



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich



Roman Engeler

Long-term Energy Forecasting of Smart Metered Customers in a Distribution Grid

Semester Thesis
PSL 1810

EEH – Power Systems Laboratory
Swiss Federal Institute of Technology (ETH) Zurich

Expert: Prof. Dr. Gabriela Hug
Supervisor: M.Sc. Thierry Zufferey

Zurich, July 15, 2018

Acknowledgment

This work is the outcome of my semester project as part of the Master's degree in mechanical engineering. The purpose of this module is to prepare students for the Master thesis and to contribute to cutting-edge research at ETH.

I would like to take the chance and thank Professor Gabriela Hug, head of the Power Systems Laboratory (PSL) at ETH, for the opportunity to learn more about future developments of the electricity grid and long-term forecasting.

Moreover, I am very grateful to my supervisor, Thierry Zufferey, for the insightful discussions and the advice during my project. I could fully profit from his knowledge about forecasting which he acquired in his Ph.D. about applications of artificial intelligence and data analytics for the operation and planning of active distribution grids.

Finally, I would like to thank Professor Konstantinos Boulouchos, head of the Aerothermochemistry and Combustion Systems Laboratory (LAV) at ETH, for his advice during my Master's program and for supporting my endeavours of acquiring applied knowledge in machine learning.

Abstract

The focus of this thesis lies on the low voltage distribution grid, which is of primary concern due to the integration of future technologies. Therefore, long-term predictions until 2035 are obtained for an aggregation of around 60 consumers. This requires to consider different scenarios for the development of the electricity consumption and the adaptation of technologies such as photovoltaic cells, heat pumps and electric vehicles. Furthermore, smart meter data from the city of Basel was available to analyze consumer behaviour. It is proposed to extract underlying patterns by K-Means clustering and to use a semi-parametric model relying solely on external variables. The model is fitted by linear regression and a neural network which are compared in terms of accuracy and speed. A characteristic of the proposed model is its dependence on temperature and irradiation which is exploited to obtain a density forecast for the electricity demand. Temperature and irradiation simulations are performed by a seasonal bootstrapping method to obtain enough variation in the samples for an appropriate estimation of the underlying distribution. The forecasts are used to identify key times defined by the most profound changes of the consumer behaviour. This knowledge at hand, a discussion is provided for the overloading of low voltage distribution lines and distribution transformers. Finally, measures are proposed to circumvent the revealed issues.

Contents

Acknowledgment	i
List of Acronyms	v
1 Introduction	1
1.1 Distribution Grid	2
1.2 Approaches for Prediction	4
1.2.1 Statistical Methods	4
1.2.2 Machine Learning Methods	5
1.3 Prediction Pipeline	6
2 Future Trends and Scenarios	7
2.1 General Development	7
2.2 Photovoltaic Cells	8
2.3 Heat Pumps	9
2.4 Electric Vehicles	11
2.5 Energy Storage	13
3 Data Preparation	14
3.1 Extraction of Residential Consumers	14
3.2 Clustering	15
3.2.1 K-Means Algorithm	15
3.2.2 Clustering of Load Profiles	16
3.3 Temperature Simulation	19
3.4 Heat-pump Simulation	20
3.5 Electric Vehicle Simulation	22
4 Semi-Parametric Model	24
4.1 Feature Selection	24
4.2 Model Structure	26
4.3 Linear Regression Model	27
4.3.1 Linear Regression Theory	27
4.3.2 Electricity Load Profile Approximation	28
4.3.3 PV Production Profile Approximation	30

CONTENTS	iv
4.4 Multi Layer Perceptron	32
4.4.1 Concepts of Artificial Neural Networks	32
4.4.2 Electricity Load Profile Approximation	33
4.5 Linear Regression versus Multi-Layer Perceptron	35
4.5.1 Fitting Accuracy	35
4.5.2 Computational Efficiency	37
5 Forecast	38
5.1 Future Prediction	38
5.2 Validation of Prediction Pipeline	40
5.3 Electricity Demand until 2035	42
5.3.1 Variation of Demand during a Year	42
5.3.2 Variation of Demand during a Day	43
5.3.3 Variation of Demand over the Years	44
5.3.4 Development of Demand for different Scenarios	46
5.4 Overloading of critical Grid Infrastructure	47
6 Conclusion	49
6.1 Summary	49
6.2 Outlook	50
Bibliography	52
Appendices	55
.1 Variation of Demand during a Day	56
.2 Variation of Demand over the Years	57
.3 Development of Demand for different Scenarios	58

List of Acronyms

A.F.	Air Forced
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
AR	Auto-Regressive
ARCH	Auto-Regressive Conditional Heteroskedastic
ARMA	Auto-Regressive Moving Average
ARIMAX	Auto-Regressive Integrated Moving Average Exogenous
BEV	Battery Electric Vehicle
CRBM	Conditionally Restricted Boltzmann Machine
DSO	Distribution System Operator
EV	Electric Vehicle
GARCH	General Auto-Regressive Conditional Heteroskedastic
GDP	Gross Domestic Product
HP	Heat Pump
LR	Linear Regression
LV	Low Voltage
MASE	Mean Absolute Scaled Error
MLP	Multi-Layer Perceptron
PHEV	Plug Hybrid Electric Vehicle
PV	Photovoltaic
REEV	Range Extended Electric Vehicle
RBM	Restricted Boltzmann Machine
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SE	Square Error
TCL	Thermostatically Controlled Loads

List of Figures

1.1	Future technologies discussed in this report.	2
1.2	Overview of the different grid levels in Switzerland (taken from SwissGrid).	3
2.1	Scenarios for the electricity consumption in Switzerland (taken from [1])..	8
2.2	Scenarios for deployment of photovoltaic cells.	9
2.3	Number of heat pumps installed each year split into categories (taken from [2]).	10
2.4	E-Mobility scenarios in Switzerland (taken from [3] (2015)). Shown is the share of EVs in sold vehicles from 2010 to 2035.	12
3.1	Electricity demand averaged over one year and sorted by increasing size.	15
3.2	Clustering cost versus number of clusters.	17
3.3	Four distinct daily profiles obtained by clustering showing different types of consumers.	18
3.4	Double Season Block Bootstrap.	20
3.5	Heat pump load profile for winter (left) and a day in November (right).	21
3.6	Procedure for HP resampling.	22
3.7	Traffic for three roads in Basel.	23
3.8	Histogram for the charging of electric vehicles obtained from synthetic profiles over a year. The blue histogram is obtained from the original synthetic profile and the red from the resampled profile.	23
4.1	Split of electricity demand for a typical household (taken from [4]).	25
4.2	Scaled demand versus current temperature for two different clusters.	26
4.3	Scaled demand for cluster 0.	29
4.4	Scaled demand for cluster 10.	30
4.5	Left: PV production profiles normalized by the maximum demand. Right: Average of normalized PV production profiles (blue) and the corresponding approximation (red).	30
4.6	Structure of a multi-layer perceptron with one hidden layer.	32
4.7	Load profile with one particular fit on the left, error versus number of hidden units on the right.	34

4.8	Load profile with one particular fit on the left, error versus number of hidden units on the right.	34
4.9	Comparison of errors for MLP fit and LS fit (without enforced sparsity).	36
4.10	Profile with better linear regression approximation and corresponding fit.	36
4.11	Profile with better multi-layer perceptron approximation and corresponding fit.	37
4.12	Comparison of errors for MLP fit and LS fit (with enforced sparsity).	37
5.1	Electricity demand at two specific times on a working day over the course of a year for node A with 67 consumers. Left: predicted demand for 2015 with actual smart meter data in green. Right: predicted demand for 2035.	41
5.2	Electricity demand at two specific times on a working day over the course of a year for node B with 68 consumers. Left: predicted demand for 2015 with actual smart meter data in green. Right: predicted demand for 2035.	41
5.3	Electricity demand for different times during a working day in summer for node A with 67 consumers. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.	43
5.4	Electricity demand for different times during a working day in winter for node A with 67 consumers. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.	44
5.5	Evolution of the density on a working day in summer at 12:00 for node A with 67 consumers.	45
5.6	Evolution of the density on a working day in winter at 20:00 for node A with 67 consumers.	45
5.7	Demand for a working day in summer at 12:00 for node A with 67 consumers. The lower bounds are given by the high PV production scenario and the upper bounds by the low PV production scenario.	46
5.8	Demand for a working day in winter at 20:00 for node A with 67 consumers.	47
1	Electricity demand for different times during a working day in summer for node B. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.	56
2	Electricity demand for different times during a working day in winter for node B. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.	56
3	Evolution of the density on a working day in summer at 12:00 for node B.	57
4	Evolution of the density on a working day in winter at 20:00 for node B.	57

5	Demand for a working day in summer at 12:00 for node B. The lower bounds are given by the high PV production scenario and the upper bounds by the low PV production scenario.	58
6	Demand for a working day in winter at 20:00 for node B.	58

List of Tables

2.1 Number of EV per low-voltage grid area of 60 consumers.	12
---	----

Chapter 1

Introduction

Energy systems build the backbone of modern society allowing businesses to grow and individuals to enjoy the benefits of high-tech devices. We depend heavily on the reliable functioning of the electricity grid as past events have shown impressively. A power outage can affect the life of millions of people as the one 2003 affecting Italy and southwest Switzerland (55 millions), the one 2009 in Brazil (87 millions) or the one 2012 in India (620 millions) [5] to name just a few. In order to minimize future impacts on our society, it is crucial to understand the complex interplay between the different levels of a power grid and to keep the infrastructure up-to-date.

