showed that all of the models present a good performance and could be developed to be production ready in smart HEMS.

Acknowledgements

This work was partially funded by Innovation Fund Denmark through the project identified with case No. 8053-00156B.

PAPER B

Economic evaluation of stochastic home energy management systems in a realistic rolling horizon

Authors:
Julian Lemos-Vinasco, Amos Schledorn, S. Ali Pourmousavi, Daniel Guericke

Submitted to:
Applied Energy, currently under review

Preprint:
https://arxiv.org/abs/2203.08639

Economic evaluation of stochastic home energy management systems in a realistic rolling horizon

Julian Lemos-Vinasco[1] [2], Amos Schledorn [1], S. Ali Pourmousavi[3], Daniel Guericke[1]

Abstract

Home energy management systems (HEMSs) are expected to become a crucial part of future smart grids. However, there is a limited number of studies that comprehensively assess the potential economic benefits of HEMSs for consumers under real market conditions and which take account of consumers' capabilities. In this study, a new optimization-based HEMS controller is presented to operate a photovoltaic and battery system. The HEMS controller considers the consumers' electrical load uncertainty by integrating multivariate probabilistic forecasting methods and a stochastic optimization in a rolling horizon. As a case study, a comprehensive simulation study is designed to emulate the operation of a real HEMS using real data from nine Danish homes over different seasons under real-time retail prices. The optimization-based control strategies are compared with a default (naive) control strategy that encourages self consumption. Simulation results show that seasonality in the consumers' load and electricity prices have a significant impact on the performance of the control strategies. A combination of optimization-based and naive control strategy presentskk the best overall results.

B.1 Introduction

As one of the major smart grid technologies, home energy management systems (HEMSs) are expected to play a key role managing energy consumption at the residential level by reacting to real-time prices and/or CO_2-based signals. In Europe, this coincides with efforts of electricity market operators and policy makers to push for a wider adoption of real-time tariffs for residential consumers that reflect the true condition of the power system and provide cost-savings for consumers [92, 34].

High expectations have been placed on HEMSs by many industry stakeholders given the systems' potential to provide a dynamic combination of production, storage, and flexible demand [77, 42, 49]. Therefore, studies on this topic have emerged from a variety of disciplines over the last decade, focusing on different components of the HEMSs. Typically, HEMSs rely on a combination of smart home technologies (SHT) such as smart meters, sensing devices, communication hardware and protocols, smart appliances, controllers, and optimization techniques [12]. The operation and coordination of these components entails technical difficulties, especially for SHT that depend on manual intervention from end users. This has led to a literature bias towards SHT solutions that require minimal consumer intervention [79].

In this regard, several studies have proposed sophisticated technical solutions by assuming a direct control of several SHT. These solutions have been used for direct control of the heating systems of homes and buildings, smart appliances, RESs, batteries, and electrical vehicle chargers (in both grid-2-vehicle and vehicle-2-grid modes), with some parameters being defined by the consumer [127, 103]. Furthermore, it is important for HEMSs to consider complex system features such as the multi-seasonality,

[1]Technical University of Denmark, Department of Applied Mathematics and Computer Science, Lyngby, Denmark
[2]Watts S/A, Section of Research and Development, Svinninge, Denmark
[3]The University of Adelaide, School of Electrical and Electronic Engineering, Adelaide, Australia

non-stationarity, and stochasticity of RESs and consumers' EL [72, 26]. Therefore, recent studies have included several of these features. In [108], a stochastic HEMS was proposed that considered consumers' satisfaction cost and fatigue towards demand-response signals. The authors of [129] proposed a two-stage stochastic model with scenarios for wind power and electric vehicles' availability. In [21], a similar approach is used with additional considerations for the battery degradation cost. Other studies apply rolling horizon approaches. Such approaches provide an opportunity to re-optimize the problem when new information about stochastic elements are available, for example, PV forecast [93, 28].

The studies in the literature on control strategies for HEMSs display several similarities. First, the studies assume direct control over different SHT (controlled laboratory conditions and/or simulations). Second, most studies are mainly oriented towards demand-response programs by assuming access to the wholesale market electricity prices (day-ahead and/or intra-day prices). Third, they present a cost-benefit comparison for a limited time period (ranging from days to weeks), typically in cold seasons with a passive consumer (consumer without SHT) as the baseline. In contrast to these publications, the results of field studies and trial projects have questioned the real benefits that consumers will be able to perceive. Results from a nine-month field trial with ten households in the UK concluded that "there is little evidence that SHT will generate substantial energy saving and, indeed, there is a risk that they may generate a form of energy intensification" [50]. These observations are aligned with the findings in [86], where in a trial with 40 households with basic SHT, minimum economic benefits were reported, with some households reporting energy intensification. However, the setups of these studies used smart appliances requiring manual consumer interventions. Moreover, [85] suggest that very limited economic benefits can be expected from these types of setups because of the inherent inflexibility of some consumers.

Although the above studies indicate a need for more research on SHT that require manual intervention from consumers, the main body of literature assumes direct access and control of most SHT elements. This is a strong assumption that may distort the studies' results [112]. Furthermore, the results mainly describe technical aspects with assumptions that may not work under current market rules. For instance, in [21, 93, 127], it is not clear if the electricity prices used correspond to prices accessible to consumers, or they assumed that consumers have access to the wholesale electricity markets, which is not possible due to the small size of individual consumers' load and RESs in the European markets [45]. Assuming access to wholesale market prices disregards the fact that consumers are subject to taxes, levies, and fees, which may have a significant impact on the results. Additionally, the cost comparisons are made with a passive consumer as baseline, disregarding the fact that simple self-consumption control strategies have proved to bring significant cost reductions [6].

On the basis of the above discussion, one can argue that there is a need for studies that assess the economic potential of HEMSs under system conditions and market rules accessible to the residential consumers. These conditions must include a realistic HEMS setup, end-consumer prices, cover a substantial period of time consisting of different seasons, and compare the results to self-consumption control strategies.

We offer a comprehensive economic assessment of a HEMS under realistic consumer and electricity market conditions. We propose a HEMS control strategy that uses a stochastic optimization framework in a rolling horizon approach and probabilistic forecasts. In particular, the HEMS setup is modeled as a stochastic MILP where the uncertainty of the consumers' EL is considered using two different multivariate probabilistic forecast methods. Moreover, the rolling horizon approach is used to allow the possibility of re-optimizing according to the latest information available to the system. The data used in the case study corresponds to nine households located in Copenhagen, Denmark, together with real hourly electricity retail prices offered by a utility company. Furthermore, a HEMS setup with only a PV and battery system is considered to emulate the possibilities that most residential consumers have at present. Although electric vehicles are a key element of the HEMSs of the future, the adaption electric vehicles is still low in Denmark [68] and therefore they were not included in the analysis. Operational

and cost results of the proposed optimization-based strategies are compared with a passive consumer as well as a self-consumption (naive) control strategy.

Overall, the results indicate that a combination of an optimization-based and a naive control strategy presents a higher economic benefit for residential consumers throughout the year. Key research findings are summarized below:

1. The stochasticity of the consumers' EL has a significant impact on the performance of the optimization-based control strategies.
2. Strong seasonality in consumption patterns shows a significant effect on the assessment and selection of the control strategies, with a self-consumption strategy (naive control) outperforming the optimization-based controllers in spring and summer.
3. Under current market rules, residential consumers are not sufficiently incentivized to actively participate in the electricity market besides covering their electricity demand.

This paper starts by presenting the HEMS setup and the mathematical details of the implemented models in Section B.2. Next, the data and the case study are explained in Section B.3. The simulation results are presented in Section B.4, which includes a comparison between different control strategies, and a comprehensive cost analysis. Finally, a discussion of the findings and perspectives for future work are outlined in Section B.5. The paper is concluded in Section B.6.

Nomenclature

Sets	
$\mathcal{T}$	Set of time steps t
$\mathcal{S}$	Set of scenarios s
Parameters	
$D_{s,t}$	Electricity load in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kWh]
λ_t^+	Electricity purchase cost in period $t \in \mathcal{T}$ [DKK/kWh]
λ_t^-	Electricity sale price in period $t \in \mathcal{T}$ [DKK/kWh]
PV_t	PV production in period $t \in \mathcal{T}$ [kWh]
PV^{peak}	PV system peak production [Wh]
S^{ini}	Initial battery SoC [Kwh]
$S^{\max}$	Battery maximum storage capacity [kWh]
$S^{\min}$	Battery minimum storage capacity [kWh]
B^{in}	Battery charge limit per period [kW]
B^{out}	Battery discharge limit period [kW]
π_s	Probability of scenario $s \in \mathcal{S}$
$\phi^{\text{DC/AC}}$	Efficiency factor when inverting power flows from direct current (DC) to alternate current (AC)
$\phi^{\text{AC/DC}}$	Efficiency factor when converting power flows from AC to DC
η^+	Battery charge efficiency factor
η^-	Battery discharge efficiency factor
M	Big M value define as $M = \max(PV_t) + \max(D_{s,t}) + B^{\text{in}}$
Variables	
$x_{s,t}^+ \in \mathbb{R}^+$	Electricity bought from the electricity retailer in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kWh]
$x_{s,t}^- \in \mathbb{R}^+$	Electricity sold to the electricity retailer in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kWh]
$g_{s,t}^{\text{D}} \in \mathbb{R}^+$	Power from the grid used to satisfy the demand in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kW]
$g_{s,t}^{\text{b}} \in \mathbb{R}^+$	Power sent from the grid to the battery in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kW]
$b_{s,t}^+ \in \mathbb{R}^+$	Battery charge power flow in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kW]
$b_{s,t}^- \in \mathbb{R}^+$	Battery discharge power flow in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kW]
$b_{t,s}^{\text{g}} \in \mathbb{R}+$	Power delivered from the battery to the grid in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kW]
$b_{t,s}^{\text{D}} \in \mathbb{R}+$	Power from the battery used to satisfy the demand in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kW]

$pv_t^{\mathrm{g}} \in \mathbb{R}+$	Power delivered directly from the PV system to the grid in period $t \in \mathcal{T}$ [kW]
$pv_t^{\mathrm{D}} \in \mathbb{R}+$	Power from the PV system used to satisfy the demand in period $t \in \mathcal{T}$ [kW]
$pv_t^{\mathrm{b}} \in \mathbb{R}+$	Power from the PV system to the battery in period $t \in \mathcal{T}$ [kW]
$s_{s,t} \in \mathbb{R}^+$	Battery SoC in scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$ [kWh]
$z_{s,t} \in \{0,1\}$	Binary variable, indicating if electricity was purchased or sold to the grid for scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$
$y_{s,t} \in \{0,1\}$	Binary variable, indicating if the battery is charging or discharging for scenario $s \in \mathcal{S}$ and period $t \in \mathcal{T}$

Table B.1: MILP mathematical nomenclature

B.2 Modeling and optimization of HEMSs

B.2.1 HEMS setup

The HEMS setup considered in this paper reflects the current network conditions in Denmark that residential electricity consumers with access to a PV and a home battery system face. Moreover, a minimal control approach is considered for the HEMS model. This means that the HEMS has direct control of the home battery, but it does not have direct control over the home appliances. A graphical overview of the setup is given in Figure B.1.

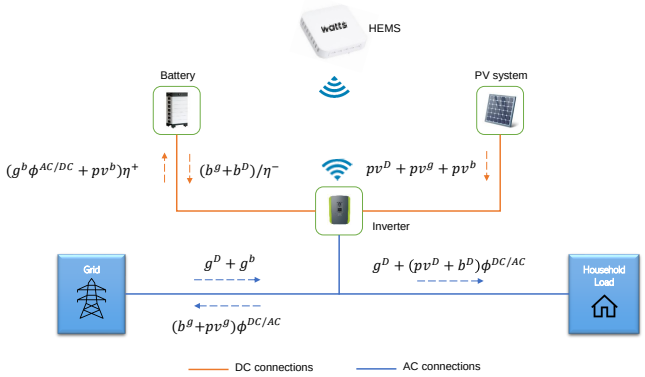

Figure B.1: Schematic diagram of the home's setup showing the AC and DC connections. The different average power flow terms, which are used in the mathematical representation of the system (Section B.2.2), are also included.

The electricity generation from the PV system can be used to charge the home battery, to meet the EL demand, or can be exported to the grid. The electricity losses due to AC/DC and DC/AC conversions are included in the formulation. We assume that the electricity retailer communicates price information to the HEMS and that the HEMS has access to NWPs. Furthermore, NWPs are input to PLF models used for the creation of EL scenarios. The data of real consumers in Denmark is used, however, these consumers did not have PV installations. Thus, a simulation model for PV production is implemented and presented in Section B.2.4. The remainder of this section introduces the mathematical model formulation for the above setup.

B.2.2 HEMS optimization model

The HEMS controller is formulated as a stochastic MILP [13], where the EL is the only uncertain parameter, i.e. having varying realizations across scenarios. The MILP in (B.1) minimizes the expected cost for fulfilling the EL in all scenarios $\mathcal{S}$ and periods $\mathcal{T}$. The main decision variables are the flows

between the PV, grid, and battery components. The model considers several time periods due to the temporal interdependence imposed by the battery SoC in a rolling horizon manner. This means that, when applying the solution to the HEMS, only the optimal solution for the first time period is applied in practice. Decisions in subsequent periods are only considered to find optimal decisions for the first time step. This allows for re-optimization and taking relevant decisions with updated forecasts.

$$\min_{x^-_{s,t}, x^+_{s,t}} \quad \sum_{s\in\mathcal{S}} \sum_{t\in\mathcal{T}} \pi_s(\lambda^+_t x^+_{s,t} - \lambda^-_t x^-_{s,t}) \tag{B.1a}$$

subject to

$$D_{s,t} = (g^{\mathrm{D}}_{s,t} + (pv^{\mathrm{D}}_t + b^{\mathrm{D}}_{s,t})\phi^{\mathrm{DC/AC}})\Delta t \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1b}$$
$$x^+_{s,t} = (g^{\mathrm{D}}_{s,t} + g^{\mathrm{b}}_{s,t})\Delta t \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1c}$$
$$x^-_{s,t} = (b^{\mathrm{g}}_{s,t} + pv^{\mathrm{g}}_t)\phi^{\mathrm{DC/AC}}\Delta t \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1d}$$
$$x^-_{s,t} \leq M z_{s,t} \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1e}$$
$$x^+_{s,t} \leq M(1 - z_{s,t}) \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1f}$$
$$PV_t = (pv^{\mathrm{g}}_t + pv^{\mathrm{D}}_t + pv^{\mathrm{b}}_t)\Delta t \qquad \forall t \in \mathcal{T} \tag{B.1g}$$
$$b^+_{s,t} = (g^{\mathrm{b}}_{s,t}\phi^{\mathrm{AC/DC}} + pv^{\mathrm{b}}_t)\eta^+ \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1h}$$
$$b^-_{s,t} = (b^{\mathrm{g}}_{s,t} + b^{\mathrm{D}}_{s,t})/\eta^- \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1i}$$
$$b^+_{s,t} \leq B^{\mathrm{in}} \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1j}$$
$$b^-_{s,t} \leq B^{\mathrm{out}} \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1k}$$
$$b^+_{s,t} \leq M y_{s,t} \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1l}$$
$$b^-_{s,t} \leq M(1 - y_{s,t}) \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1m}$$
$$s_{s,t} = S^{\mathrm{ini}} + (b^+_{s,t} - b^-_{s,t})\Delta t \qquad \forall s \in \mathcal{S},\, t = 1 \tag{B.1n}$$
$$s_{s,t} = s_{s,t-1} + (b^+_{s,t} - b^-_{s,t})\Delta t \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S},\, t > 1 \tag{B.1o}$$
$$S^{\min} \leq s_{s,t} \leq S^{\max} \qquad \forall t \in \mathcal{T},\, s \in \mathcal{S} \tag{B.1p}$$
$$b^+_{s,t} = b^+_{j,t}\,,\; b^-_{s,t} = b^-_{j,t} \qquad \forall s, j \in \mathcal{S},\, s \neq j,\, t = 1 \tag{B.1q}$$
$$s_{s,t} = s_{j,t} \qquad \forall s, j \in \mathcal{S},\, s \neq j,\, t = 1 \tag{B.1r}$$

The objective function (B.1a) minimizes the expected cost of electricity. Furthermore, constraint (C.1d) ensures that the consumer's demand is satisfied in all scenarios. Electricity purchase and sale quantities are set in constraints (B.1c) and (B.1d). Constraints (B.1e) and (B.1f) ensure that electricity sale and purchase are mutually exclusive. The PV power balance is set in constraint (B.1g) such that the total generation meets the sum of PV production to grid, demand and battery. Constraints (B.1h), (B.1i), (B.1j), and (B.1k) model the physical battery behaviour in terms of power flow. Simultaneous charging and discharging of the battery is disallowed in constraints (B.1l) and (B.1m). The evolving SoC is modelled by constraints (B.1n) and (B.1o), while constraint (B.1p) limits the SoC to the battery capacity.