The electricity grid will undergo major changes particularly on the lowest level connecting the individual consumers with the electricity grid. Fig 1.1 depicts a few technologies that will have an impact on the future grid and that will thus be considered in this work. The rapid expansion of renewables and the electrification of the heating and transport section pose challenges for the grid operators. Transitioning towards a smart grid is required to tackle them. An important step in this direction is the deployment of monitoring devices at all grid levels. It allows the operator to track the system's state in near real-time and take measures in case of local grid overloading. A counter-measure could be for example to incentivize the customers to consume at specific times of the day which can be achieved by dynamic pricing [6]. For smart meters with a bidirectional communication system, demand-side management allows to exploit the flexibility of certain devices, namely Thermostatically Controlled Loads (TCL) such as heat pumps, boilers, etc.

With the knowledge about today's consumption pattern, future developments can be studied. By considering different scenarios of for example the uptake of electric vehicles, one can gain valuable insights into future requirements for the grid. This allows operators to update appropriately their infrastructure in order to keep the power systems up to the highest standards. An important consideration lies in the time frame over which smarting of the grid should be achieved. [7] discusses different approaches for the different grid levels as for example no smarting or gradual smarting in terms of reliability, cost, accessibility and efficiency.

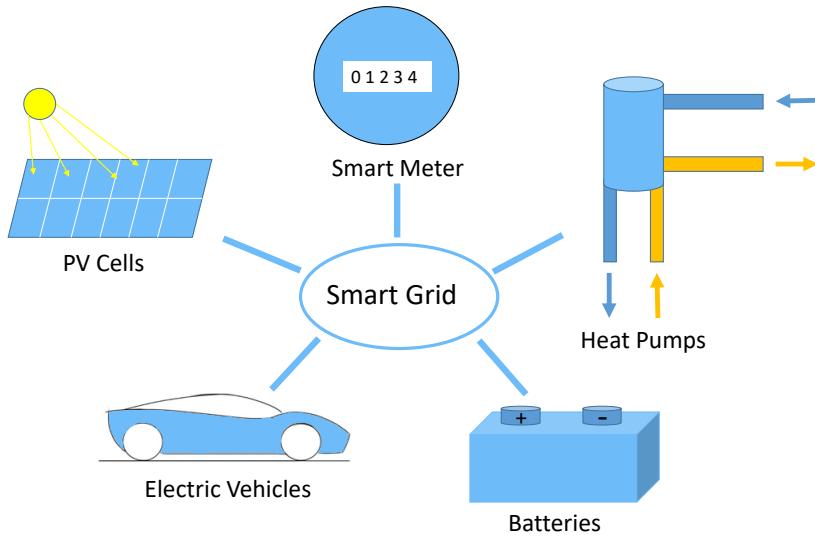


Figure 1.1: Future technologies discussed in this report.

Compared to the existing literature, the key contribution of this project are density distributions for a small aggregation of residential load profiles (50- 80) combined with different future scenarios to study the long-term development and identify critical times for grid overloading. Therefore, smart meter data is available of around 30'000 individual consumers in the city of Basel over the span of approximately three years (Apr '14 - Sept '17). For the accurate prediction of such a small aggregations, K-means clustering is proposed to group similar consumers. The extracted patterns are enriched with future scenarios for the development of electricity demand, deployment of PhotoVoltaic (PV) cells and Heat Pumps (HP) as well as the growing number of Electric Vehicles (EV). Finally, density predictions are provided for year 2019 until 2035. This should build a starting point for further considerations and planning activities of distribution grid operators.

1.1 Distribution Grid

To understand the framework of this thesis, some knowledge is required about the distribution grid. The first paragraph relates the grid level of interest to the overall energy system. In the second paragraph, a discussion is provided about the particular interests into the distribution grid. Finally, the need for data acquisition for control purposes is emphasized.

The power grid is split into different levels, ranging from the transmission grid (grid level 1) down to the local low-voltage distribution grid (grid level 7) as depicted in Fig 1.2. To connect the different levels, transformers are used to decrease or increase the alternating voltages. Power plants are connected to the extra-high-voltage level grid,

which transports the electricity over large distances. On one side, it is used to directly connect large cities to the power generating facilities and on the other side, it grants access to the European electricity grid. The high-voltage level provides the different regions with electricity. From there, the network is becoming thinner and thinner the closer it gets to the end customer. Medium-voltage level lines are usually used within a region to access smaller cities and villages. Finally, individual consumers are connected to the local distribution grid, which is organized in a radial topology. Supply security in the distribution grid is guaranteed by loading each line with less than 50% such that in the case of an outage, a neighboring branch can take over the load [8]. Hence, it is very important for the Distribution System Operator (DSO) to know the capacity of the grid and to predict how the future load will change. This provides opportunities for optimization by reducing the safety margin for a better usage of the existing grid infrastructure [9].

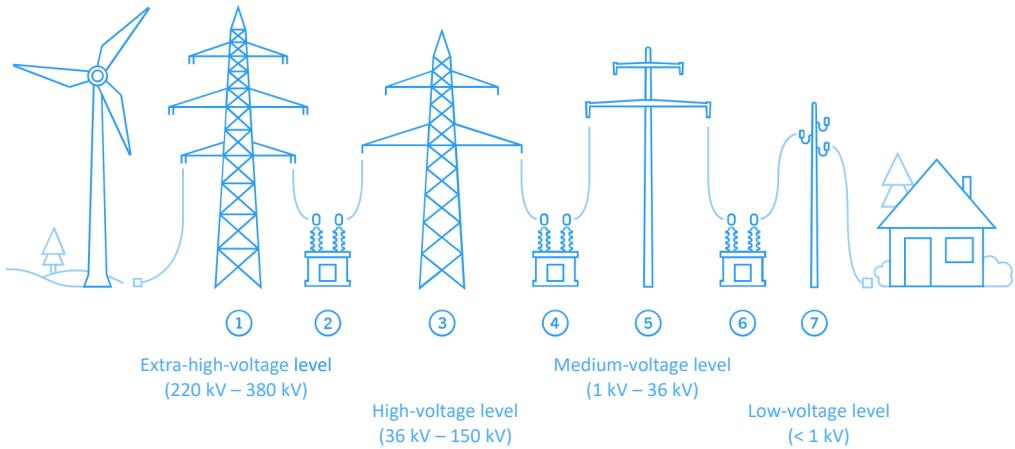


Figure 1.2: Overview of the different grid levels in Switzerland (taken from SwissGrid).

Reliable planning on the distribution grid level is rendered difficult since most renewable energy sources will be connected to this grid level [9]. As a consequence, the challenges due to the Energy Strategy 2050 are largest there. A large set of a grid operator's assets consists of the Low Voltage (LV) grid infrastructure [9]. Distribution grid lines and transformers have to be able to cope with the demanded load at any time and need to be updated appropriately. Replacing a distribution transformer is very costly, thus it is important for a DSO to be able to predict accurately future loads. Grid reinforcements due to PV power are expected to increase up to 3'000 per year in 2035 and cost an additional 180 million CHF for a specific rural distribution grid operator [9]. To reduce costs, the power flow can be controlled but this requires insights into the load pattern.

On the transmission and sub-transmission level, monitoring and control of power

flows is already incorporated today [10]. This provides the operator with large amounts of data in near real time [8] giving insights into the state of the grid, which allows for a reliable planning and prediction of demands. However, the local distribution grid is still without near real time monitoring and control [8]. With an increased share of renewables in the energy mix, the need for near real time sensoring arises due to the intermittency of these energy sources. Recently, large numbers of smart meters were installed in the different distribution grids [9, 11], which provides the basics for big data analytics and an efficient monitoring and control of the grid-state.