Since it is possible to re-optimize the solution after one time period, and the energy exchange with the grid is unrestricted, we can frame the problem as a two-stage stochastic problem. The first stage of the problem defines the operational schedules of the battery in the first period, as given by constraints (B.1q) and (B.1r). Thus, the battery charging and discharging in the first time period needs to be the same for all scenarios.

B.2.3 Electrical load forecast

The HEMS optimization model presented in Section B.2.2 uses EL scenarios as input. The scenarios must consider the temporal correlation inherent to the EL. Thus, the multivariate PLF methods

presented in [72] are used in this paper to generate the required scenarios. The methods use either RLS with a full covariance model for the residuals or the quantile-copula with a full covariance model of the temporal correlation under the Gaussian domain and are referenced in [72] as *RLS-Free* and *Copula-Free*.

B.2.4 PV simulation

Another key element of the HEMS is the PV system. In this study, PV generation data were not available. Thus, a simulation approach is used to estimate a rooftop PV production. The simulation model is based on the guidelines provided in the energy data catalogue by the Danish Energy Agency [24]. The report suggests that electricity produced by a PV system should be estimated as

$$PV_t = PV^{peak} \cdot \frac{GHI_t}{1000} \cdot PV^{\mathrm{tf}} \cdot PV^{\mathrm{pr}} \tag{B.2}$$

where PV^{peak}corresponds to the PV production under laboratory standard test conditions (1000 W/m^2 irradiation with a cell temperature of 25°C), PV^{tf} is the PV transposition factor and PV^{pr} is the PV performance ratio. Moreover, the GHI_t (global horizontal irradiation) values are calculated using the deterministic cloud cover to GHI model described in [69] and given by

$$GHI_t = GHI_t^{CS} \cdot [0.35 + 0.65 \cdot (1 - CC)] \tag{B.3}$$

where cc is the normalized cloud coverage (0 = clear, 1 = overcast) that is obtained from a NWP model, and the GHI_t^{CS} is the clear sky GHI that is estimated by the clear sky methods provided by the *pvlib* Python package [53].

B.3 Case study

In this section, we describe the input data used by the HEMS models presented in Section B.2.1, and the technical details of the simulation setup used to calculate the results.

B.3.1 Electrical load

The EL demand profiles of nine residential consumers are shown in Figure B.2. The consumption data results from smart meters sampled at an hourly resolution for the year 2020. Information given about these consumers includes the number of inhabitants, the approximate house location given by its longitude and latitude coordinates, and the fact that they use heat pumps as heating technology. The use of heat pumps explains the seasonality of the EL, i.e., a significantly higher consumption during winter in comparison to summer (see Figure B.2). Although not visible in yearly plots, intra-day patterns can be found on a closer inspection of the data (see Figure B.3). These patterns may be explained by the daily routine activities of the tenants, e.g., having breakfast and dinner at regular times. These factors together with the data presented in Section B.3.2 were considered when building the PLF used for the scenario generation of each consumer load demand. While it is out of the scope of this paper to describe and analyze the PLF methods, they are described in detail in [72].

B.3.2 Numerical weather prediction

Both the PV simulation model (Section B.2.4) and the PLF methods (Section B.2.3) rely on the NWP values as their primary input parameters. Here, the weather forecast provided by the OpenWeatherMap service at an hourly resolution was used. The NWP data are described in [90]. In particular, the ambient temperature and solar irradiation were used for the PLF and PV models. Please note that the

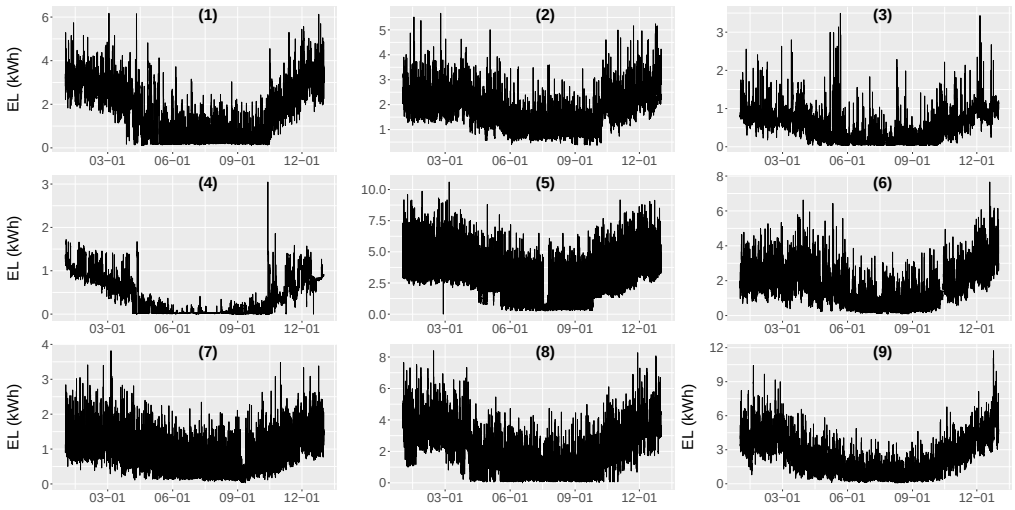

Figure B.2: Consumers' electricity consumption for the time period 2020-01-01 to 2020-12-31.

solar irradiation signal was derived using the expected cloud coverage from the NWPs in combination with the Global Horizontal Irradiation (GHI), as described in Section B.2.4.

The EL and NWP data are combined to produce a coherent dataset used for the HEMS. An example week for one user is shown in Figure B.3. Please note that the NWP is updated every hour with the forecast values covering several hours ahead (k horizons). The NWP for one hour ahead $k = 1$ and eighteen hours ahead $k = 18$ are presented in Figure B.3 (b) and (c).

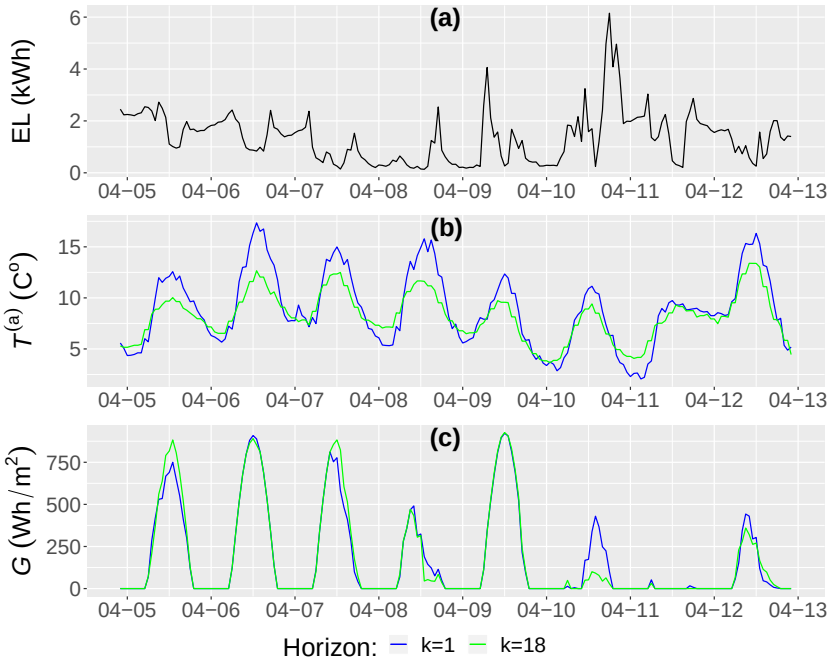

Figure B.3: Consumer 1's electricity consumption (a), ambient temperature (b), and solar irradiation (c) for the time period 2020-04-05 to 2020-04-13.

B.3.3 Electricity prices

Given the high penetration of smart meters in Denmark, it is common for electricity retailers to offer hourly prices to residential consumers. Typically, this type of tariff is derived from the day-ahead (DA) wholesale electricity market prices, also known as ELSPOT [44]. In the ELSPOT market, different zones/regions have their unique DA prices. In Denmark, two price zones zones exist: Western Denmark (DK1) and Eastern Denmark (DK2) [43]. To obtain the electricity prices for the residential consumers, retailers add taxes, levies, and fees to the DA prices. In this study, actual retail electricity prices provided by the Danish electricity retailer Watts are used [123]. All consumers are located in the greater Copenhagen area, which is a part of DK2 region. Figure B.4 shows the retailer prices and DA electricity prices for the period of 2020-01-01 to 2021-12-03. The current Danish regulations allow residential consumers to sell their surplus electricity back to the grid. The feed-in-tariff is decided by the retailers. Most of them offer the ELSPOT price adjusted for associated operational fees as feed-in-tariff to residential consumers, as described by [116].

In our case study, Watts electricity prices are used as real-time price paid by the consumers to purchase electricity, while DA prices are used as feed-in-tariff for selling. This corresponds to the parameters λ_t^+ and λ_t^-, respectively (see Section B.2.2).

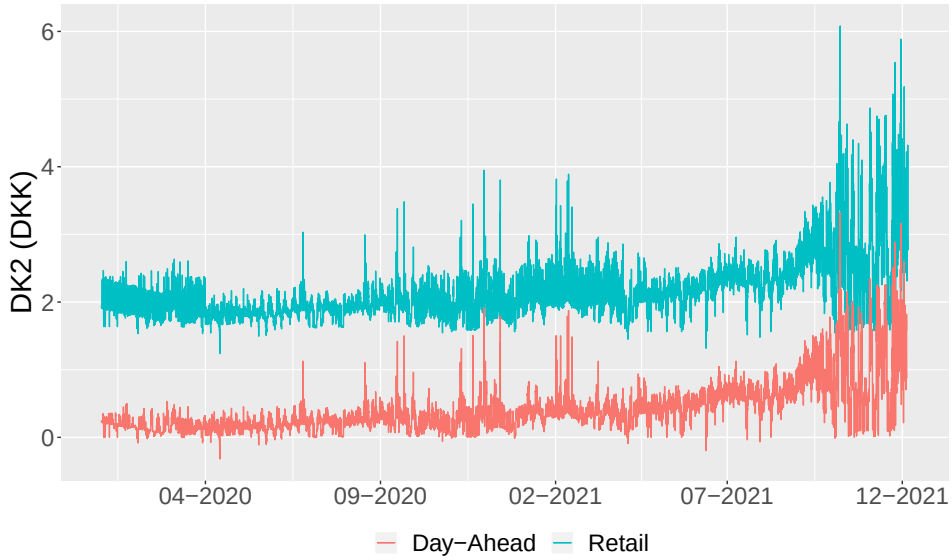

Figure B.4: Retailer prices and DA electricity prices for the time period 2020-01-01 to 2021-12-03 for DK2.

Please note the differences in prices in 2020 and 2021. On the one hand, a tariff regulation change was introduced in 2021 that stipulates a low (between 00.00 to 17.00 and 20.00 to 00.00) and a peak (between 17.00 to 20.00) electricity distribution fee from October to March [98]. On the other hand, 2021 was a year with unusual electricity prices, i.e., high prices and high volatility [19], as can be seen in the final quarter of 2021 in Figure B.4.

The HEMS setup and formulation allows cost reduction by selling excess electricity (excess electricity from PV), trading (buy at low-price hours to sell at high-price ones), and load shifting. This can be done by exploiting price volatility. Thus, for further reference, the mean and standard deviation of the prices are presented in Table B.2. Note that the statistics are only presented for January, April, July, and October of 2021. These are the months that are included in the simulation setup described

in Section B.3.4. Please note that at present, consumers are mainly passive users of electricity, which has no significant effect on prices as they are price-takers [55].

Time Period	Purchase price Mean	Purchase price SD	Selling price Mean	Selling price SD
January 2020	1.943	0.164	0.206	0.074
January 2021	2.085	0.230	0.379	0.126
April 2020	1.797	0.107	0.129	0.086
April 2021	2.008	0.120	0.357	0.160
July 2020	1.858	0.144	0.191	0.115
July 2021	2.330	0.198	0.615	0.159
October 2020	1.921	0.193	0.202	0.118
October 2021	2.624	0.780	0.810	0.591

Table B.2: Purchase and selling prices mean and standard deviation (SD) for January, April, July and October.

B.3.4 Simulation setup

The simulation study is designed to resemble a real-time application. The aim of the simulation is to optimize the battery's operational setpoints for the next hour when considering a 24-hour horizon. Therefore, a rolling horizon approach is used, which means that the PLF, PV simulation, and HEMS optimization will be updated every hour to determine the new operation schedules. A graphical representation of the rolling horizon simulation setting at time t is presented in Figure B.5. Please note the PLF models are re-fitted using $t - N$ historical values at each time step t. Thus, more accurate prediction can be expected by using the latest available information from the forecasting models.

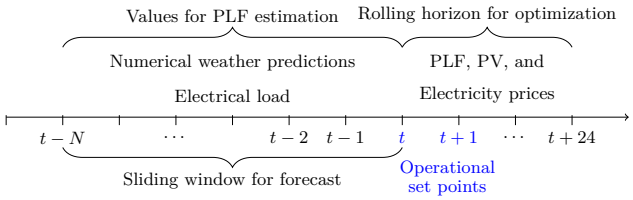

Figure B.5: Graphical representation of the simulation setting for time t.

Four months of data (January, April, July and October) are selected as representative for seasonal variations in order to analyse one year of operation. Moreover, the prices used in the simulation are the prices in DK2 from 2021. This is motivated by the regulation changes introduced in that year. Please note that the consumers' EL data were only available for 2020. Therefore, we used consumers' data from 2020 with prices from 2021 in our simulations, assuming that the EL in the selected months of 2020 is likely to be similar to the EL in the same months in 2021 and the fact that residential consumers are price-takers.

B.4 Simulation results

In this section, we compare consumers' cost savings when using different control strategies. Two such strategies are considered: a naive controller and an optimization-based controller. A naive controller refers to a consumer with PV and battery system without a HEMS. This controller maximizes self consumption by only selling electricity to the grid when the battery is fully charged. The naive controller

uses neither forecasting nor optimization methods and it is usually the default controller in the HEMS setup presented in Section B.2.1 [6]. The optimization-based controller refers to a consumer using a HEMS with optimization and forecast capabilities as presented in this paper. With this in mind, the rest of the results section is organized as follows. In Subsection B.4.1, we determine the most suitable PLF method for the optimization-based controller by comparing the performance of the proposed HEMS optimization model using different forecasting methods. A comparison between the best optimization-based controller and a naive controller is presented in Subsection B.4.2. Results indicate the optimal strategy is a combination of a naive and a optimal controller. This is presented in detail in Subsection B.4.3.

B.4.1 Comparison of different optimization-based controllers

The PLF methods presented in [72] allow different combinations of forecasting and optimization methods. This section discusses the performance of such combinations in order to select the best method. The analyzed combinations are the following:

- **RLS-SP**: the proposed HEMS optimization using 100 scenarios generated by the RLS forecasting method.
- **RLS-EV**: the proposed HEMS optimization using the expected value of the 100 scenarios generated by the RLS forecasting method.
- **Copula-SP**: the proposed HEMS optimization using 100 scenarios generated by the Copula forecasting method.
- **Copula-EV**: the proposed HEMS optimization using the expected value of the 100 scenarios made by the Copula forecasting method.
- **PI-RH**: perfect information (PI) in a rolling horizon, i.e, using the proposed HEMS optimization assuming that the consumer's load is known. This method is not applicable in practice, since the PI-RH method assumes a perfect knowledge of the future demand. But it can be used to give performance bounds on the optimization in the other settings.

Please note that "EV" methods correspond to a deterministic version of the HEMS stochastic model presented in Section B.2.2. This allows us to assess the impact of modeling the EL uncertainty by comparing with "SP" methods. Table B.3 presents the total electricity cost for the simulated months for the different combinations of forecasting and optimization methods. Moreover, the table reports the cost relative to the theoretical approach PI-RH expressed as percentages. The simulation results indicate that the optimization using scenarios-based stochastic programming outperforms the other solutions for all consumers, with the RLS-SP being the best method. It presents the smallest difference to the PI-RH. Considering that we re-optimize the solution after one time period, these results are aligned with the results in [72]. In [72], although the Copula-based forecast presented an overall better performance, the RLS-based forecast showed higher accuracy for the initial time period. This may indicate that the performance of the optimization depends to a high degree on the precision of the forecast in the initial time step.

A cost comparison between the different methods on a monthly level is shown in Figure B.6. The figure shows the additional cost incurred in each method in comparison to the PI-RH case as our baseline. Similar results as the aggregated results presented in Table B.3 are seen, where the RLS-SP and Copula-SP methods outperform the deterministic methods in every season. Moreover, the simulation results indicate that in some cases, the solutions found through SP methods tend to be more robust than those determined assuming perfect information. This can be seen in the RLS-SP results for consumer 1 in July, where RLS-SP solutions outperform those of PI-RH. This may seem counterintuitive considering that PI-RH assumes perfect knowledge of uncertain EL, however, the PI is limited by the 24 hours

Consumer no.	PI-RH DKK	%	Copula-EV DKK	%	Copula-SP DKK	%	RLS-EV DKK	%	RLS-SP DKK	%
1	5798.61	-	6220.66	7.28	5976.54	3.07	6143.84	5.95	5895.99	1.68
2	12142.80	-	12887.78	6.14	12553.15	3.38	12745.82	4.97	12438.41	2.43
3	6689.17	-	7259.67	8.53	6977.89	4.32	7398.95	10.61	6984.19	4.41
4	2071.31	-	2438.35	17.72	2210.56	6.72	2514.62	21.40	2183.16	5.40
5	1247.45	-	1638.04	31.31	1445.67	15.89	1695.82	35.94	1402.20	12.41
6	7435.47	-	8322.06	11.92	7887.49	6.08	8254.65	11.02	7824.03	5.23
7	9294.46	-	9945.68	7.01	9610.14	3.40	9903.44	6.55	9557.67	2.83
8	3157.25	-	3547.44	12.36	3312.33	4.91	3639.82	15.28	3316.83	5.05
9	6616.65	-	6970.68	5.35	6752.09	2.05	6954.68	5.11	6715.13	1.49

Table B.3: Total cost of the optimization and forecasting methods. All percentages are calculated relative to the PI-RH method.

in the rolling horizon. This behavior may indicate that considering the EL stochasticity may lead to more robust solutions towards the unpredicted EL and/or the optimization may benefit from a longer forecasting horizon.

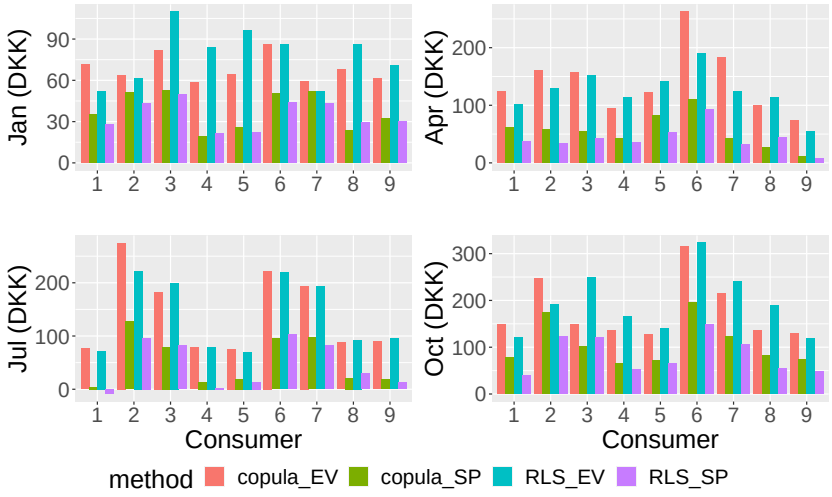

Figure B.6: Additional cost by forecast and optimization framework in comparison with the PI-RH as baseline.

B.4.2 Comparison of naive and optimal control

In this section, we present the operational results for the optimization-based control method (RLS-SP) and the naive controller. Figure B.7 shows the total load, total PV production, cost savings relative to the naive controller, total amount of electricity imported from the grid to the battery, the total power exported from the battery to the grid, and the average battery SOC for each consumer.

From the cost saving plots, we can observe significant differences between seasons. The RLS-SP outperforms the naive controller in winter (Jan) and autumn (Oct). In January, the additional cost savings can be explained by more load shifting. Here, the lower PV production available for self-consumption incentivizes more transactions with the grid, which can be seen by the higher grid-to-battery energy flow. Note that the battery-to-grid energy flow is non-significant. This indicates the optimization is taking advantage of the price volatility (see Table B.2) by charging during low-price hours in order

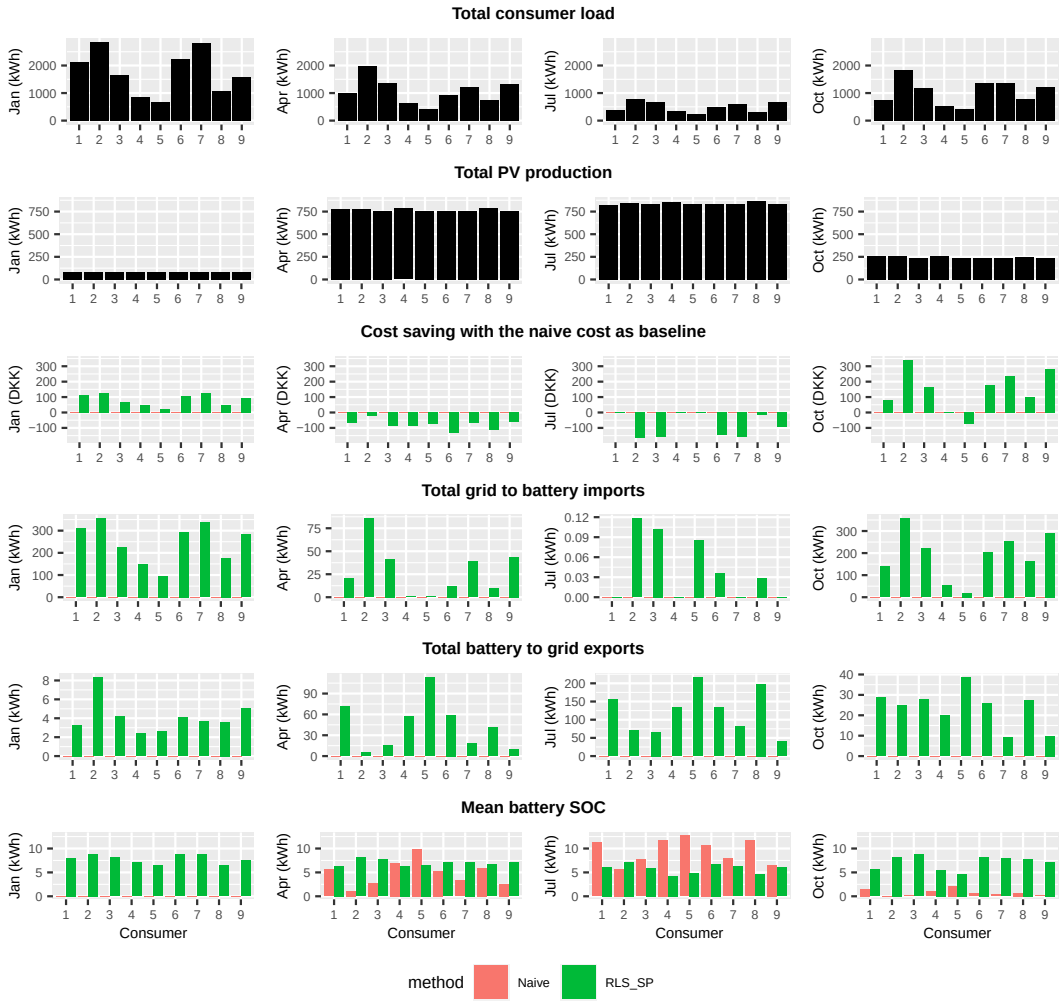

Figure B.7: Monthly cost and operation results for the naive controller and RLS-SP.

to use the stored electricity during high-price hours. October yields the highest cost savings of the simulated months. This is achieved by exploiting the high price volatility (the highest price volatility among the simulated months) in a similar fashion to January.

In contrast, spring (Apr) and summer (Jul) show the lowest cost savings. In these months, the RLS-SP controller is outperformed by the naive strategy. In particular, April presents a considerable EL with high PV production but low price volatility (see Table B.2). This implies that the optimization minimizes cost by selling excess PV generation, which could be a sub-optimal strategy in the long-term given the gap between purchase and sale electricity prices (see Figure B.4). In this case, the daily rolling optimization horizon might not be able to capture the longer-term effects and saving PV generation for posterior use will benefit the consumer the most. Furthermore, this argument matches the behavior seen in July, where the low EL, high PV production, and low price volatility leave almost no space for additional cost savings relative to the naive approach. Therefore, under these conditions, a naive approach is able to minimize consumers' costs in the long-term, and the application of more sophisticated optimization-based methods could have a negative impact on cost minimization.

B.4.3 Combination of naive controller and optimization

As we have seen previously, the naive controller (which maximizes self consumption) performs better than the optimization-based methods in the presence of high PV production and low EL. This behavior could be explained as a result of the price structure and volatility, and the implications of the finite optimization horizon. The optimization minimizes short-term cost by selling excess electricity, since it plans no more than 24 hours ahead. However, sale prices are very unfavorable in comparison with purchase prices, which makes self-consumption a better long-term strategy. Thus, one option could be to have a HEMS switching between a naive controller in the spring and summer and to use an optimization-based control strategy in the autumn and winter seasons. Hence, a full cost comparison between a passive consumer (without PV and battery), a naive controller, and the proposed strategy (Naive+RLS-SP) is presented in Table B.4. The results show that significant cost savings can be achieved in the naive and the proposed strategy in comparison with a passive consumer. In particular, the simulation results show that consumers with higher EL benefit most from installing the hardware and controllers. Moreover, the differences between the two control strategies show that the combined controller (Naive+RLS-SP) provides on average 8.05% additional savings for consumers with higher load (excluding consumers 4 and 5 with the lowest load of all consumers) in comparison with the naive controller, as can be seen in detail in Table B.5.

Consumer no.	Passive		Naive		Naive+RLS-SP	
	DKK	%	DKK	%	DKK	%
1	9219.20	-	6015.12	34.75	5825.17	36.81
2	16734.67	-	12715.96	24.01	12248.38	26.81
3	10799.02	-	6972.54	35.43	6735.15	37.63
4	5353.58	-	2140.78	60.01	2094.73	60.87
5	3901.01	-	1281.02	67.16	1332.69	65.84
6	11208.06	-	7827.33	30.16	7544.91	32.68
7	13370.53	-	9695.89	27.48	9333.01	30.20
8	6455.34	-	3331.51	48.39	3184.64	50.67
9	10806.51	-	6935.86	35.82	6563.42	39.26

Table B.4: Total cost of the simulated months for passive consumers, and naive and optimal (naive+RLS-SP) control strategies. All percentage values are calculated relative to the passive consumers' cost.

Consumer No.	Naive	Naive+RLS-SP	Difference	%
1	3204.09	3394.03	189.95	5.93
2	4018.70	4486.29	467.59	11.64
3	3826.48	4063.87	237.39	6.20
4	3212.80	3258.85	46.05	1.43
5	2619.99	2568.32	-51.67	-1.97
6	3380.73	3663.15	282.42	8.35
7	3674.64	4037.51	362.87	9.88
8	3123.83	3270.70	146.88	4.70
9	3870.65	4243.09	372.44	9.62

Table B.5: Naive and optimal control strategies (naive+RLS-SP) total cost savings and their difference in value and percentage.

B.5 Dicussion

The simulation results from Section B.4.1 show that EL uncertainty has a significant impact on the economic performance of the optimization-based strategies, with the stochastic methods outperforming

the deterministic methods. While it appears clear that including EL uncertainty increases the consumers' economic benefits, the lack of access to real PV data limits the ability to study the impact of modeling the uncertainty inherent to the PV generation on the system operation. To consider the PV uncertainty, one needs to account for the correlation between the PV and EL time series, which could be technically challenging in a multivariate setting. This could be a topic for future research.