1.2 Approaches for Prediction

Time series forecasting is of wide interest in many disciplines as for example econometric forecasts, weather forecasts or energy demand forecasts. Depending on the forecasting horizon, it is differentiated between short-term (often half an hour to one week ahead), mid-term (one month up to one year ahead) and long-term (more than one year ahead) predictions [12]. In the case of short-term forecasting, a point prediction might be appropriate where a single value at a specified time in the future is predicted. However, for long-term predictions, the uncertainty is substantial and a density forecast should be favoured providing some kind of uncertainty measure. In the following two Subsections 1.2.1 and 1.2.2, different methodologies for forecasting are discussed and the choice for this work will be justified.

1.2.1 Statistical Methods

Some popular approaches in the context of statistical electricity consumption prediction will be introduced and the time horizon as well as the resolution of each method will be emphasized. Finally, the choice of method in this work will be given.

[13] applies a linear regression model to predict the monthly residential electricity consumption per unit at national level in Brazil for a ten-year horizon. Inputs are demographic, macroeconomic and microeconomic models.

In the landscape of electricity forecasting methods, ARMA (AR: Auto-Regressive, MA: Moving-Average) models and its derivatives are very prominent. In the case of an underlying trend, the series is first differentiated leading to a so-called ARIMA (I: Integrated) model. To include other drivers, exogenous variables can be added to the model which is then called ARIMAX (X: exogenous). For example in [14], an ARIMAX model is used to predict the net annual electricity consumption in Morocco several years into the future based on the demography and the Gross Domestic Product (GDP).

In the framework of ARMA models, the residuals are assumed to satisfy the requirements of Gaussian white noise: uncorrelated and normally distributed with zero mean and constant variance [15]. If the variance of the residuals changes over time, the assumptions for the ARMA class of models are not met anymore. This issue can be resolved by using an ARCH (Auto-Regressive Conditional Heteroskedastic) model

for the residuals, which can accommodate for the serial correlation. [15] used different variants of ARMA-ARCH models to predict the hourly electricity price for New England one-step-ahead.

[16] makes use of an ARIMA-GARCH (G: generalized) model to predict the aggregated monthly electricity demand for a region under an U.S. regional transmission operator up to one year into the future (12-steps-ahead).

As outlined earlier, long-term predictions ask for density estimation to quantify the uncertainty. Although it is possible to derive some kind of measure for the predictive uncertainty for above models, e.g. by confidence intervals, another approach will be used as introduced by [17]. The author used a semi-parametric model to estimate the weekly and annual maximum electricity demand for South Australia up to ten years into the future. A distribution for the demand is obtained by a mixture of temperature simulations as explained in Section 3.3, future economic scenarios, and residual bootstrapping. This approach is favoured due to two main reasons. First of all, a true density distribution is obtained by exploiting the dependence of individual consumers on the outside temperature. This also gives some indication about the sensitivity with respect to the weather. Further, it provides a convenient way to include other technologies as for example photovoltaic cells or heat pumps in the densities.

1.2.2 Machine Learning Methods

In recent years, machine learning algorithms have gained popularity in time series prediction. Particularly the expressive power of Artificial Neural Networks (ANNs) is exploited to obtain accurate predictions for time series. Selected recent developments in the field of electricity demand predictions will be summarized in this subsection. Subsequently, the method applied in this thesis will be introduced.

In the case of short-term forecasts, Multi-Layer Perceptrons (MLPs) have been in use for some years as for example in [18]. The author predicted the electricity demand for different levels of aggregation (one up to around 40'000 consumers) with a resolution of 15 minutes and a time horizon of one day ahead. As features, processed previous values of the consumption and exogenous variables (weather data, calendar features and public holidays) are fed into the network.

[19] compared different machine learning algorithms for predicting the hourly energy consumption of an office building (35 people) up to one-week ahead based on its history of energy consumption. Among the methods are non-linear auto-regressive neural networks with a single hidden layer and a Conditionally Restricted Boltzmann Machine (CRBM), with the latter outperforming the others.

Furthermore, various deep learning models have been adopted to predict the electric power consumption of a single household for short- and mid-term scenarios with different resolutions (1 min for 15 min horizon, 1 week for 1 year horizon) in [20]. Extensions of a Restricted Boltzmann Machine (RBM) are compared with a MLP and a Recurrent

Neural Network (RNN) whereby the RBM outperforms the others in most scenarios. In contrast to other studies, data available from different submeterings is used to train the models.

An application of MLPs for long-term forecasts is shown in [12], in which the monthly electricity consumption on a state-level of Australia is predicted for up to two years ahead. Input variables are besides socio-economic factors also environmental inputs. In the same work, it is suggested to use a deep network derived by a combination of stacked auto-encoders and MLPs for improved performance. Auto-encoders offer the possibility to learn a reduced representation of the data by for example ignoring noise or redundancies.

In this work, an MLP with one hidden layer will be applied for predicting the time series. First of all, it is always advisable to use the most simple model able to explain the data since the interpretation is facilitated. Further, the implementation of a MLP with the Keras library in Python is straightforward rendering it a convenient choice for the scope of this thesis. Furthermore, the model fitting process will be automated, and hence, a very standard model is preferred with few necessary choices of parameters.

1.3 Prediction Pipeline

The steps followed to obtain the long-term predictions will be the topic of this section. In Chapter 2, the technologies are discussed which will influence the way of how electricity will be consumed in the future. Subsequently, the model will be prepared and predictions will be obtained. A list of the necessary tasks is given below:

1. Standardization of load profiles
2. Clustering load profiles into K clusters
3. Logarithmic transformation of the resulting cluster centers
4. Function fitting of the transformed centers
5. Long-term prediction for cluster centers
6. Inverse transformation of predicted cluster centers
7. Inverse standardization of aggregated predicted profiles

Chapter 3 addresses the standardization of the profiles as well as the clustering. Furthermore, simulations are outlined for temperature and irradiation, heat pump, and electric vehicle data which are later used for the prediction. Number 3 and 4 of above list are subject of Chapter 4. Therein, the applied model will be introduced and the feature selection will be motivated. Then, the linear regression model and the multi-layer perceptron are defined. Both models are used to approximate the load curves and their performance is compared in the last section of the chapter. Subsequently, in Chapter 5, long-term predictions are obtained by the linear regression model. An analysis of the development reveals key times for an overloading of the grid. Finally, the work in this thesis is summarized and conclusions are drawn in Chapter 6.

Chapter 2

Future Trends and Scenarios

Different factors will increase the risk of grid overloading but also provide the opportunity for load balancing for the local distribution grid operator. A few of these changes will be mentioned and discussed in further detail in this chapter. In Section 2.1, a closer look at the general development of the electricity demand is provided. New means of production will be established as PV cells, wind power plants and small decentralized biomass and hydro power plants. A discussion of the opportunities and challenges of PV cells is given in Section 2.2. Furthermore, demand patterns are also expected to change due to the electrification of the heating sector, discussed in Section 2.3, and the transport sector, discussed in Section 2.4. Finally, Section 2.5 addresses the capabilities of batteries even though they are not incorporated in this thesis.

2.1 General Development

In this section, it will be investigated how the future electricity demand will develop. The scenarios are taken from [1] (2011), which estimated the electricity consumption up to 2050 in the framework of the "Energiestrategie 2050". Key drivers for the development of the demand were identified therein and a reduced list is given below:

- Technological advances through research
- Population growth
- Increase of GDP per capita
- Increase in electricity intensity (ratio of electricity consumption to GDP)
- Electricity price
- Partial electrification of heating and mobility sector

As can be seen, a variety of different factors influence the demand in the future ranging from technological to demographical to economical factors. Particularly the partial electrification of the heating and mobility sector will have a profound influence on the load pattern. Thus, it will be further analyzed in subsequent sections. For the

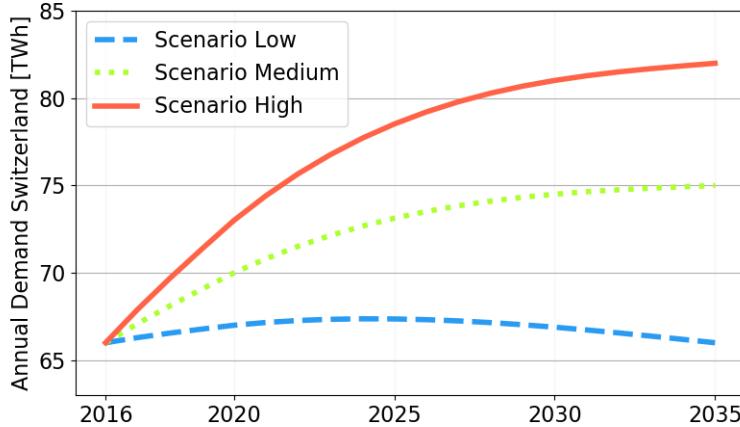


Figure 2.1: Scenarios for the electricity consumption in Switzerland (taken from [1]).

general development, three different scenarios will be analyzed as suggested in [1] and drawn in Fig 2.1. Note that the value for 2016 is not explicitly given in [1], and hence, it was linearly interpolated between the values for 2010 and 2020 for the medium scenario.