A comparison between naive and optimal control strategies, presented in Section B.4.2, shows that the seasonal variations and the electricity price structure have a significant impact on the performance of the control strategies. During cold seasons with low PV production and higher EL, an optimization-based strategy is preferable in order to exploit price volatility by load shifting, charging the battery at low-price hours and discharging it during high-price hours. In warmer seasons with low price volatility, lower EL and PV make the naive strategy a better option. This could be explained by the fact that under these conditions, a long-term self-consumption strategy might be more profitable, which the optimization-based controller fails to capture due to its finite optimization horizon. Please note that this dynamic is tied to the price structure, where high differences between purchase and sale prices do not incentivize consumers to trade electricity. This may have a direct impact on the consumers in studies where market entities such as aggregators are present. Under current market conditions, aggregators will have to offer prices that surpass the retail prices. Otherwise, the consumers will not participate in the aggregation programs. Thus, studying different business models for aggregators under realistic price structures could be a line of future research.

The simulation results presented in Section B.4.3 indicate that the overall best control among the studied strategies is a combination of a naive and optimization-based controllers. This is shown by the proposed Naive+RLS strategy that outperforms the naive controller by 8.05% on average (excluding users 4 and 5). Please note that these extra savings result from a software update on the default system controller, which does not incur additional cost to the consumers. Additionally, the results obtained for customers 4 and 5 suggest a self-consumption strategy benefits the consumer the most when the EL is low, making more elaborated control strategies unnecessary in practice.

As smart grids continue to develop and price schemes evolve into a real-time structure, the proposed setup and control strategy shows that there is an incentive for consumers to adopt setups such as the one presented in this paper. Moreover, if we consider that a higher energy demand and a higher penetration of RES is expected in the residential sector, we could argue that these types of automation system might become a much needed tool for consumers to protect themselves against higher price volatility. Furthermore, although the proposed HEMS omits elements such as direct control of smart appliances and electric vehicles, we would expect their inclusion to bring higher savings for the consumer, which is a further element that could be explored in future research. While the economic results presented here show potential economical advantages of HEMS, the calculations were made for 2021 only. Future scenarios of electricity prices could be included and analyzed in future research in this area.

B.6 Conclusion

In this paper, a HEMS setup tailored for residential consumers under current electricity market rules in Denmark was presented. The HEMS was formulated as a MILP using stochastic programming, where the EL is treated as an uncertain parameter, which is modelled by multivariate probabilistic forecast models. Moreover, the HEMS formulation allows selling surplus electricity back to the grid depending on the price incentives and the re-optimization of the solution in a rolling horizon fashion.

The simulation results suggest that the optimization-based controllers, which consider the EL uncertainty, outperform their counterparts that consider only the expected value of the EL for all users and seasons investigated. Specifically, by comparing the RLS-SP controller with the naive one, we found that the seasonality of the data and prices have a profound impact on the cost-savings of all users. One

possible explanation could be that the finite optimization horizon and the gap between purchase and sale prices lead to sub-optimal solutions of the optimization-based controller in spring and summer, which are seasons with high PV production and low EL. In these seasons, a naive controller performs better. On the other hand, in autumn and winter with low PV production and high EL, optimization-based controllers exploit price volatility by shifting load to minimize electricity cost. Thus, the proposed combination of a naive and optimization-based strategy was shown to reduce the annual consumers' electricity cost in comparison with the naive method as a year-around controller. These savings are specifically significant considering that the proposed change of strategies in different seasons may come as a software update with small cost for the consumers. Finally, the case study showed that the proposed setup and control strategy offer a significant step towards a comprehensive assessment of the potential of HEMSs, indicating that the setup and control strategy could be developed in real-world applications of HEMSs.

acknowledgements

This work was partially funded by Innovation Fund Denmark through the project identified with case No. 8053-00156B.

PAPER C

Stochastic bi-objective home energy management without articulation of user preferences

Authors:
Amos Schledorn, Julian Lemos-Vinasco, Daniela Guericke, S. Ali Pourmousavi, Jon A. R. Liisberg, Henrik Madsen

Status:
In preparation to be submitted to IEEE transactions on smart grid

Stochastic bi-objective home energy management without articulation of user preferences

Amos Schledorn [1], Julian Lemos-Vinasco[1] [2], Daniel Guericke[1] S. Ali Pourmousavi[3], Jon A. R. Liisberg[2], Henrik Madsen[1]

Abstract

Smart buildings are a source of energy flexibility and will play an important role in the decarbonising of energy systems. Home energy management systems are often formulated as cost-minimising controllers. However, control that is entirely based on cost minimisation might fall short of emission reductions achievable when operating under an emission-minimising regime, while emission-minimising control might not be economically desirable for consumers. This motivates bi-objective setups. At the same time, such optimisation problems are subject to uncertainty in forecast data. In this paper, we propose a bi-objective extension to a stochastic cost-minimising home energy management system. The control problem is formulated as a stochastic program under uncertain demand which is forecast probabilistically. Non-dominated solutions to the bi-objective problem are generated through the ϵ-constraint method. We propose a method for selecting a compromise solution from the set of non-dominated solutions that explicitly takes uncertainty into account and is tailored towards stochastic programming. In a simulation study, both bi-objective and single-objective and single-objective controllers are evaluated in a rolling-horizon fashion under realistic market conditions using actual retail prices and measured demand. Our results suggest significant differences between optimising for either costs or emissions or bi-objectively, while we do not observe a significant impact of our proposed compromise solution selection method compared to traditional methods.

C.1 Introduction

The International Energy Agency estimates that COP26 pledges could lead to a 1.8 degree path, if implemented [14]. In order to meet these pledges, signatory governments face the challenge of integrating a large amount of renewable electricity into today's energy systems. Home energy management systems (HEMS) can play an important role in tackling this challenge: Buildings are responsible for 30 % of final energy consumption world-wide [56] and digital technologies will play a key role in integrating as much as 240 million PV systems and 1.6 billion electric cars by 2050 in a Net Zero Emissions scenario [56]. Here, smart buildings are an essential source of flexibility in energy systems [62].

In order for flexible demand to contribute to transitioning to low-carbon energy systems, appropriate incentives through market signals are essential [105]. In Denmark specifically, the weakness of correlation between price levels and carbon intensity can lead to different outcomes when controlling smart buildings with the aim of CO_2 versus cost minimisation [22]. Hence, if a smart home owner wishes to not only operate based on monetary, but also environmental objectives, a purely cost-minimising HEMS would not be sufficient. This raises the question of bi-objective HEMS optimising for low-cost and low-emissions simultaneously. On top of that, the nature of HEMS problems suggests an optimisation in a rolling horizon fashion, where forecasts and operational control are updated sequentially,

[1]Technical University of Denmark, Department of Applied Mathematics and Computer Science, Lyngby, Denmark
[2]Watts S/A, Section of Research and Development, Svinninge, Denmark
[3]The University of Adelaide, School of Electrical and Electronic Engineering, Adelaide, Australia

and uncertainty is modelled explicitly [REF single-obj paper]. In a bi-objective context, this requires a HEMS to not only take a set of possible solutions into account, but also select a solution that provides an acceptable trade-off between cost and emission minimisation. In this paper, we propose a solution approach for bi-objective stochastic HEMS in a rolling horizon.

The remainder of this paper is structured as follows: In Section C.2, we review related work and summarise our contributions. The proposed problem formulation and solution approach are described in Section C.3 and applied in a case study presented in Section C.4. Section C.5 concludes the paper and gives an outlook on future work.

Nomenclature

Sets	
$\mathcal{T}$	Set of time steps t
$\mathcal{S}$	Set of scenarios s
$\mathcal{M}$	Set of objectives m
Parameters	
$D_{s,t}$	Electricity load in scenario $s \in \mathcal{S}$ and time step $t \in \mathcal{T}$ [kWh]
λ_t^+	Electricity purchase price in time step $t \in \mathcal{T}$ [EUR/kWh]
λ_t^-	Electricity sale price in time step $t \in \mathcal{T}$ [EUR/kWh]
π_s	Probability of scenario $s \in \mathcal{S}$
Variables	
$x_{s,t}^+ \in \mathbb{R}^+$	Electricity purchased from the retailer in scenario $s \in \mathcal{S}$ and time step $t \in \mathcal{T}$ [kWh]
$x_{s,t}^- \in \mathbb{R}^+$	Electricity sold to the retailer in scenario $s \in \mathcal{S}$ and time step $t \in \mathcal{T}$ [kWh]
$b_{s,t} \in \mathbb{R}$	Battery discharge quantity in scenario $s \in \mathcal{S}$ and time step $t \in \mathcal{T}$ [kWh]
$p_t \in \mathbb{R}+$	Power output from the PV system in time step $t \in \mathcal{T}$ [kWh]
c	Total system costs [EUR]
e	Total system emissions [kg CO_2-eq]
Misc.	
f	vector of generic objective functions $f_1, f_2, ...f_{\lvert\mathcal{M}\rvert}$
$\mathcal{X}$	feasible space for generic decision variables $\mathbf{x}$

C.2 Related Work

We propose a HEMS optimisation model, which 1) is stochastic; 2) is multi-objective; 3) operates in a rolling horizon fashion; and 4) is formulated as a MIP or LP. While research on models that meet all four of these requirements seems to be limited, there exist papers which fulfil one or more of them. In the following, we briefly review a selection of them and refer the reader to reference [71] for an extensive review on HEMS literature, touching uncertainty handling and multi-objective approaches.

C.2.1 Single-objective HEMS

Single-objective HEMS planning problems are solved through a wide range of optimisation approaches and heuristic methods. For instance, the authors of [94] use time series analysis to forecast PV production and a deterministic single-objective HEMS is applied in a rolling horizon fashion. In [29], a convex optimisation model is used to solve a deterministic HEMS problem and an autoregressive model to forecast photovoltaic generation and electric load in a rolling horizon. The authors of [128] propose an online queuing algorithm to solve a HEMS problem. Besides exact solution methods to HEMS problems, there exists research on heuristic methods: For instance, the authors of [47] compare different

meta-heuristic approaches to solving a model-predictive control problem for a Canadian HEMS and conclude particle swarm optimisation to be the most suitable.

Applications of stochastic programming to HEMS include reference [108], where costs are minimised in a single-objective stochastic HEMS while setting an upper bound on user dissatisfaction, and reference [21], where costs are minimised under uncertain electric load and photovoltaic generation.

C.2.2 Multi-objective HEMS

We give a brief introduction to multi-objective optimisation in Appendix C.3. In this section, we focus on solution methods applied in HEMS with multiple objectives.

C.2.2.1 Scalarising approaches

One way to accommodate a vector of multiple objective functions is to combine them in a scalar function. For instance, a scalarised objective function combining costs and user comfort in a deterministic optimisation model is minimised in reference [5]. The authors of reference [60] propose a HEMS framework where a single-objective model-predictive control problem is solved based on weightings of user preferences of different objectives. In reference [131], a stochastic HEMS problem optimising for the weighted sum of costs and user comfort is solved using particle swarm optimisation. The authors of reference [99] compare different heuristic approaches to solving a multi-objective HEMS problem minimising the weighted sum of their objectives. An example of risk modelling in HEMS is the work in reference [58], where a stochastic program with a conditional value-at-risk is applied to minimise the weighted sum of costs and discomfort.

C.2.2.2 Genetic algorithms

Applications of genetic algorithms include the work of Veras et al. [115], where a multi-objective non-linear deterministic HEMS problem through NSGA-II and reference [41], where a stochastic multi-objective micro-grid management problem is solved. On the other hand, Zupančič et al. [133] find that a decision-tree-based approach outperforms a genetic algorithm in solving a deterministic bi-objective HEMS problem (cost/self-sufficiency). Other meta-heuristics are used in references [119] and [3], the approach in the latter article also proposing an analytical hierarchy process to select a solution from the Pareto-optimal front.

C.2.2.3 Epsilon-constraint programming and lexicographic methods

ϵ-constraint programming, which is used in our study and described in detail in Appendix C.3, is used by Javadi et al. [57] to construct the Pareto-optimal front of a bi-objective HEMS problem comprising cost and user discomfort minimisation and applying a fuzzy decision making approach to select a trade-off solution. The authors apply a similar approach in reference [59]. Nikkhah et al. [88] meet some of the requirements outlined in the introduction of this section by applying information gap theory and ϵ-constraint programming to optimise building flexibility for system costs and user comfort in a rolling horizon-based approach. Another solution method to multi-objective HEMS problems are lexicographic methods (e.g. [113]).

C.2.3 Selection of Compromise Solution from the Pareto Front

We simulate the real-world application of our proposed HEMS controller in a rolling horizon fashion. This requires a technique to select a compromise solution from the front of Pareto-optimal solutions. Oprivic and Tzeng [91] define the compromise solution as "[...] a feasible solution, which is the closest to the ideal [...]", where the ideal is the Utopian point. Wang and Rangaiah [121] give an overview on methods for selecting such compromise solutions including Linear Programming Technique for Multi-dimensional Analysis of Preference (LINMAP), which we apply in this study based on [2] with slight adjustments (c.f. Appendix C.3). While several methods exist for selecting compromise solutions, research on approaches tailored towards stochastic programming seems to be limited. [1] We also propose an extension of the method that takes stochasticity into account tailored to stochastic programming.

We refer the reader to [121] for an extensive methodological overview for solution selection in multi-objective optimisation.

C.2.4 Contribution

Based on this research gap, the following contributions are made:

- We propose a bi-objective extension of a stochastic HEMS model that minimises costs and emissions, but could be applied to other objectives.
- We propose a solution approach based on stochastic ϵ-constraint programming that selects and applies a compromise solution in a rolling horizon fashion.
- We apply the aforementioned contributions in a realistic case study on eight users in Denmark using probabilistic load forecasting and measured consumption data.

C.3 Methodology

The methodology in this paper is based on the work in reference [73], where a cost-minimising HEMS controller based on stochastic programming and probabilistic forecasting is proposed and analysed extensively. Here, we only give a brief overview of this controller (Section C.3.2.1) and refer to [73] for a detailed formulation.

In this study, the controller is extended by a bi-objective component minimising for both costs and carbon emissions (Section C.3.2.2). Since bi-objective programming produces a set of non-dominated solutions rather than a single optimum, implementation in practice requires the selection of one solution from this set. Hence, we propose a compromise solution selection method that is tailored to stochastic programming (Section C.3.3). Consumer demand scenarios are generated via probabilistic forecasting [72] and reduced via partitioning around medoids (Section C.3.4).

C.3.1 Overview

We assume a HEMS problem of optimally controlling household battery operation under uncertain electric load. This control takes electricity generation from a rooftop PV system as well as electricity trading into account. Trading comprises both electricity purchases and sales at time-varying prices, whereas price levels can be different when selling or purchasing (in our study, sale price levels are substantially lower since they do not include a set of tariffs and fees purchase prices do). These prices

[1] A review of related work is out of the scope of this paper. Notably, a Scopus search query for `"LINMAP" AND "STOCHASTIC PROGRAM`, `"TOPSIS" AND "STOCHASTIC PROGRAM"` or `"VIKOR" AND "STOCHASTIC PROGRAM"` in article title, abstract or keywords did not yield any results. VIKOR and TOPSIS are methods similar to LINMAP.

are supplied by an electricity retailer. Hence, the controller faces no uncertainty in this regard.

The controller operates at limited foresight, i.e. optimising only for a rolling time window $\mathcal{T}$. In a given time step t_0, electric load scenarios are generated $|\mathcal{T}|$ time steps ahead and electricity prices are sent by the retailer for the same time frame. For the window $\mathcal{T}$, photo-voltaic generation levels and emission intensity in the electricity mix are assumed fully known. After an optimal solution is determined, only the solution for the initial time step is implemented, as all subsequent decisions can still be adapted when the most recent information is available. That means that all decision variables but battery operation in the current time step only serve as a means of optimising battery operation in that time step. After the implementation of these decisions, the optimisation window is shifted by one time step.

C.3.2 Optimisation models

C.3.2.1 Simplified single-objective problem

In this study, the HEMS planning program formulated in reference [73] is applied and extended by emission minimisation. In a condensed version, the single-objective problem of minimising either operational costs c or emissions e can be formulated as (C.1). In relation to the formulation in reference [73], we introduce e alongside a binary parameter α. This parameter allows us to write both single-objective problems in one program, where $\alpha = 1$ leads to cost minimisation and $\alpha = 0$ minimises emissions ((C.1a)). An important notion is that α is not so much a weighting factor to scalarise a vector of objectives but rather a switch between cost and emission minimisation.

The problem is formulated as a stochastic mixed-integer program minimising the expected value of either objective over a horizon $\mathcal{T}$. Electricity demand $D_{s,t}$ in scenario s and time step t is the only uncertain parameter in the two-stage stochastic program. The probability of scenario $s \in \mathcal{S}$ is denoted by π_s. The system consists of a battery storage system and a rooftop photo-voltaic unit.

$$\min \quad \alpha c + (1-\alpha)e \tag{C.1a}$$

$$\text{s.t.} \quad c = \sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}} \pi_s(\lambda_t^+ x_{s,t}^+ - x_{s,t}^- \lambda_t^-) \tag{C.1b}$$

$$e = \sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}} \pi_s e_t x_{s,t}^+ \tag{C.1c}$$

$$x_{s,t}^+ - x_{s,t}^+ + p_t + b_{s,t} = D_{s,t} \quad \forall t \in \mathcal{T}, s \in \mathcal{S} \tag{C.1d}$$

$$b_{s,t} = b_{s',t} \quad t = 1, \forall s \in \mathcal{S}, s\prime \in \mathcal{S} \setminus s \tag{C.1e}$$

Operational costs c are the difference between costs for power purchase $x_{s,t}^+$ at price $\lambda_{s,t}^+$ and electricity sales $x_{s,t}^-$ at price $\lambda_{s,t}^-$ ((C.1b)). Total emissions e are the expected value of hourly emissions, i.e. the product of scenario emission intensity e_t and power purchase $x_{t,s}^+$ ((C.1c)). The energy balance is modelled by (C.1d), which ensures that demand $D_{s,t}$ is covered by the sum net electricity purchases, PV power output p_t and battery discharge $b_{s,t}$. Battery operation in the initial time step constitutes the first-stage decision, which must be equal across scenarios ((C.1e)).

In addition to the formulation in (C.1), the model updates the battery state of charge, disallows simultaneous charging and discharging as well as simultaneous power purchase and sale. It also accounts for inversion and conversion losses. The model is described in detail in reference [73]. In the reminder of the

paper, we refer to the full model including the above mentioned parts, when referring to (C.1b) to (C.1e).

C.3.2.2 Bi-objective problem

In single-objective optimisation problems, the aim is the minimisation of a scalar objective function $f \in \mathbb{R}$ over decision variables $\mathbf{x} \in \mathcal{X}$. In multi-objective optimisation, the assumption of f being one-dimensional is dropped, such that it is a vector of objective functions $\mathbf{f} \in \mathbb{R}^M$ with M being the cardinality $|\mathcal{M}|$ of the set of objectives $m \in \mathcal{M}$. This problem can be written in a generic form as [82]:

$$\min \quad \mathbf{f}(\mathbf{x}) \tag{C.2a}$$
$$s.t. \quad \mathbf{x} \in \mathcal{X} \tag{C.2b}$$

Here, optimisation refers to finding solutions $\mathbf{x}^* \in \mathcal{X}$ that are Pareto-optimal. A solution $\mathbf{x}^* \in \mathcal{X}$ is considered Pareto-optimal if it is not dominated by any other solution $\mathbf{x}' \in \mathcal{X}$ in the sense that $\mathbf{f}(\mathbf{x}') \in \mathcal{X}$ yields a lower objective value (for minimization) than $\mathbf{f}(\mathbf{x}^*) \in \mathcal{X}$ for at least one $m \in \mathcal{M}$ and is not worse for all other $n \in \mathcal{M} \setminus \{m\}$ [25]. Important terminology in multi-objective optimisation includes the Pareto-front (the set of all Pareto-optimal solutions [18]), Nadir point (the vector of upper bounds on each objective in the set of Pareto-optimal objective space) and the Utopian point (the vector of lower bounds on each objective in the set of Pareto-optimal objective space) [25].

In this study, the bi-objective problem of minimising operational costs and carbon emissions can be formulated as (C.3). As described in **??**, the model features constraints in addition to (C.1b) to (C.1e), which are formulated in reference [73].

$$\min \quad [c, e] \tag{C.3a}$$
$$s.t. \quad (C.1b) to\ (C.1e) \tag{C.3b}$$

C.3.2.3 Solution approach

There exists a plethora of methods to determine a set of Pareto-optimal solutions (c.f. [82] for an overview). The ϵ-Constraint method [48] is an iterative approach, where in each iteration i, one objective $n \in \mathcal{M}$ is optimised while all other objectives $m \in \mathcal{M} \setminus \{n\}$ are bound by a parameter ϵ_m.

In this paper, the ϵ-Constraint method is applied to solve the problem formulated in (C.3) to explore the solution space of a bi-objective minimisation of total costs $c \in \mathbb{R}$ and carbon emissions $e \in \mathbb{R}^+$. In our solution approach (Algorithm 3), the application of the ϵ-constraint method proceeds as follows: Iteratively, the mixed-integer linear program formulated in (C.1) is solved with decreasing upper bounds on the emissions e. These bounds represent the ϵ-constraints and are linearly decreasing within a pre-defined interval. This interval ranges from the highest reasonable emission value (the emission level when optimising for costs only) to the lowest possible emission value (the emission level when optimising for emissions only).

As this sets the range of reasonable values for e, the approach allows to determine an adequate representation of the Pareto-front. Here, $I = 20$ iterations are chosen leading to I solutions on the Pareto-front $\mathbf{X} = \{x_1, ..., x_i, ..., x_I\}$.

Algorithm 3 ϵ-constraint implementation

1: Solve (C.1) for $\alpha = 1$ and save objective values c^+ and e^-
2: Solve (C.1) for $\alpha = 0$ and save objective values c^- and e^+
3: **for** i = 1, ..., I **do**
4: Solve (C.1) with bound on emissions: $e \leq e^- - \frac{i}{I}(e^- - e^+)$
5: **end for**

C.3.3 Solution selection from Pareto-front

After the optimisation problem is solved, one first-stage solution (here battery operation in the initial time step) must be chosen from the Pareto front to be implemented. This raises the question of how to determine an appropriate compromise solution between the two objectives. An overview of such compromise solution selection methods is given in reference [121]. We propose an extension of LINMAP to stochastic programming. LINMAP selects that solution from the Pareto-front that yields the lowest normalised distance to the Utopian point. Here, we follow the formulation of Ahmadi et al. [2] with the exception that objective values are minimised by the Utopian point. That means that we apply that solution $\mathbf{x_i}$ among all solutions on the Pareto-front $\mathbf{X}$, for which:

$$\sqrt{\sum_{n=1,...,M} \left(\frac{f_n(\mathbf{x_i}) - f_n(\mathbf{x^U})}{f_n(\mathbf{x^U})}\right)^2} \leq \sqrt{\sum_{n=1,...,M} \left(\frac{f_n(\mathbf{x_j}) - f_n(\mathbf{x^U})}{f_n(\mathbf{x^U})}\right)^2} \quad \forall j = 1, ..., I \tag{C.4}$$

We extend that metric to stochastic programming, such that the selected solution minimises the distance between the set of objective values corresponding to the selected solution $\mathbf{x_i}$ in scenario $s \in S$ and the discrete distribution of objective values in the Utopian point, i.e. select that solution $\mathbf{x_i}$, for which:

$$\sqrt{\sum_{n=1,...,M, s \in S} \left(\frac{f_{n,s}(\mathbf{x_i}) - f_{n,s}(\mathbf{x^U})}{f_n(\mathbf{x^U})}\right)^2} \leq \sqrt{\sum_{n=1,...,M, s \in S} \left(\frac{f_{n,s}(\mathbf{x_j}) - f_{n,s}(\mathbf{x^U})}{f_{n,s}(\mathbf{x^U})}\right)^2} \quad \forall j = 1, ..., I \tag{C.5}$$

Hence, the bi-objective controller applied in a given time step can be formulated as Algorithm 4.

Algorithm 4 Bi-objective controller

1: Apply Algorithm 3, save Utopian point $[c^+, e^+]$ & Nadir point $[c^-, e^-]$
2: Determine compromise solution $\mathbf{x}^*$ based on (C.4) or (C.5)
3: Implement battery operation corresponding to $\mathbf{x}^*$
4: Move optimisation window by one time step and apply controller

In the reminder of this paper, we refer to the deterministic solution selection based on (C.4) as LINMAP and to the stochastic solution selection based on (C.5) as S-LINMAP. Notably, the extension from (C.4) to (C.5) is applicable to other techniques for selecting a compromise solution aside from LINMAP.

C.3.4 Scenario Generation and Reduction

Electric load is the only uncertain parameter in this study. Scenarios are generated by a probabilistic forecasting model which considers the temporal correlation between forecast values, as introduced in reference [72]. In particular, the model applies recursive least squares estimation (RLS), which allows to update model parameters when new observations become available [76]. Results in reference [72] show that considering the temporal correlation can significantly improve of multivariate distribution forecasts. This is supported by the results in reference [73], where the RLS method outperformed other forecast methods when it was applied to a HEMS.

To ease the computational burden imposed by multiple iterations in the ϵ-constraint method, we first generate representative scenarios using partitioning around medoids [64] following the approach in reference [15].

C.4 Case Study

C.4.1 Case Study Setup

The setup of our case study follows a similar structure as the experiments in [73]. Hence, only a condensed overview is given here and we refer to [73] for a detailed description.

C.4.1.1 General Setup

Based on electricity load measurements for nine customers of the Danish Power retailer Watts A/S [124], the proposed modelling framework is tested for four months from different seasons (January, April, July, October) in the year 2020. We simulate the application of four different controllers, namely cost minimisation, CO_2 minimisation and a bi-objective model applying both a deterministic and stochastic solution selection from the Pareto front. The simulations follow a rolling horizon approach, where in each hour, the optimisation problem is solved looking 24 hours ahead and the state of charge of the home's battery system is set by the operation in the hour before, i.e. the only first-stage decision in the stochastic program.

C.4.1.2 PV Generation Data

We assume all home systems in this study to be equipped with a rooftop PV system with a capacity of 6.3 kWp [24]. As no data on PV power output is available, simulated data is used following the same approach as in [72]: Solar irradiance is simulated based on the location of the respective smart homes using the implementation in pvlib [53]. Then, based on the relationship between solar irradiance and power output given in [24], generation profiles are computed. The impact of cloud cover is modelled based on [70].

C.4.1.3 Power Prices

We apply the same power purchase prices as Watts A/S, which follow Danish day-ahead prices [44] adjusted by taxes, tariffs and fees. Sale prices are assumed equal to spot market prices [44].

C.4.1.4 Emissions

Similarly to power prices, emissions are assumed to be supplied by the retailer and fully known to the HEMS. Here, they are supplied by Watts A/S, who makes estimates based on the power generation mix in Denmark [125].

C.4.1.5 Implementation

While electric load forecasts were implemented in R using the onlineforecast package [9], scenario reduction and optimisation were implemented in Python 3.8.2. For scenario reduction by means of partitioning around medoids, we used the implementation of K-medoids clustering in scikit-learn-extra [106]. The optimisation was implemented using the package MIP [40] and solved with Gurobi 9.1.0 [46] and standard settings on DTU's high-performance cluster computer using 2 Intel Xeon 2660v3 processors with 2.6 GHz and 32 GB of RAM.

C.4.2 Numerical Results

We first present aggregated results to compare the four different controller setups (cost minimisation, emission minimisation and bi-objective optimisation using both LINMAP and S-LINMAP). This is followed by a illustrations of the bi-objective models for an exemplary user and time step.

C.4.2.1 Aggregated Comparison of Controller Setups

The four controller configurations have been applied for all users and months and first-stage decisions were saved. Then, the solutions were evaluated on the same optimisation model with fixed first-stage decisions and electricity load observations instead of forecast scenarios. Both levelised costs of electricity (LCOE) and per-MWh CO_2 emissions have been computed per MWh alongside cost/emission minimisation under perfect information (PI), meaning perfect foresight of the entire optimisation horizon (744h) without a rolling horizon approach. These are theoretical lower bounds on the respective objectives and we refer to them as PI costs and PI emissions in the following.

On average, PI costs amount to 165.7 EUR/MWh (Table C.1). The cost minimising controller exceeds this value by 10.9 EUR/MWh ranging from 121.9 EUR/MWh for user 5 to 198.3 EUR/MWh for user 2 (Table C.1). When minimising emissions the difference to the PI costs increases to 19.6 EUR/MWh on average with a similar pattern across users. Unsurprisingly, electricity bills under the bi-objective optimisation models are in-between the single-objective results at 179.1 EUR/MWh with no significant difference between the two solution selection methods.

Table C.1: Average LCOE in EUR/MWh across users under perfect information (cost minimising) and across controller setups (total costs across all months divided by total demand across all months).

User	PI	Bi-objective		Single-objective	
	Cost min.	LINMAP	S-LINMAP	Cost min.	CO_2 min.
1	170.3	181.9	181.8	179.2	188.4
2	190.9	200.6	200.6	198.3	206.0
3	163.0	176.7	176.8	174.3	182.0
4	108.1	127.6	127.5	124.8	134.3
5	96.2	125.0	124.9	121.9	134.4
6	177.2	192.3	192.3	190.4	198.4
7	184.8	196.2	196.1	193.8	201.4
8	135.6	153.8	153.7	150.0	161.4
9	161.2	171.7	171.6	168.7	178.1
Mean	165.7	179.1	179.1	176.6	185.3

Total emissions (Table C.2) follow the same structure as total costs (Table C.1): While average emissions are highest under a cost minimising setup (112.7 kg CO_2-eq/MWh), CO_2 levels are lowest when optimising for emissions at 107.6 kg CO_2-eq/MWh. Under either of the bi-objective controllers, emissions are slightly higher at 108.1 kg CO_2-eq/MWh. On average, minimising emissions in a single-objective program exceeds PI emissions by 9.3%.

Table C.2: Average carbon emissions in kg CO_2-eq/MWh across users under perfect information (cost minimising) and across controller setups (total emissions across all months divided by total demand across all months).