2.2 Photovoltaic Cells

With the vote to phase out nuclear power, an alternative is needed to substitute for the missing energy production. PV cells have the potential to provide access to a large power reservoir, which explains their boom in recent times. In the first paragraph, the scenarios for the PV cell deployment will be introduced. Subsequently, challenges of solar power will be analyzed before the final paragraph addresses the relationship between consumers and producers.

The total potential for suitable locations of photovoltaic cells on buildings is estimated as 18 TWh per year [21] (2015). If possible sites in the mountains and on public infrastructure (parking slots, reservoirs, dam walls, etc.) are included, Swissolar estimates the potential in Switzerland as high as 35 TWh per year. However, a precise prediction of the development is difficult since the incentives provided by different players (government, companies) are hard to predict. This results in a large uncertainty about the consumer behaviour which is accounted for by considering two different scenarios for the deployment rate. As "PV scenario medium" denoted is the scenario according to the "Energiestrategie 2050" [22] (2017) (see Fig 2.2). A rather optimistic scenario was published by SwissSolar [23] (2017) (therein referred to as "Roadmap Swissolar"), which assumes the use of more than 80% of the available potential until 2050. It will be denoted as "PV scenario high" and the corresponding development is shown in Fig 2.2. In [21] it is noted that exploiting this potential would not be profitable for a household and would require substantial government subsidy.

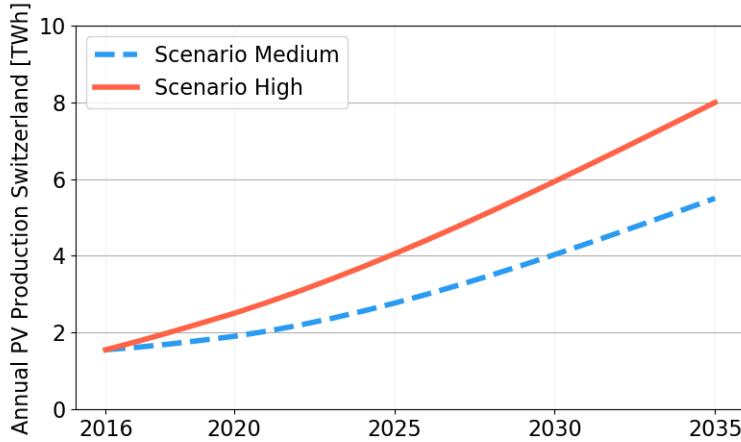


Figure 2.2: Scenarios for deployment of photovoltaic cells.

Despite its many advantages, there remains one last problem with solar power: its intermittent nature. Consequently, the planned installation of PV cells on a large number of rooftops will lead to a decreasing planning reliability for grid operators [9]. This challenge is accentuated by the expectation that around 90-95% of all renewable power plants (installed power) will be connected to the medium- and low-voltage power grid [9]. On a sunny summer day, solar power production will steeply increase in the morning and sharply decrease in the late afternoon, making it difficult for large power plants to provide the difference to the consumption level. Hence, additional power generation capabilities with a fast response time or the import of electricity is necessary to meet the demand during low photovoltaic production [24]. Another means to provide load balancing are batteries, which will be discussed in more detail in Section 2.5.

Besides the influence on the stability of the grid, PV power generation will also transform the relationship between consumers and the grid operator. In today's power grid, the power flow is top-down from the producer (e.g. a large nuclear power plant) to the consumer (e.g. a household). In the future, due to local generation of electricity, the power flow will be bidirectional [25]. Thus, consumers will become prosumers, consuming and producing electricity. As a consequence, new forms of power grid management will be required in a decentralized fashion.

2.3 Heat Pumps

There is an ongoing endeavour of households to rely less on external parties providing oil and gas for heating purposes. Among the main drivers are fluctuating oil and gas prices, economic growth and environmental concerns [26]. An alternative heating system not running on fossil fuels is a heat pump, which achieves high efficiency while facilitating the de-carbonization of the heating sector. First, the heat pump market will be analyzed

before challenges and opportunities are pointed out and the effective number of heat pumps per node is estimated.

Figure 2.3 depicts the number of newly installed HP from 2005 until 2035 taken from [2] (2008). It is observed that the deployment of heat pumps in new buildings has already saturated and is slightly decreasing. In contrast, the replacement of heat pumps with other heat pumps is increasing and it will be the largest share in the future. A different behaviour is observed for conventional heating systems (mainly oil or gas), of which only very few are replaced by heat pumps. Their replacement by heat pumps is harmed by the high economic burden [26].

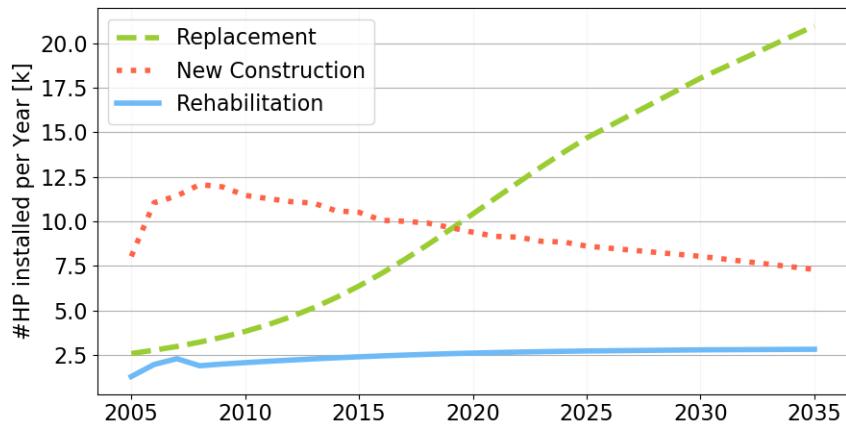


Figure 2.3: Number of heat pumps installed each year split into categories (taken from [2]).

When not adequately managed, heat pumps can introduce additional stress on the LV distribution grid by their irregular demand peaks. Already a few per node can lead to significant variations in the peak load rendering it difficult for operators to accurately predict the demand. This uncertainty will be further addressed in Section 3.4. On the other side, in a future smart grid, the flexibility of heat pumps could be exploited to provide stability and cost efficiency to the grid by different means [26]:

- Reducing heat pump active power demand in times of local under-voltage and increasing it in times of over-voltage
- Planning of operation to avoid simultaneity of load peaks on household and grid level
- Using aggregated load profiles to shift peaks
- Increasing the self-consumption rate of PV cells

For the prediction pipeline introduced in Chapter 1, the number of heat pumps per low-voltage grid area will be required. According to the Federal Statistical Office [27], the number of buildings in use for living is estimated as 1.7 millions in 2016. In the same

year, approximately 200k heat pumps (without boilers and single room heat pumps, see Fig 4.19 in [2]) were installed. This leads to a ratio of 1 HP/ 8.5 households. For the development of the number of installed heat pumps the assumption is made that the number of households per low-voltage grid area remains constant. This seems to be reasonable since most cities are assumed to grow mainly in the urban agglomeration. Furthermore, it is assumed that heat pumps are replaced by heat pumps and conventional heating systems by conventional heating systems. To justify the latter assumption: 2'500 replacements per year correspond to 1 system per 50 households in 15 years.

The discussed scenario herein is very general and does not take into account local policies in terms of development of new technologies, which may vary heavily between cities. For example in Basel, it is intended to replace most of the gas based heating systems by electrical heat pumps. For an accurate prediction for this specific city, this fact should be taken into account.

2.4 Electric Vehicles

The rising numbers of Electric Vehicles (EV) will cause additional stress on the distribution grid if not properly managed. On the other side, control of the charging pattern can provide opportunities for load balancing. In order to analyze the impact of EV further, the notion "electric vehicle" needs to be explained. Different types of electric vehicles are known in the literature: Battery Electric Vehicles (BEV), Plug Hybrid Electric Vehicles (PHEV) and Range Extended Electric Vehicles (REEV). BEV and PHEV have by far the largest share [28] and, for the seek of notation, they will be denoted as EV from now on. The first part of this section addresses the environmental pollution caused by the transport sector. Subsequently, the scenario for the uptake of EV will be discussed. Finally, a more detailed view will be considered for the challenges and opportunities.