User	PI	Bi-objective		Single-objective	
	CO_2 min.	LINMAP	S-LINMAP	Cost min.	CO_2 min.
1	107.5	115.0	114.9	119.6	114.5
2	111.8	118.4	118.4	123.2	117.7
3	95.5	105.2	105.2	109.4	104.4
4	66.2	79.8	79.8	83.9	79.8
5	64.5	83.9	83.9	87.6	84.8
6	105.5	116.1	116.0	120.2	115.8
7	110.3	118.1	118.0	122.7	117.3
8	80.4	93.5	93.4	97.6	93.1
9	92.8	100.0	100.0	105.7	99.1
Mean	98.8	108.1	108.1	112.7	107.6

Aggregating savings monthly instead of user-wise yields similar observations (Tables C.3 and C.4). Notably, in April and July, Total emissions under the emission-minimising controller are higher than those under either bi-objective controller (Table C.4). These are the months with the highest electricity generation from PV (and consequently a lower, but volatile residual demand to be covered by grid imports). Hence, a possible explanation could be that volatile carbon intensity in the electricity mix (Figure C.1) in the light of uncertain and volatile residual demand triggers higher import volumes (Table C.5).

Table C.3: Average LCOE in EUR/MWh across months under perfect information (cost minimising) and across controller setups (total costs across all users divided by total demand across all users).

Month	PI	Bi-objective		Single-objective	
	Cost min.	LINMAP	S-LINMAP	Cost min.	CO2 min.
Jan.	242.7	248.2	248.2	244.7	255.1
Apr.	98.1	115.1	115.0	111.9	120.5
Jul.	-11.3	25.0	24.9	24.7	30.3
Oct.	189.6	202.2	202.1	200.7	208.1
Mean	165.7	179.1	179.1	176.6	185.3

Table C.4: Average carbon emissions in t CO_2-eq/MWh across months under perfect information (cost minimising) and across controller setups (total emissions across all users divided by total demand across all users).

Month	PI	Bi-objective		Single-objective	
	CO_2 min.	LINMAP	S-LINMAP	Cost min.	CO2 min.
Jan.	156.3	160.2	160.2	167.1	158.6
Apr.	64.2	79.0	78.9	81.0	80.2
Jul.	3.9	19.2	19.2	19.9	19.9
Oct.	82.7	92.6	92.6	97.7	91.4
Mean	98.8	108.1	108.1	112.7	107.6

Table C.5: Average purchase intensity in percentage across users under perfect information and across controller setups (total purchase volume across all users divided by total demand across all users).

Month	PI		Bi-objective		Single-objective	
	Cost min.	CO_2 min.	LINMAP	S-LINMAP	Cost min.	CO2 min.
Jan.	95.9	96.7	96.5	96.5	95.9	97.2
Apr.	43.2	44.0	49.6	49.5	48.4	51.3
Jul.	6.3	6.4	19.7	19.7	20.1	21.1
Oct.	76.3	76.8	78.2	78.2	78.3	79.4

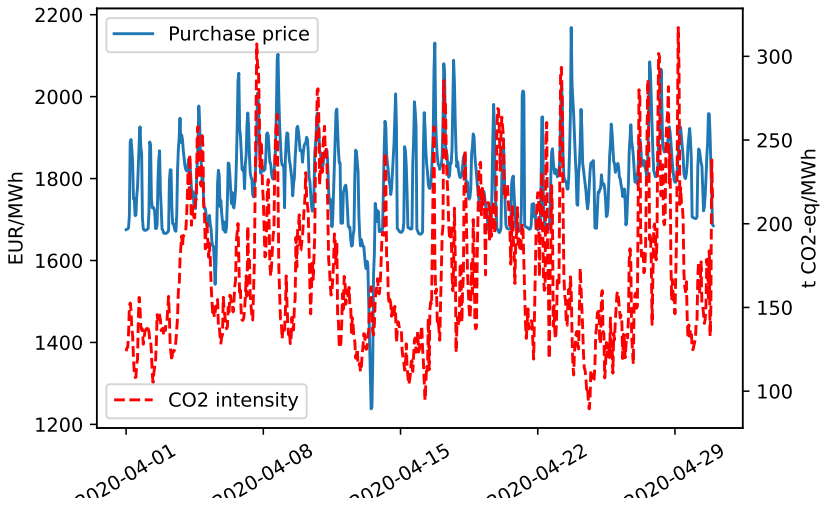

Figure C.1: Purchase prices and carbon intensity in April 2020.

C.4.2.2 Exemplary control trajectory

Operational decisions under the different control regimes are shown in Figure C.2 alongside purchase prices and carbon intensity in the electricity mix of electricity purchases. The bottom plot shows (for this example) different shapes for electricity prices and carbon intensity, underlining the potential for a bi-objective control. While the two bi-objective controllers do not show a significantly different behaviour, it clearly showcases differences between the two single-objective models: Decisions are roughly similar until ca. the beginning of the third day, a rise in carbon intensity during that day triggers electricity purchases of the emission minimising controller charging the battery and discharging it ca. 12 hours later. The cost-minimising controller, on the other hand, makes suggest to purchase electricity during the following price valley, when emissions are peaking. Even though not being entirely evident from this figure, one could argue that the behaviour of either bi-objective controller falls in-between these two extremes.

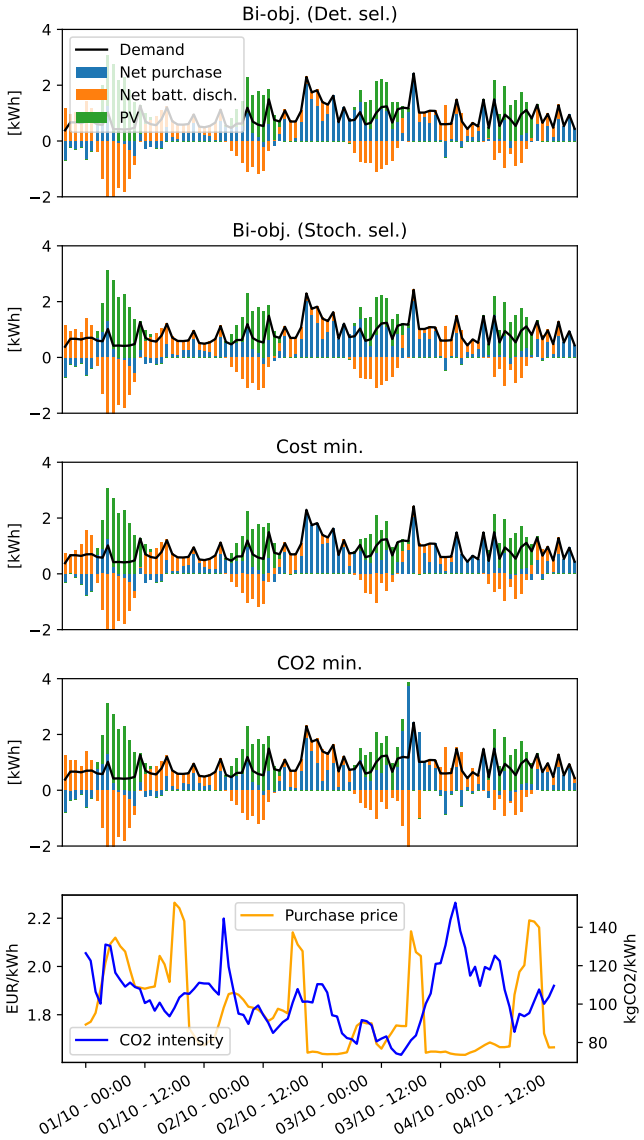

Figure C.2: Controller decisions in the first 96 hours for user 9 in October.

C.4.2.3 Exemplary Deterministic and Probabilistic Pareto-Fronts

In the following, we illustrate a probabilistic and deterministic Pareto-fronts exemplarily for user 3 in October, time step 151. The objective values of Pareto-optimal solutions in a bi-objective problem are traditionally plotted as a convex front. In the case of stochastic programming, we can extend this to a probabilistic Pareto-front. This is illustrated in Figure C.3. The solid black line indicates the deterministic front, i.e. the expected value of costs and emissions for each solution found via the ϵ-constraint method. A scatter plot of costs and emissions for each scenario and solution is added. Here, the color denotes the respective solution and the point size indicates the scenario probability. We can clearly see that both the scenario and iteration counter give shape to the plot, as the shape of the deterministic front is roughly repeated for each scenario, though shifted. At the same time, the objective values of individual solutions appear to rank similarly across scenarios.

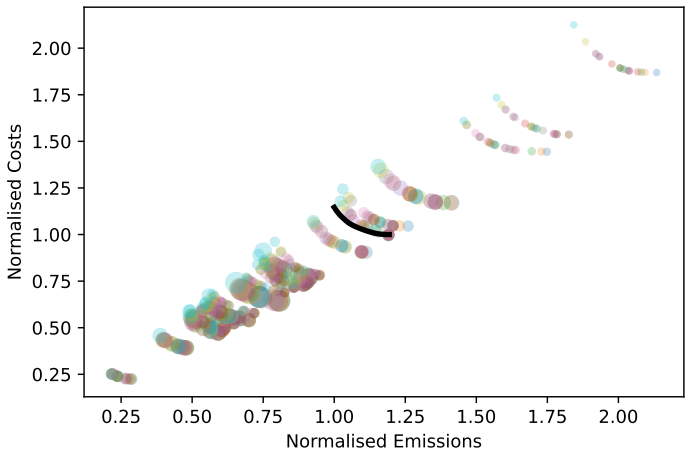

Figure C.3: Exemplary deterministic and probabilistic Pareto-front for user 3, October, time step 151. Point sizes indicate scenario probabilities. The black line shows the deterministic front. For illustration purposes, only 25 scenarios are displayed.

While Figure C.3 illustrates the shape of the probabilistic Pareto-fronts in each scenario, it lacks a representation of the individual solution across scenarios - which would be of interest for a decision maker. This is illustrated in Figure C.4 where the distributions of the objective values of three solutions (instead of twenty, to enhance readability) have been approximated via Kernel Density Estimation with scenario probabilities as weights [122].

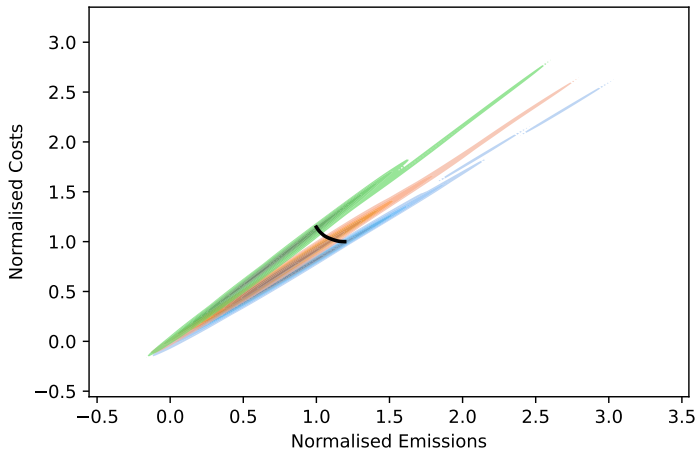

Figure C.4: Exemplary deterministic and probabilistic Pareto-front for user 3, October, time step 151. The black line indicates the deterministic front.

C.5 Conclusion

In this study, we have proposed a solution approach for HEMS using a bi-objective two-stage stochastic program under uncertain electric load and applied it in a rolling horizon.

We solved the above mentioned program using ϵ-constraint programming [48] and proposed a method for selecting a compromise solution from the set of Pareto-optimal solutions that is tailored towards stochastic HEMS optimisation problems. We extend the implementation of LINMAP in [2] by a stochastic component, which we refer to as S-LINMAP.

An application was simulated under real-world data for nine Danish households for one month across four seasons each. Our results suggest a significant difference in carbon emissions and system costs if optimising solely for either objective or selecting a compromise solution. This means that consumers face the decision between a low-cost or low-emission approach in their HEMS. This observation contradicts the conception of Danish electricity prices providing a sufficient indirect control signal to incentivise shifting consumption to low-emission periods.

While advocating for a probabilistic representation of the Pareto front, i.e. considering the entire distribution of the vector of objective values, we cannot observe significant difference in either cost or emission levels in the application of LINMAP or S-LINMAP.

With regard to this analysis, our study has the clear shortcoming of being both problem- and domain-specific. Hence, we suggest the application of S-LINMAP, or stochastic extensions to other compromise solution selection methods, to further HEMS problems and problems in different domains as a future line of research.

Acknowledgment

This work was partially funded by Innovation Fund Denmark through the project identified with case No. 8053-00156B.

Bibliography

[1] Ahmed G. Abo-Khalil et al. "Electric vehicle impact on energy industry, policy, technical barriers, and power systems." In: *International Journal of Thermofluids* 13 (February 2022), page 100134. ISSN: 2666-2027. DOI: 10.1016/J.IJFT.2022.100134.

[2] Mohammad H. Ahmadi et al. "Application of the multi-objective optimization method for designing a powered Stirling heat engine: Design with maximized power, thermal efficiency and minimized pressure loss." In: *Renewable Energy* 60 (2013), pages 313–322. ISSN: 09601481. DOI: 10.1016/j.renene.2013.05.005. URL: http://dx.doi.org/10.1016/j.renene.2013.05.005.

[3] Masoud Alilou, Behrouz Tousi, and Hossein Shayeghi. "Home energy management in a residential smart micro grid under stochastic penetration of solar panels and electric vehicles." In: *Solar Energy* 212 (2020), pages 6–18. ISSN: 0038092X. DOI: 10.1016/j.solener.2020.10.063. URL: https://doi.org/10.1016/j.solener.2020.10.063.

[4] Kasun Amarasinghe, Daniel L. Marino, and Milos Manic. "Deep neural networks for energy load forecasting." In: *IEEE International Symposium on Industrial Electronics*. Institute of Electrical and Electronics Engineers Inc., August 2017, pages 1483–1488. ISBN: 9781509014125. DOI: 10.1109/ISIE.2017.8001465.

[5] Amjad Anvari-Moghaddam, Hassan Monsef, and Ashkan Rahimi-Kian. "Optimal smart home energy management considering energy saving and a comfortable lifestyle." In: *IEEE Transactions on Smart Grid* 6.1 (2015), pages 324–332. ISSN: 19493053. DOI: 10.1109/TSG.2014.2349352.

[6] Donald Azuatalam et al. "Energy management of small-scale PV-battery systems: A systematic review considering practical implementation, computational requirements, quality of input data and battery degradation." In: *Renewable and Sustainable Energy Reviews* 112 (September 2019), pages 555–570. ISSN: 1364-0321. DOI: 10.1016/J.RSER.2019.06.007.

[7] Peder Bacher, Hjörleifur G. Bergsteinsson, and Henrik Madsen. *onlineforecast: An adaptive forecasting package in R*. URL: https://onlineforecasting.org/index.html (visited on 12 March 2022).

[8] Peder Bacher, Henrik Madsen, and Henrik Aalborg Nielsen. "Online short-term solar power forecasting." In: *Solar Energy* 83.10 (October 2009), pages 1772–1783. ISSN: 0038-092X. DOI: 10.1016/J.SOLENER.2009.05.016.

[9] Peder Bacher et al. "onlineforecast: An R package for adaptive and recursive forecasting." In: *Arxiv preprint* (2021), pages 1–36. URL: http://arxiv.org/abs/2109.12915.

[10] Peder Bacher et al. "Short-term heat load forecasting for single family houses." In: *Energy and Buildings* 65 (October 2013), pages 101–112. ISSN: 0378-7788. DOI: 10.1016/J.ENBUILD.2013.04.022.

[11] Peder Bacher et al. "Short-term heat load forecasting for single family houses." In: *Energy and Buildings* 65 (October 2013), pages 101–112. ISSN: 03787788. DOI: 10.1016/j.enbuild.2013.04.022.

[12] R. Balakrishnan and V. Geetha. "Review on home energy management system." In: *Materials Today: Proceedings* 47 (January 2021), pages 144–150. ISSN: 2214-7853. DOI: 10.1016/J.MATPR.2021.04.029.

[13] John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer New York, 2011. DOI: 10.1007/978-1-4614-0237-4. URL: http://link.springer.com/10.1007/978-1-4614-0237-4.

[14] Fatih Birol. *COP26 climate pledges could help limit global warming to 1.8 °C, but implementing them will be the key*. 2021. URL: https://www.iea.org/commentaries/cop26-climate-pledges-could-help-limit-global-warming-to-1-8-c-but-implementing-them-will-be-the-key.

[15] Ignacio Blanco et al. "Operational planning and bidding for district heating systems with uncertain renewable energy production." In: *Energies* 11.12 (2018). ISSN: 19961073. DOI: 10.3390/en11123310.

[16] John Bjørnar Bremnes. "Probabilistic wind power forecasts using local quantile regression." In: *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology* 7.1 (2004), pages 47–54.

[17] Joilson de Assis Cabral, Luiz Fernando Loureiro Legey, and Maria Viviana de Freitas Cabral. "Electricity consumption forecasting in Brazil: A spatial econometrics approach." In: *Energy* 126 (May 2017), pages 124–131. ISSN: 03605442. DOI: 10.1016/j.energy.2017.03.005.

[18] Carlos A. Coello Coello. "A Short Tutorial on Evolutionary Multiobjective Optimization." In: *Evolutionary Multi-Criterion Optimization*. Edited by E. Zitzler et al. Springer, 2001. ISBN: 3540417451.

[19] European Commission. *EU Energy prices*. URL: https://energy.ec.europa.eu/topics/markets-and-consumers/eu-energy-prices (visited on 22 January 2022).

[20] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. *Integer programming*. 2016, pages 3.1–3.16. ISBN: 9781420009712. DOI: 10.1145/1216965.1216966. URL: https://link-springer-com.proxy.findit.cvt.dk/book/10.1007/978-3-319-11008-0.

[21] Carlos Adrian Correa-Florez et al. "Stochastic operation of home energy management systems including battery cycling." In: *Applied Energy* 225 (September 2018), pages 1205–1218. ISSN: 0306-2619. DOI: 10.1016/J.APENERGY.2018.04.130.

[22] Michael Dahl Knudsen and Steffen Petersen. "Demand response potential of model predictive control of space heating based on price and carbon dioxide intensity signals." In: *Energy and Buildings* 125 (2016), pages 196–204. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.04.053. URL: http://dx.doi.org/10.1016/j.enbuild.2016.04.053.

[23] Danish Energy Agency. *About the Danish Energy Agency | Energistyrelsen*. URL: https://ens.dk/en/about-us/about-danish-energy-agency (visited on 9 February 2022).

[24] Danish Energy Agency. *Technology Data for Generation of Electricity and District Heating | Energistyrelsen*. (Visited on 20 September 2021).

[25] Kalyanmoy Deb. *Multi-Objective Optimization using Evolutionary Algorithms*. 2nd edition. Chichester: John Wiley & Sons, Ltd, 2002.

[26] Seyyed Reza Ebrahimi et al. "Home energy management under correlated uncertainties: A statistical analysis through Copula." In: *Applied Energy* 305 (January 2022), page 117753. ISSN: 0306-2619. DOI: 10.1016/J.APENERGY.2021.117753.

[27] Niematallah Elamin and Mototsugu Fukushige. "Modeling and forecasting hourly electricity demand by SARIMAX with interactions." In: *Energy* 165 (December 2018), pages 257–268. ISSN: 03605442. DOI: 10.1016/j.energy.2018.09.157.

[28] Mahmoud Elkazaz et al. "Optimization based Real-Time Home Energy Management in the Presence of Renewable Energy and Battery Energy Storage." In: *SEST 2019 - 2nd International Conference on Smart Energy Systems and Technologies* (September 2019). DOI: 10.1109/SEST.2019.8849105.

[29] Mahmoud Elkazaz et al. "Optimization based Real-Time Home Energy Management in the Presence of Renewable Energy and Battery Energy Storage." In: *SEST 2019 - 2nd International Conference on Smart Energy Systems and Technologies* (2019). DOI: 10.1109/SEST.2019.8849105.

[30] Energinet. *Data about the energy system | Energinet*. URL: https://en.energinet.dk/Electricity/Energy-data (visited on 8 February 2022).

[31] Energinet. *Learn more about an aggregator's function and role | Energinet*. URL: https://en.energinet.dk/Electricity/Green-electricity/Demand-side-response/What-is-an-aggregator (visited on 11 February 2022).

[32] Energinet. *Pilot and market projects | Energinet*. URL: https://en.energinet.dk/Electricity/Green-electricity/Demand-side-response/Pilot-and-market-projects (visited on 11 February 2022).

[33] Energinet. *Rates | Energinet*. URL: https://energinet.dk/El/Elmarkedet/Tariffer (visited on 22 January 2022).

[34] Energinet. *Rates | Energinet*. https://energinet.dk/El/Elmarkedet/Tariffer. URL: https://energinet.dk/El/Elmarkedet/Tariffer (visited on 22 January 2022).

[35] Energinet. *What is DataHub? | Energinet*. URL: https://en.energinet.dk/Electricity/DataHub#Documents (visited on 23 March 2022).

[36] ENTSO-E. *ENTSO-E Mission Statement*. URL: https://www.entsoe.eu/about/inside-entsoe/objectives/ (visited on 7 March 2022).

[37] ENTSO-E. *ENTSO-E Transparency Platform*. URL: https://transparency.entsoe.eu/ (visited on 7 March 2022).

[38] European Commission. *Energy transition in cities | European Commission*. URL: https://ec.europa.eu/info/eu-regional-and-urban-development/topics/cities-and-urban-development/priority-themes-eu-cities/energy-transition-cities (visited on 24 March 2022).

[39] Tingting Fang and Risto Lahdelma. "Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system." In: *Applied Energy* 179 (October 2016), pages 544–552. ISSN: 03062619. DOI: 10.1016/j.apenergy.2016.06.133.

[40] Haroldo G. Santos and Túlio A.M. Toffolo. *Mixed Integer Linear Programming with Python*. Technical report. 2020, pages 1–58. URL: https://buildmedia.readthedocs.org/media/pdf/python-mip/latest/python-mip.pdf.

[41] Farhad Samadi Gazijahani, Sajad Najafi Ravadanegh, and Javad Salehi. "Stochastic multi-objective model for optimal energy exchange optimization of networked microgrids with presence of renewable generation under risk-based strategies." In: *ISA Transactions* 73 (2018), pages 100–111. ISSN: 00190578. DOI: 10.1016/j.isatra.2017.12.004.

[42] Murray Goulden et al. "Smart grids, smart users? The role of the user in demand side management." In: *Energy Research & Social Science* 2 (June 2014), pages 21–29. ISSN: 2214-6296. DOI: 10.1016/J.ERSS.2014.04.008.

[43] Nord pool group. *Bidding areas*. https://www.nordpoolgroup.com/the-power-market/Bidding-areas/. URL: https://www.nordpoolgroup.com/the-power-market/Bidding-areas/ (visited on 1 February 2022).

[44] Nord pool group. *Day-ahead market*. URL: https://www.nordpoolgroup.com/the-power-market/Day-ahead-market/ (visited on 1 November 2021).

[45] Nord pool group. *Rules and regulations*. https://www.nordpoolgroup.com/trading/Rules-and-regulations/. URL: https://www.nordpoolgroup.com/trading/Rules-and-regulations/ (visited on 26 January 2022).

[46] Gurobi Optimization. *Gurobi Optimizer Reference Manual*. 2019. URL: http://www.gurobi.com.

[47] Cristina Guzman, Alben Cardenas, and Kodjo Agbossou. "Evaluation of Meta-heuristic Optimization Methods for Home Energy Management Applications." In: *IEEE International Symposium on Industrial Electronics* (2017), pages 1501–1506. DOI: 10.1109/ISIE.2017.8001468.

[48] HAIMES YV, LASDON LS, and WISMER DA. "On a bicriterion formation of the problems of integrated system identification and system optimization." In: *IEEE Transactions on Systems, Man and Cybernetics* SMC-1.3 (1971), pages 296–297. ISSN: 00189472. DOI: 10.1109/TSMC.1971.4308298.

[49] Lena Hansson, Ulrika Holmberg, and Helene Brembeck. "Making sense of consumption : selections from the 2nd Nordic Conference on Consumer Research 2012." In: *2nd Nordic Conference on Consumer Research*. 2012, page 393. ISBN: 978-91-974642-6-0.

[50] Tom Hargreaves, Charlie Wilson, and Richard Hauxwell-Baldwin. "Learning to live in a smart home." In: *https://doi.org/10.1080/09613218.2017.1286882* 46.1 (January 2017), pages 127–139. DOI: 10.1080/09613218.2017.1286882. URL: https://www.tandfonline.com/doi/abs/10.1080/09613218.2017.1286882.

[51] Yaoyao He et al. "Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network." In: *Applied Energy* 233-234 (January 2019), pages 565–575. ISSN: 03062619. DOI: 10.1016/j.apenergy.2018.10.061.

[52] Yaoyao He et al. "Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory." In: *Applied Energy* 185 (January 2017), pages 254–266. ISSN: 03062619. DOI: 10.1016/j.apenergy.2016.10.079.

[53] William F. Holmgren, Clifford W. Hansen, and Mark A. Mikofski. "pvlib python: a python package for modeling solar energy systems." In: *Journal of Open Source Software* 3.29 (September 2018), page 884. ISSN: 2475-9066. DOI: 10.21105/JOSS.00884. URL: https://joss.theoj.org/papers/10.21105/joss.00884.

[54] Sayed Saeed Hosseini et al. "Non-intrusive load monitoring through home energy management systems: A comprehensive review." In: *Renewable and Sustainable Energy Reviews* 79 (November 2017), pages 1266–1274. ISSN: 18790690. DOI: 10.1016/j.rser.2017.05.096.

[55] Anne Immonen, Jussi Kiljander, and Matti Aro. "Consumer viewpoint on a new kind of energy market." In: *Electric Power Systems Research* 180 (March 2020), page 106153. ISSN: 0378-7796. DOI: 10.1016/J.EPSR.2019.106153.

[56] International Energy Agency. *World Energy Outlook 2021 - revised version October 2021*. Technical report. 2021. URL: www.iea.org/weo.

[57] Mohammad Javadi et al. "A Multi-Objective Model for Home Energy Management System Self-Scheduling using the Epsilon-Constraint Method." In: *Proceedings - 2020 IEEE 14th International Conference on Compatibility, Power Electronics and Power Engineering, CPE-POWERENG 2020*. 2020, pages 175–180. ISBN: 9781728142180. DOI: 10.1109/CPE-POWERENG48600.2020.9161526.

[58] Mohammad Sadegh Javadi et al. "Conditional Value-at-Risk Model for Smart Home Energy Management Systems." In: *e-Prime* 1.August (2021), page 100006. ISSN: 27726711. DOI: 10.1016/j.prime.2021.100006. URL: https://doi.org/10.1016/j.prime.2021.100006.

[59] Mohammad Sadegh Javadi et al. "Self-scheduling model for home energy management systems considering the end-users discomfort index within price-based demand response programs." In: *Sustainable Cities and Society* 68 (2021). ISSN: 22106707. DOI: 10.1016/j.scs.2021.102792.

[60] Xin Jin et al. "Foresee: A user-centric home energy management system for energy efficiency and demand response." In: *Applied Energy* 205.August (2017), pages 1583–1595. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.08.166. URL: http://dx.doi.org/10.1016/j.apenergy.2017.08.166.

[61] Andreas Jossen. "Fundamentals of battery dynamics." In: *Journal of Power Sources* 154.2 (March 2006), pages 530–538. ISSN: 0378-7753. DOI: 10.1016/J.JPOWSOUR.2005.10.041.

[62] Rune Grønborg Junker et al. "Characterizing the energy flexibility of buildings and districts." In: *Applied Energy* 225.February (2018), pages 175–182. ISSN: 03062619. DOI: 10.1016/j.apenergy.2018.05.037. URL: https://doi.org/10.1016/j.apenergy.2018.05.037.

[63] Isao Kanda and J. M.Quintana Veguillas. "Data preprocessing and quantile regression for probabilistic load forecasting in the GEFCom2017 final match." In: *International Journal of Forecasting* 35.4 (October 2019), pages 1460–1468. ISSN: 01692070. DOI: 10.1016/j.ijforecast.2019.02.005.

[64] Leonard Kaufmann and Peter Rousseeuw. "Clustering by Means of Medoids." In: *Data Analysis based on the L1-Norm and Related Methods* (1987), pages 405–416.

[65] Selina Kerscher and Pablo Arboleya. "The key role of aggregators in the energy transition under the latest European regulatory framework." In: *International Journal of Electrical Power & Energy Systems* 134 (January 2022), page 107361. ISSN: 0142-0615. DOI: 10.1016/J.IJEPES.2021.107361.

[66] Roger Koenker. *Quantile regression*. Cambridge University Press, January 2005, pages 1–349. ISBN: 9780511754098. DOI: 10.1017/CBO9780511754098. URL: https://www.cambridge.org/core/books/quantile-regression/C18AE7BCF3EC43C16937390D44A328B1.

[67] M. Kyung and S. K. Ghosh. "Maximum likelihood estimation for directional conditionally autoregressive models." In: *Journal of Statistical Planning and Inference* 140.11 (November 2010), pages 3160–3179. ISSN: 03783758. DOI: 10.1016/j.jspi.2010.04.012.

[68] L, Sevdari, and K. Marinelli. *Status e-mobility DK*. Technical report. https://orbit.dtu.dk/en/publications/status-e-mobility-dk. Technical University of Denmark (DTU), 2021. URL: https://orbit.dtu.dk/en/publications/status-e-mobility-dk.

[69] David P. Larson, Lukas Nonnenmacher, and Carlos F.M. Coimbra. "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest." In: *Renewable Energy* 91 (June 2016), pages 11–20. ISSN: 0960-1481. DOI: 10.1016/J.RENENE.2016.01.039.

[70] David P. Larson, Lukas Nonnenmacher, and Carlos F.M. Coimbra. "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest." In: *Renewable Energy* 91 (2016), pages 11–20. ISSN: 18790682. DOI: 10.1016/j.renene.2016.01.039. URL: http://dx.doi.org/10.1016/j.renene.2016.01.039.

[71] Joaquim Leitao et al. "A Survey on Home Energy Management." In: *IEEE Access* 8 (2020), pages 5699–5722. ISSN: 21693536. DOI: 10.1109/ACCESS.2019.2963502.

[72] Julian Lemos-Vinasco, Peder Bacher, and Jan Kloppenborg Møller. "Probabilistic load forecasting considering temporal correlation: Online models for the prediction of households' electrical load." In: *Applied Energy* 303 (December 2021), page 117594. ISSN: 0306-2619. DOI: 10.1016/J.APENERGY.2021.117594. URL: https://linkinghub.elsevier.com/retrieve/pii/S0306261921009685.