The transport sector is responsible for around 36% of the total energy consumption in Switzerland, whereof 96% is covered by fossil fuels [28]. This demand has been relatively stable over the past 15 years and continues to pollute the environment. It is estimated that mobility emits 33.3% of CO₂ with the largest share of 65% emitted by passenger cars [28]. Therefore, the transition of the transport sector from internal combustion engine driven vehicles to electric driven vehicles is key to meet the goals of the Swiss Energy Strategy 2050. However, EVs are only "greener" if the electricity is produced by renewable energy sources, whereas if the energy is from a coal-fired power plant the emission will be higher than for an efficient conventional vehicle [29].

Purchasing an EV is becoming more appealing as the energy density, and hence the range, of batteries steadily increases while costs are falling [3]. Depending on the scenario, the share of EVs in sold automobiles will expand to 25% for the most conservative scenario or 55% for the most optimistic scenario in 2035 [29] (2015) (see Fig 2.4). Thus, 13% to 20% of all passenger cars will be electric at that time [29]. In this work, the middle scenario "EFF" (efficiency) in Fig 2.4 will be considered.

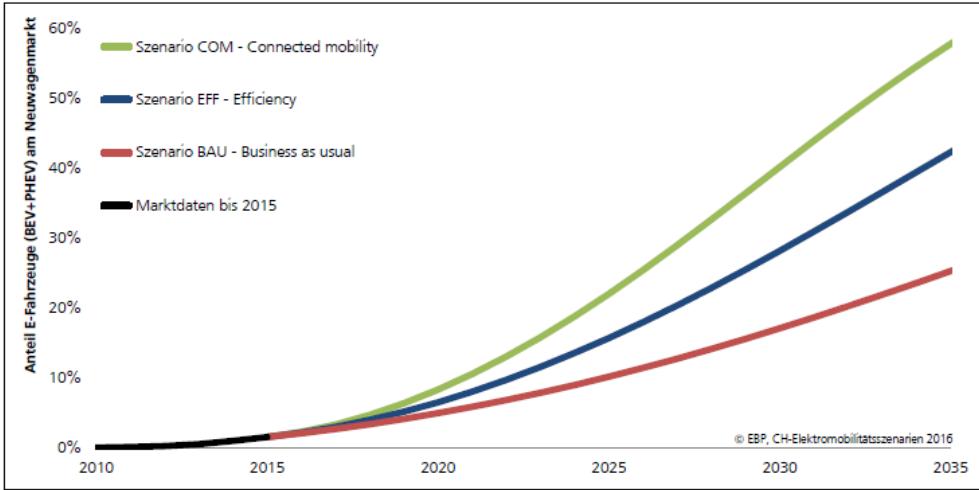


Figure 2.4: E-Mobility scenarios in Switzerland (taken from [3] (2015)). Shown is the share of EVs in sold vehicles from 2010 to 2035.

A small calculation is provided to obtain the number of EVs per low-voltage grid area of 60 consumers. The stock of EVs in 2016 is estimated to around 10'000 [30]. Furthermore, each year, 300k vehicles were newly registered for the past 40 years [30]. Hence, it will be assumed that the same number of vehicles will be registered in the future. The yearly growth rate is taken from Fig 2.4 and a life span of 8 years is assumed (note: there are no long-term studies yet available). Finally, the assumption is made that EVs are equally spread across all households in Switzerland (1.7 millions). As a result, following numbers are obtained for the development of EVs per low-voltage grid area with 60 households:

Table 2.1: Number of EV per low-voltage grid area of 60 consumers.

2015	2020	2025	2030	2035
< 1	2	8	16	26

A recent study in Great Britain suggests that increased peak loads will occur in the evenings due to EV charging if no actions are taken. In weak LV grids, local problems can be faced due to this behaviour [31]. In contrast, incentivizing consumers to charge their car during night could support the grid by balancing the increased production of wind energy. Furthermore, EVs provide the potential to store electricity decentralized allowing to balance the intermittent renewables and to optimize the use of the distribution grid [29]. This technology is known as vehicle-to-grid and allows the EV to communicate with the power grid to provide demand response services by either returning electricity

to the grid or by reducing its consumption.

In the framework of this thesis, the charging pattern of EV will be included by synthetic charging profiles. However, simulating the intelligent management of the additional battery capacity in the grid is out of scope.

2.5 Energy Storage

The focus of this section lies on the role of energy storage in electricity grids and particularly in future smart grid applications. Storage systems have the main task of adding flexibility and balancing to the grid. In the past, pumped hydro plants were for example used to consume energy during night, when excess electricity is available (mainly from nuclear power plants) and to produce energy around noon, when the demand is highest. In the future, storage will be required to balance the intermittency of PV and wind power. [32] states that grid operators are only able to compensate for the fluctuations in production when the intermittent renewable share is lower than 15% to 20%. Hence, storage will play a key role in the upcoming years.

Another problem arising with particularly solar energy is the discrepancy between the production by PV cells and the consumption by heat pumps. The first will generate mainly in summer around noon [24] whereas the latter will use during early and late hours in winter [33]. Storage on different scales will be required for balancing the difference. Not only the scale but also the right storage duration and response time is of importance. On a transmission grid level, pumped hydro and compressed air systems are a suitable choice [34]. On a local level, battery storage systems convince due to the lower capital outlays and faster response times [34].

In 2010, the storage in the EU energy system was estimated to around 5% of total installed capacity [32]. A broad differentiation between technologies can be made based on the scale i.e. if it is suitable for centralized or decentralized storage. A comprehensive overview of storage technologies and their main characteristics (response time, power rating, storage duration, etc.) can be found in [34]. Large scale pumped hydro plants, heat storage as e.g. molten salt, or compressed air storage support load balancing on transmission grid level [34]. On distribution level, batteries are promising mainly due to a fast response time. However, [1] (2011) expected that this technology will most likely not be important on a national level before 2030 due to the high costs. However, it is again noted that it might vary heavily between different regions as for example some DSOs require a consumer installing a PV cell to also install a battery. Nevertheless, storage of electricity will not be considered any further in this work.

Chapter 3

Data Preparation

In this chapter, the pre-processing of the data is explained. Real world measurements usually contain inconsistencies, noise or data gaps with the need of preparation before use. Thus, prepared data is taken from [35] which was processed in terms of anomaly detection, data cleaning and missing value imputation.

In this thesis, only residential consumers are considered. Therefore, Section 3.1 explains how the data is split into small households and large industries. The prepared data at hand, Section 3.2 introduces the mathematical framework of K-means clustering which is used to group similar load profiles. For the density predictions, temperature simulations are performed in Section 3.3. Even though the procedure is just outlined for temperature data this applies in the same way for irradiation data. Next, starting from raw heat pump data, resampling is used to explore possible consumption patterns in Section 3.4. A similar procedure is applied for the electric vehicle charging pattern in the last Section 3.5.

3.1 Extraction of Residential Consumers

For a DSO it is important to know how the load for a substation will change in order to plan for updates of power grid lines and transformers. Therefore, the focus of this work lies on residential areas where a number of individual households are bundled together in a low-voltage grid area. Industrial consumers are excluded from the analysis since further economic insights would be needed to predict changes in electricity consumption. Furthermore, large industries often have their own transformer which requires a separate planning for the DSO if for example a new warehouse needs to be built. To decide about the cut-off, the consumption of each consumer is averaged over the period of a year and the sorted data is plotted in Fig 3.1. One observes that more than half of the consumers have an average consumption of less than 250 W. Thus, the cut-off is chosen at 1500 W, which includes around 4'900 consumers out of the 5'000 available.

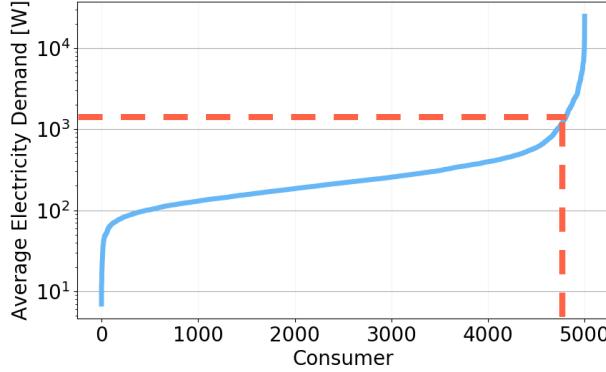


Figure 3.1: Electricity demand averaged over one year and sorted by increasing size.