[73] Julian Lemos-Vinasco et al. "A Rolling Horizon Approach for Stochastic Optimization of Home Energy Management Systems." In: *Submitted to Applied Energy* (2022).

[74] Jon Liisberg. "Data-driven models for energy advising leading to behavioural changes in residences." PhD thesis. Technical University of Denmark, 2019. URL: https://orbit.dtu.dk/en/publications/data-driven-models-for-energy-advising-leading-to-behavioural-cha.

[75] Bidong Liu et al. "Probabilistic Load Forecasting via Quantile Regression Averaging on Sister Forecasts." In: *IEEE Transactions on Smart Grid* 8.2 (March 2017), pages 730–737. ISSN: 19493053. DOI: 10.1109/TSG.2015.2437877.

[76] Henrik Madsen. *Time Series Analysis*. Chapman & Hall, 2008. ISBN: ISBN-10: 142005967X | ISBN-13: 978-1420059670 0. URL: http://henrikmadsen.org/wordpress/time-series-analysis/.

[77] D. Mariano-Hernández et al. "A review of strategies for building energy management system: Model predictive control, demand side management, optimization, and fault detect & diagnosis." In: *Journal of Building Engineering* 33 (January 2021), page 101692. ISSN: 2352-7102. DOI: 10.1016/J.JOBE.2020.101692.

[78] James E. Matheson and Robert L. Winkler. "Scoring rules for continuous probability distributions." In: *Management Science* 22.10 (1976), pages 1087–1096. ISSN: 00251909. DOI: 10.1287/mnsc.22.10.1087.

[79] Claire McIlvennie, Angela Sanguinetti, and Marco Pritoni. "Of impacts, agents, and functions: An interdisciplinary meta-review of smart home energy management systems research." In: *Energy Research & Social Science* 68 (October 2020), page 101555. ISSN: 2214-6296. DOI: 10.1016/J.ERSS.2020.101555.

[80] Mckinsey & Company. *How electric vehicles could change the load curve | McKinsey*. URL: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-potential-impact-of-electric-vehicles-on-global-energy-systems (visited on 29 March 2022).

[81] Leonardo Meeus. *The Evolution of Electricity Markets in Europe*. Edward Elgar Publishing, November 2020. DOI: 10.4337/9781789905472.

[82] Kaisa Miettinen. "Some Methods for Nonlinear Multi-objective Optimization." In: *Evolutionary Multi-Criterion Optimization*. Edited by Eckart Zitzler et al. Springer, 2001. ISBN: 3540417451.

[83] Annette Möller, Alex Lenkoski, and Thordis L. Thorarinsdottir. "Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas." In: *Quarterly Journal of the Royal Meteorological Society* 139.673 (April 2013), pages 982–991. ISSN: 1477-870X. DOI: 10.1002/qj.2009. URL: https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.2009.

[84] Juan M. Morales et al. "Integrating renewables in electricity markets - Operational problems." In: *Springer* 205 (2014), page 429. DOI: 10.1007/978-1-4614-9411-9.

[85] Larissa Nicholls and Yolande Strengers. "Peak demand and the 'family peak' period in Australia: Understanding practice (in)flexibility in households with children." In: *Energy Research & Social Science* 9 (September 2015), pages 116–124. ISSN: 2214-6296. DOI: 10.1016/J.ERSS.2015.08.018.

[86] Larissa Nicholls, Yolande Strengers, and Sergio Tirado. "Smart home control: Exploring the potential for off-the-shelf enabling technologies in energy vulnerable and other households." In: *Energy Consumers Australia* (2017). URL: https://researchrepository.rmit.edu.au/esploro/outputs/report/Smart-home-control-Exploring-the-potential/9921861798501341.

[87] Henrik Aalborg Nielsen and Henrik Madsen. "Modelling the heat consumption in district heating systems using a grey-box approach." In: *Energy and Buildings* 38.1 (January 2006), pages 63–71. ISSN: 03787788. DOI: 10.1016/j.enbuild.2005.05.002.

[88] Saman Nikkhah et al. "Optimising Building-to-Building and Building-for-Grid Services under Uncertainty: A Robust Rolling Horizon Approach." In: *IEEE Transactions on Smart Grid* Preprint (2021). ISSN: 1949-3053. DOI: 10.1109/tsg.2021.3135570.

[89] Nord Pool Group. *About us | Nord Pool*. URL: https://www.nordpoolgroup.com/About-us/ (visited on 9 February 2022).

[90] OpenWeather Ltd. *Point forecast description*. URL: https://openweathermap.org/api/hourly-forecast (visited on 26 August 2020).

[91] Serafim Opricovic and Gwo Hshiung Tzeng. "Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS." In: *European Journal of Operational Research* 156.2 (2004), pages 445–455. ISSN: 03772217. DOI: 10.1016/S0377-2217(03)00020-1.

[92] European parliament. *Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02012L0027-20210101. (Visited on 22 January 2022).

[93] Nikolaos G. Paterakis et al. "Optimal operation of smart houses by a real-time rolling horizon algorithm." In: *IEEE Power and Energy Society General Meeting* 2016-Novem (November 2016). DOI: 10.1109/PESGM.2016.7741507.

[94] Nikolaos G. Paterakis et al. "Optimal operation of smart houses by a real-time rolling horizon algorithm." In: *IEEE Power and Energy Society General Meeting* 2016-Novem (2016). ISSN: 19449933. DOI: 10.1109/PESGM.2016.7741507.

[95] Ignacio J. Pérez-Arriaga. *Regulation of the power sector*. Volume 61. 2013. DOI: 10.1007/978-1-4471-5034-3.

[96] Oscar Perpiñán. "solaR : Solar Radiation and Photovoltaic Systems with R." In: *Journal of Statistical Software* 50.9 (August 2012), pages 1–32. ISSN: 1548-7660. DOI: 10.18637/jss.v050.i09. URL: https://www.jstatsoft.org/index.php/jss/article/view/v050i09/v50i09.pdf%20https://www.jstatsoft.org/index.php/jss/article/view/v050i09.

[97] Pierre Pinson et al. "From probabilistic forecasts to statistical scenarios of short-term wind power production." In: *Wind Energy* 12.1 (2009), pages 51–62. ISSN: 10954244. DOI: 10.1002/we.284.

[98] Radius A/S. *Tariffer og netabonnement - få et overblik her på siden*. https://radiuselnet.dk/elkunder/priser-og-vilkaar/tariffer-og-netabonnement/. URL: https://radiuselnet.dk/elkunder/priser-og-vilkaar/tariffer-og-netabonnement/ (visited on 23 December 2021).

[99] Sahar Rahim et al. "Exploiting heuristic algorithms to efficiently utilize energy management controllers with renewable energy sources." In: *Energy and Buildings* 129 (2016), pages 452–470. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.08.008. URL: http://dx.doi.org/10.1016/j.enbuild.2016.08.008.

[100] Mashud Rana and Irena Koprinska. "Forecasting electricity load with advanced wavelet neural networks." In: *Neurocomputing* 182 (March 2016), pages 118–132. ISSN: 18728286. DOI: 10.1016/j.neucom.2015.12.004.

[101] Lisa Buth Rasmussen et al. "Load forecasting of supermarket refrigeration." In: *Applied Energy* 163 (February 2016), pages 32–40. ISSN: 0306-2619. DOI: 10.1016/J.APENERGY.2015.10.046.

[102] Lisa Buth Rasmussen et al. "Load forecasting of supermarket refrigeration." In: *Applied Energy* 163 (2016), pages 32–40.

[103] Marcelo Salgado et al. "A low-complexity decision model for home energy management systems." In: *Applied Energy* 294 (July 2021), page 116985. ISSN: 0306-2619. DOI: 10.1016/J.APENERGY.2021.116985.

[104] Michael Scheuerer and Thomas M. Hamill. "Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities." In: *Monthly Weather Review* 143.4 (2015), pages 1321–1334. ISSN: 15200493. DOI: 10.1175/MWR-D-14-00269.1.

[105] Amos Schledorn et al. "Frigg: Soft-Linking Energy System and Demand Response Models." In: *Preprint submitted to Applied Energy* (2022).

[106] scikit-learn-extra developers. *scikit-learn-extra*. 2019. URL: https://scikit-learn-extra.readthedocs.io/en/stable/contributing.html.

[107] Eva Lucas Segarra, Germán Ramos Ruiz, and Carlos Fernández Bandera. "Probabilistic load forecasting for building energy models." In: *Sensors (Switzerland)* 20.22 (November 2020), pages 1–20. ISSN: 14248220. DOI: 10.3390/s20226525.

[108] Miadreza Shafie-Khah and Pierluigi Siano. "A stochastic home energy management system considering satisfaction cost and response fatigue." In: *IEEE Transactions on Industrial Informatics* 14.2 (2018), pages 629–638. ISSN: 15513203. DOI: 10.1109/TII.2017.2728803.

[109] M Sklar. "Fonctions de repartition a n dimensions et leurs marges." In: (1959).

[110] Julija Tastu, Pierre Pinson, and Henrik Madsen. "Multivariate conditional parametric models for a spacio-temporal analysis of short-term wind power forecast errors." In: *Scientific Proceedings of the European Wind Energy Conference*. Warsaw, 2010, pages 77–81.

[111] Julija Tastu, Pierre Pinson, and Henrik Madsen. "Space-time trajectories of wind power generation: Parametrized precision matrices under a Gaussian copula approach." In: *Lecture Notes in Statistics*. Volume 217. Springer Science and Business Media, LLC, 2015, pages 267–296. ISBN: 9783319187310. DOI: 10.1007/978-3-319-18732-7_14. URL: https://link.springer.com/chapter/10.1007/978-3-319-18732-7_14.

[112] Sergio Tirado Herrero, Larissa Nicholls, and Yolande Strengers. "Smart home technologies in everyday life: do they address key energy challenges in households?" In: *Current Opinion in Environmental Sustainability* 31 (April 2018), pages 65–70. ISSN: 1877-3435. DOI: 10.1016/J.COSUST.2017.12.001.

[113] Marcos Tostado-Véliz, Samundra Gurung, and Francisco Jurado. "Efficient solution of many-objective Home Energy Management systems." In: *International Journal of Electrical Power & Energy Systems* 136 (2022). DOI: 10.1016/j.ijepes.2021.107666.

[114] Jaclason Veras et al. "A Multi-Objective Demand Response Optimization Model for Scheduling Loads in a Home Energy Management System." In: *Sensors* 18 (September 2018), page 3207. DOI: 10.3390/s18103207.

[115] Jaclason M. Veras et al. "A Multi-Objective Demand Response Optimization Model for Scheduling Loads in a Home Energy Management System." In: *Sensors* 18.3207 (2018). ISSN: 14248220. DOI: 10.3390/s18103207.

[116] Vindstød. *Billig el fra danske vindmøller | Vindstød*. https://www.vindstoed.dk/solcelle. URL: https://www.vindstoed.dk/solcelle (visited on 10 November 2021).

[117] PJC Vogler–Finck, Peder Bacher, and Henrik Madsen. "Online short-term forecast of greenhouse heat load using a weather forecast service." In: *Applied Energy* 205 (2017), pages 1298–1310.

[118] Julian Vossen, Baptiste Feron, and Antonello Monti. "Probabilistic forecastic of household electrical load using Artificial Neural Networks." In: *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)* (2018), pages 1–6.

[119] Xiuwang Wang, Xinna Mao, and Hossein Khodaei. "A multi-objective home energy management system based on internet of things and optimization algorithms." In: *Journal of Building Engineering* 33 (2021), page 101603. ISSN: 23527102. DOI: 10.1016/j.jobe.2020.101603. URL: https://doi.org/10.1016/j.jobe.2020.101603.

[120] Yi Wang et al. "Probabilistic individual load forecasting using pinball loss guided LSTM." In: *Applied Energy* 235 (February 2019), pages 10–20. ISSN: 03062619. DOI: 10.1016/j.apenergy.2018.10.078.

[121] Zhiyuan Wang and Gade Pandu Rangaiah. "Application and Analysis of Methods for Selecting an Optimal Solution from the Pareto-Optimal Front obtained by Multiobjective Optimization." In: *Industrial and Engineering Chemistry Research* 56 (2017), pages 560–574. ISSN: 15205045. DOI: 10.1021/acs.iecr.6b03453.

[122] Michael Waskom. "seaborn: statistical data visualization." In: *Journal of Open Source Software* 6.60 (2021), page 3021. ISSN: 2475-9066. DOI: 10.21105/joss.03021.

[123] Watts A/S. *How much is a kWh? • Green energy • Watts - energy assistant*. https://watts.dk/en/elprodukt/hvad-koster-en-kwh/. URL: https://watts.dk/en/elprodukt/hvad-koster-en-kwh/ (visited on 2 November 2021).

[124] Watts A/S. *Watts Energy Assistant • Use power with green care | Download the app.* URL: https://watts.dk/en/ (visited on 23 March 2022).

[125] Watts A/S. *Watts Features.* URL: https://watts.dk/en/funktioner/.

[126] Yandong Yang et al. "Power load probability density forecasting using Gaussian process quantile regression." In: *Applied Energy* 213 (March 2018), pages 499–509. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.11.035.

[127] Mojtaba Yousefi et al. "Profit assessment of home energy management system for buildings with A-G energy labels." In: *Applied Energy* 277 (November 2020), page 115618. ISSN: 0306-2619. DOI: 10.1016/J.APENERGY.2020.115618.

[128] Liang Yu, Tao Jiang, and Yulong Zou. "Online Energy Management for a Sustainable Smart Home with an HVAC Load and Random Occupancy." In: *IEEE Transactions on Smart Grid* 10.2 (2019), pages 1646–1659. ISSN: 19493053. DOI: 10.1109/TSG.2017.2775209.

[129] Saeed Zeynali et al. "Two-stage stochastic home energy management strategy considering electric vehicle and battery energy storage system: An ANN-based scenario generation methodology." In: *Sustainable Energy Technologies and Assessments* 39 (June 2020), page 100722. ISSN: 2213-1388. DOI: 10.1016/J.SETA.2020.100722.

[130] Shu Zhang et al. "Load probability density forecasting by transforming and combining quantile forecasts." In: *Applied Energy* 277 (November 2020), page 115600. ISSN: 03062619. DOI: 10.1016/j.apenergy.2020.115600.

[131] Yanyu Zhang et al. "A Novel Multiobjective Optimization Algorithm for Home Energy Management System in Smart Grid." In: *Mathematical Problems in Engineering* (2015). ISSN: 15635147. DOI: 10.1155/2015/807527.

[132] Qiang Zhou et al. "A grey-box model of next-day building thermal load prediction for energy-efficient control." In: *International Journal of Energy Research* 32.15 (December 2008), pages 1418–1431. ISSN: 0363907X. DOI: 10.1002/er.1458. URL: http://doi.wiley.com/10.1002/er.1458.

[133] Jernej Zupančič, Bogdan Filipič, and Matjaž Gams. "Genetic-programming-based multi-objective optimization of strategies for home energy-management systems." In: *Energy* 203 (2020). ISSN: 03605442. DOI: 10.1016/j.energy.2020.117769.