3.2 Clustering

A single load profile shows a high level of randomness since the start-up of a single device, as for example an oven or a washing machine, is visible [18]. However, only patterns due to some internal coherence of the signal or external influences can be accurately captured by a forecasting algorithm. It is proposed to use K-Means clustering to cluster consumers of similar behaviour, thus extracting the common underlying patterns. Another advantage of this approach is the scalability in the sense that for new profiles, only the classification into one of the existing clusters is necessary to make use of the already known predictions (note: this assumes that the profiles used for obtaining the clusters are representative for the data set).

A summarized version of the K-Means algorithm is provided in Subsection 3.2.1 derived from [36] and [37]. Subsequently, the application to smart meter data is explained in Subsection 3.2.2.

3.2.1 K-Means Algorithm

K-Means is an unprobabilistic method for unsupervised learning which helps to discover underlying patterns in the data. Similar data points are grouped into clusters represented by a cluster center μ_k . Each data point is hard-assigned to a single cluster (therefore unprobabilistic). K-Means tries to minimize intra-cluster distances in a metric space, where the to be minimized objective function is given by:

$$J = \sum_{i=1}^N \sum_{j=1}^K r_{ij} \|\mathbf{x}_i - \mu_j\|^2, \quad (3.1)$$

with \mathbf{x}_i a data point and r_{ij} the assignment of point i to cluster j i.e. $r_{ij} = 1$ if \mathbf{x}_i is assigned to μ_j and $r_{ij} = 0$ otherwise. An analytical solution does not exist, and

hence, an iterative scheme is used. First, initial centers are chosen which is often done by K-means++ (also implemented in the Scikit library of Python):

1. Start with random data point as center.
2. Add centers 2 to k randomly, with probability proportionally to squared distance to closest selected center:

Data: x_i data points
Result: μ_j cluster centers
 Z normalizer of probability distribution;
for $j = 2$ to k **do**
 i_j sampled with probability $P(i_j = i) = \frac{1}{Z} \min_{1 \leq l \leq j} \|x_i - \mu_l\|_2^2$;
 $\mu_j \leftarrow x_{i_j}$;

Then, the iterations $t = 1 : N$ consist of two alternating steps. In the E-step, the cluster assignments z_i are updated:

$$z_i \leftarrow \arg \min_{j \in [1, \dots, K]} \|\mathbf{x}_i - \mu_j^{(t-1)}\|_2^2, \quad (3.2)$$

before in the so-called M-step, the new cluster centers are calculated based on the assignments:

$$\mu_j^{(t)} \leftarrow \frac{1}{n_j} \sum_{i:z_i=j} \mathbf{x}_i \quad (3.3)$$

where n_j is the number of data points assigned to a given cluster. This update scheme is guaranteed to converge to a local minimum.

3.2.2 Clustering of Load Profiles

As an objective, the load profiles should be clustered according to their similar shapes. Through an analysis of several load profiles, it has been found that usually the demand during the week is higher than at the weekend. Thus, the consumption during the week is more important for overloading of the grid and clustering will be performed based on a typical weekday.

Although only small consumers are considered, the individual load profiles have usually different scales. For grouping consumers with similar consumption profiles into the same cluster, the scale needs to be similar as K-Means minimizes an intra-cluster distance. To select the most appropriate standardization/ re-scaling of the data, the complete prediction pipeline needs to be taken into account. Dividing each profile by its mean allows a convenient extrapolation of the mean consumption later. It has also been experimented with applying the logarithmic transform to the raw profiles before

clustering but the results were inconsistent. Hence, the log-transform is applied to the cluster centers after the clustering.

Two different strategies are applied for obtaining the features used in K-Means clustering:

1. An individual profile experiences a lot of randomness as noted earlier, and hence, averaging is performed of the week days over the course of a year i.e. $5 \cdot 52$ days are taken into account.
2. To account for the seasonal structure, an averaged week day profile is obtained for each season i.e. $5 \cdot 13$ days are taken into account for each of the averages of spring, summer, autumn and winter. The number of features is reduced by down-sampling each profile by a factor of four (to have the same number of features as in (1)).

To compare the two strategies, the later (Chapter 5) introduced density plots are compared against the smart meter data for three different nodes in 2016. The number of not captured values served as metric, by which no obvious favourite could be determined. However, some of the clusters obtained by the latter approach posed a problem for the linear regression model, resulting in a constant line as best fit. Taking a closer look at the respective clusters, it was observed that they consist of only very few profiles exhibiting a large jump of consumption roughly in the middle of the training data set. Since no feature could model this behavior, the best fit was a constant line in between the lower and the higher consumption level. Therefore, the first approach is chosen to cluster similar consumption profiles.

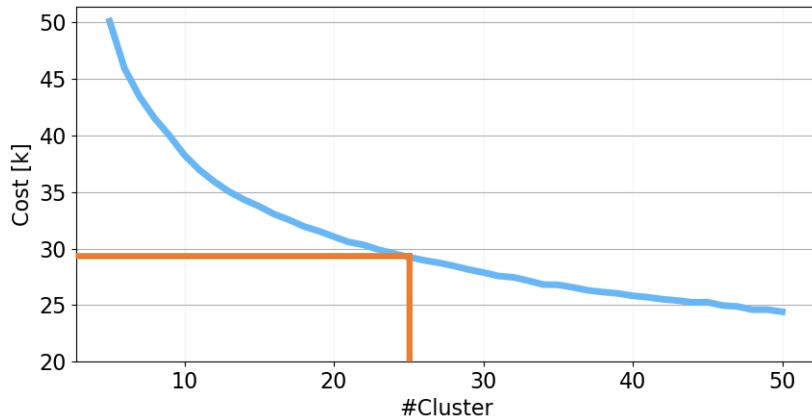


Figure 3.2: Clustering cost versus number of clusters.

K-Means requires to choose the number of cluster centers in advance. A parameter study is conducted with different numbers of clusters and the clustering cost (objective function) is reported. Fig 3.2 depicts the number of clusters on the x -axis and the corresponding clustering cost on the y -axis. A common heuristic is to choose the number

of centers according to the so-called elbow-methods in which k is chosen such that increasing k leads to a negligible decrease in cost. As a result, $k = 25$ is chosen for the number of centers.

A selection of four resulting cluster centers is shown in Fig 3.3. Some obvious features distinguishing the different clusters will be pointed out whereby the discussion is adapted from [35]. It should be noted that the profiles are an average over a large number of days. Hence, they only provide an idea of the consumption pattern but do not reflect the actual consumption on a specific day. An increased demand around noon and in the evening suggests that cluster 1 probably consists of restaurants and cafeterias. On the other side, a relatively constant profile between 08:00 and 20:00 indicates that cluster 2 might belong to a shop or a department store. Contrarily, cluster 3 shows a distinct peak at 12:00 and 18:00 with an increased consumption until around 22:00 which would be expected from a typical household. A few people will have lunch at home explaining the smaller peak but most of the people will return home around 18:00 and cook or use electronic devices consuming energy until late evenings. Finally, cluster 4 mainly consists of devices sensitive to electricity prices and with some flexibility in their demand such as boilers and heat pumps programmed to consume at night when prices are lower (Switzerland knows a day and a night tariff, with latter starting at 22:00).

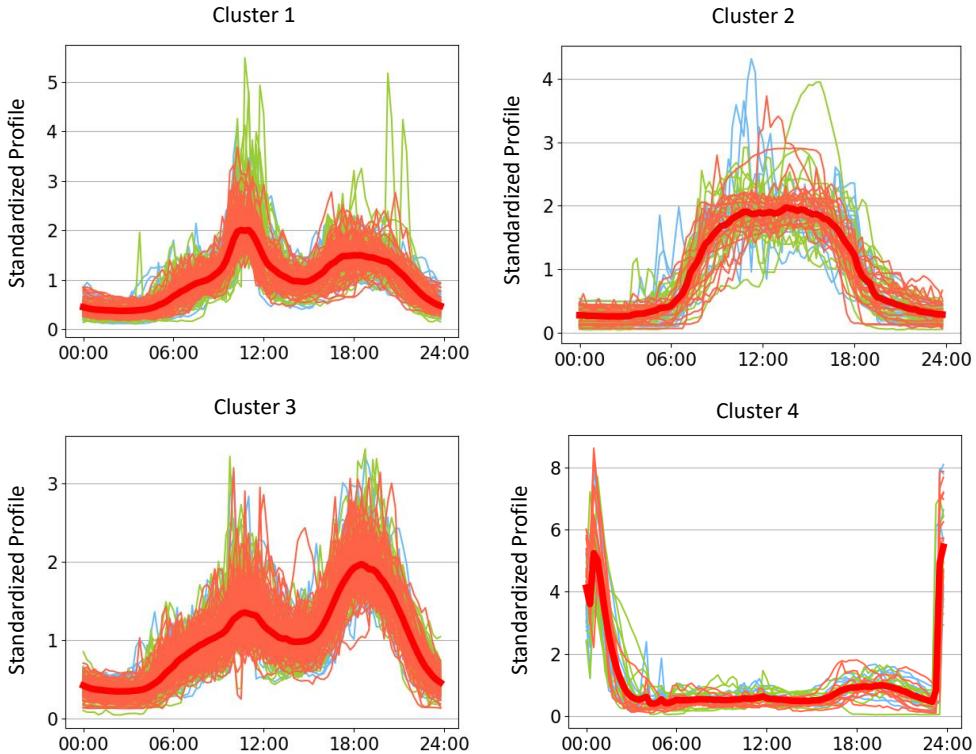


Figure 3.3: Four distinct daily profiles obtained by clustering showing different types of consumers.

3.3 Temperature Simulation

As will be shown in Section 4.1, the electricity consumption behaviour depends heavily on the outside temperature. This fact is exploited to generate a density distribution for the demand in a specific year. Weather data is available for the city of Basel from the last 10 years and will be used for this purpose. As noted in [17], insufficient variations in the data cause problems when calculating statistical properties. Therefore, the approach suggested in the same paper will be used to generate artificial temperature series.

The methodology for the temperature bootstrapping is referred to as *Double Season Block Bootstrap*. Each temperature series is divided into segments, from which one randomly samples to obtain a new series. It is important to preserve the structure of the data during this process as for example the seasonal pattern, the intra-day pattern and the inherent serial correlation of the series. For reasons of illustration, the algorithm to assemble the series (without start and end block) of a single year is depicted in Algorithm 1 (note: temperature data is available in 15 minutes intervals, and hence, a day consists of 96 data points). Figure 3.4 illustrates the generation of a new profile from three available time series.

Algorithm 1: Double Season Block Bootstrap

Data: Temperature series for N years $T_{\text{year } n \in [1, N]}$
Result: Double Season Block Bootstrap Temperature Series $T_{\text{bootstrap}}$

```

 $T_{\text{bootstrap}} \leftarrow \text{start};$ 
 $m \leftarrow \text{integer, number of days};$ 
 $\Delta m \leftarrow \text{integer with } \Delta m < m, \text{ number of days};$ 
 $\Delta s \leftarrow \text{integer with } \Delta s < m, \text{ number of days};$ 
while not end of  $T_{\text{bootstrap}}$  do
     $n \leftarrow \text{select year at random};$ 
     $s \leftarrow \text{select start uniformly at random from [current position in } T_{\text{bootstrap}} \pm \Delta s];$ 
     $l \leftarrow \text{select length uniformly at random from } [m \pm \Delta m];$ 
     $T_{\text{bootstrap}} \leftarrow T_{\text{year } n}[s \cdot 96 : (s + l) \cdot 96]$ 

```

where Δs is a small integer number for shifting the starting point which has little effects on the time series patterns [17]. Further, m is an integer number representing the mean length of the block and Δm is an integer number for varying the length of the blocks. Both need to be chosen such that the inherent serial correlation is preserved. Considerable auto-correlation of the temperature is observed for time delays up to six days, which led to the choice of $m = 8$, $\Delta m = 2$ and $\Delta s = 3$.

Moreover, special treatment of the start and the end block is necessary to satisfy the constraints (e.g. we cannot choose a starting points which is before the start of the series). Using this approach a high variability is achieved with a relatively small amount of data.

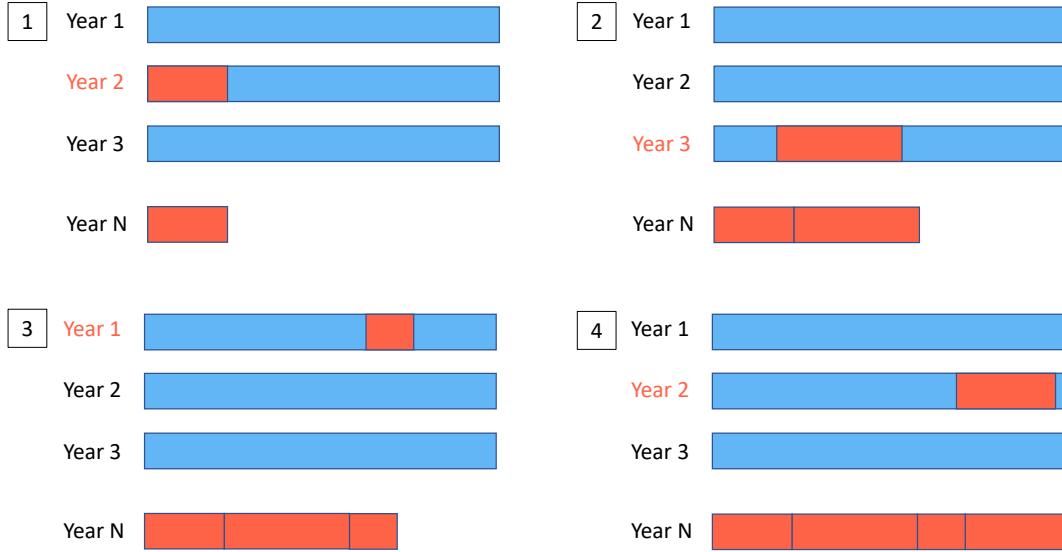


Figure 3.4: Double Season Block Bootstrap.

3.4 Heat-pump Simulation

In Section 2.3, it has been argued that about each 9th household has a heat-pump installed. This translates to about six to seven HP per low-voltage grid area with 60 consumers. HP data is available for a total of 20 heat pumps with a resolution of 1 minute. Hence, down-sampling is necessary to match the sampling frequency of the smart meters. Some of the profiles have missing values, which are imputed by inserting the values of the preceding period. This preserves the later discussed seasonal pattern for relatively small intervals of missing values. It is pointed out that the outlined generation of artificial HP profiles preserves the underlying pattern but is not coupled to the temperature bootstrapping. It is out of scope to model the HP consumption as a function of temperature which would achieve this coupling. The remaining part of this section is divided into three paragraphs, the first analyzing a specific HP profile and the last two introducing the heat pump simulation procedure.

A different resampling procedure than for the temperature needs to be applied since data is only available for a single winter and the HP profiles are very different to one another such that no exchange of segments is possible. Fig 3.5 shows a consumption profile of a heat pump over the course of a winter (left) and for a week in the middle of November (right). For extended periods, the electricity demand is very close to zero before it suddenly jumps up to full-power. This type of control strategy is called "bang-bang" control [38]. The turning on of the heat pump is almost random, and hence, with a very high probability, the consumption could be some time earlier or later. Despite this randomness, a certain daily pattern is observable with higher consumption in the early

morning hours. To appropriately include this information in a density plot, resampling is applied as explained in the next paragraph. Another observation is that the electricity consumption heavily depends on the season as indicated by the rolling mean (14 days) in Fig 3.5. During January, the frequency of turning the heat pump on is the highest which has to be taken into account during resampling.

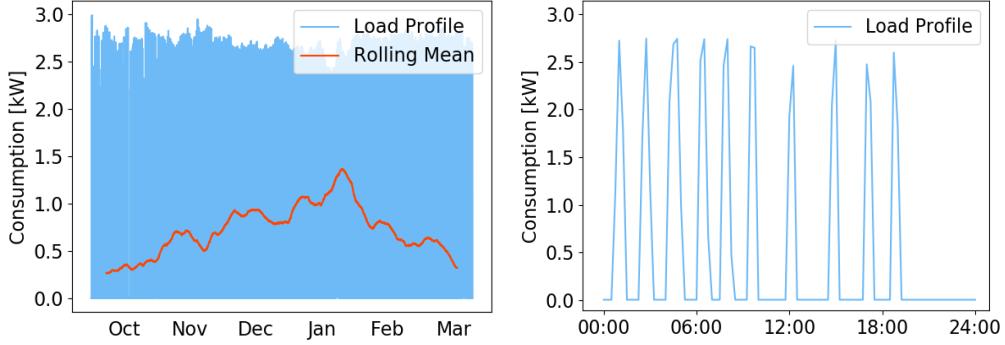
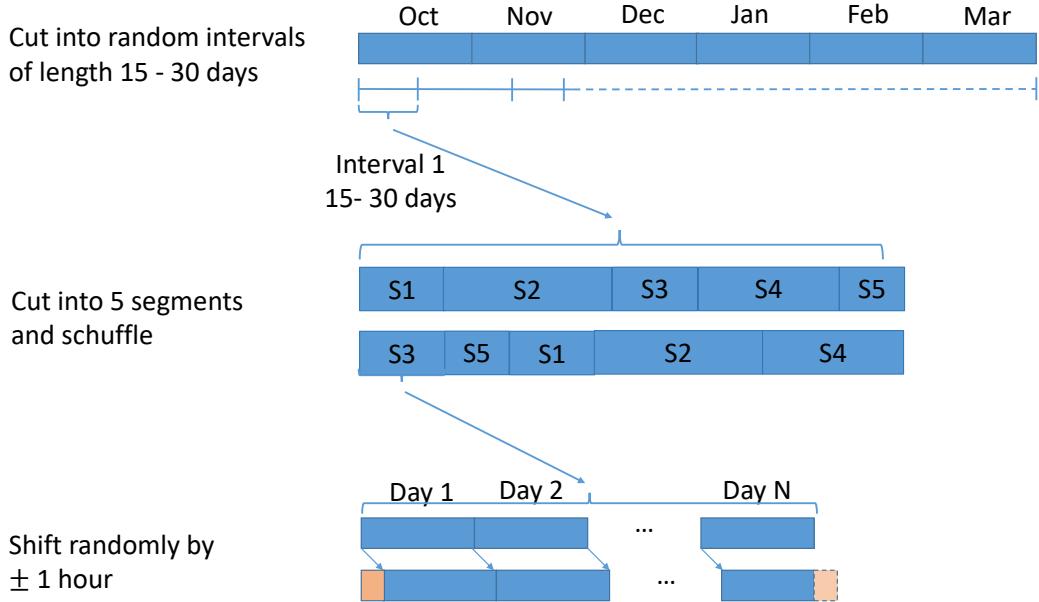


Figure 3.5: Heat pump load profile for winter (left) and a day in November (right).

When generating new heat pump profiles from the existing ones, the two described patterns have to be preserved. An illustration of the process is provided in Fig 3.6. The seasonal structure is maintained by cutting the HP profile into intervals of length between 15 and 30 days. Reshuffling is only allowed within an interval but the intervals remain at the same position in the time series. Each of these intervals is split into 5 segments of random length which are then reshuffled. To preserve the intra-day pattern, the segments are cut in between days. The randomness in turning on a HP is accounted for by randomly shifting each segment by shift $\in [-60 \text{ min}, -45 \text{ min}, \dots, 45 \text{ min}, 60 \text{ min}]$. The part exceeding the time frame is appended at the end or beginning, respectively (light orange section in Fig 3.6). In the final step, the individual pieces are re-assembled to form an artificial heat pump time series.

One last issue remains to be addressed, namely that heat pump data is only available during the winter period. Generally, it can be assumed that heat pump operation takes place mainly during that time of the year. Evidence is provided by the rolling mean in Fig 3.5, which decays towards both sides. However, since it decays not completely to zero, there would be an abrupt transition from the period with HP operation to the period without HP operation. To resolve this issue, the series is allowed to shift uniformly by shift $\in [-14 \text{ days}, \dots, 14 \text{ days}]$. It can be interpreted as the possibility of the period with cold temperatures to start already earlier in the year or to persist until later in the year, thus shifting HP operation. Using the described approach, 500 artificial heat pump profiles are generated from each heat pump profile.

**Figure 3.6:** Procedure for HP resampling.

3.5 Electric Vehicle Simulation

This section addresses the generation of a density distribution for electric vehicles. To model the consumption pattern, synthetic load profiles of 76 EVs are available with a sampling time of 15 minutes. In contrast to the heat pumps, data is available around the year but without a specific seasonal pattern. Nevertheless, a resampling procedure similar to the one explained in the last section will be applied, which would allow to use profiles with a seasonal structure. This is motivated by the observation that the traffic volume exhibits a maximum during summer as depicted in Fig 3.7. The figure shows the monthly average of the traffic on three different roads in the city of Basel. On the other side however, a distinct intra-day pattern is observed as has been discussed in Section 2.4. For the further analysis, the density plot of a synthetic load profile is shown in blue in Fig 3.8. A peak demand is observed at around 18:00 with significant charging activities until 01:00. Even though there exists a distinct pattern, the exact time of consumption is to some extend random.

The procedure for resampling is the same as for the heat pumps due to the similar structure identified (intra-day pattern, possibility to include seasonal pattern). One difference is that the random shift of the segments is increased to shift $\in [-1hr, \dots, +3hrs]$. This is motivated by the observation that the synthetic profiles show a very strong peak at 18:00 (Fig 3.8). This led to unrealistically high consumption peaks in the long-term forecasts since one would expect that consumer behaviour gets more diversified with larger numbers of EVs. A second difference lies in the fact that no special treatment of

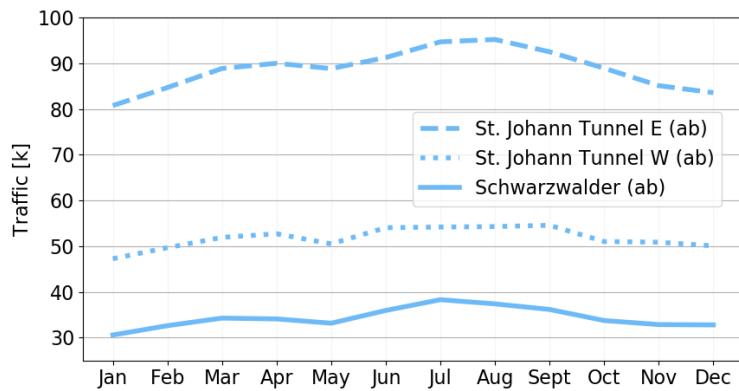


Figure 3.7: Traffic for three roads in Basel.

boundaries is necessary since data is available for a complete year.

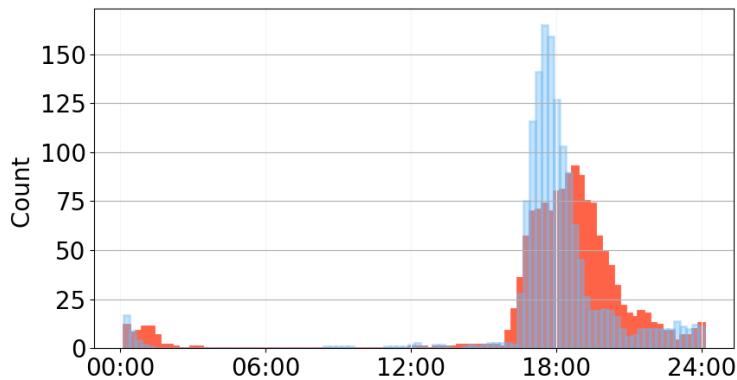


Figure 3.8: Histogram for the charging of electric vehicles obtained from synthetic profiles over a year. The blue histogram is obtained from the original synthetic profile and the red from the resampled profile.

Chapter 4

Semi-Parametric Model

This chapter deals with the implementation of the model derived from [17]. An important consideration for every model is the feature selection, which is addressed in Section 4.1. Then, Section 4.2 points out the specific structure of the model. Therein, a logarithmic transform of the data is motivated to improve the fitting accuracy. The next two sections take a look at the fitting of the scaled cluster centers. In Section 4.3, the linear regression method is introduced as a first method for the function approximation. A methodology to enforce sparsity and improve the fitting accuracy is applied. Furthermore, it is recognized that filtering of the data is necessary to remove serial correlation of the residuals and make the underlying assumptions of linear regression valid. As a second method, a multi-layer perceptron will be fitted in Section 4.4 with the previously introduced filtering process in place as well. Finally, the two models will be compared in terms of accuracy and computational cost in Section 4.5

4.1 Feature Selection

As for all statistical and machine learning methods, the selection of useful features is key for a high accuracy. For the modelling approach introduced in the next section, the prediction relies solely on external variables. Hence, one needs to understand which are the main influences on the electricity demand of a household. Therefore, the consumption split depicted in Fig 4.1 will be analyzed. It is noted that the consumption of an individual household relies heavily on the behaviour of the inhabitants, which is hardly predictable. Nevertheless, the general trends extracted by the clustering approach can be predicted based on external variables as will be discussed next. A large share is influenced by the outside temperature as for example heating (21.7%), hot water (13%) and air conditioning (6.8%). Moreover, it is also expected that people spend more time indoors during winter time, which could lead to a higher electricity consumption for entertainment (6.6%) and lightning (7.3%). Additionally, the latter will be influenced by the irradiation.

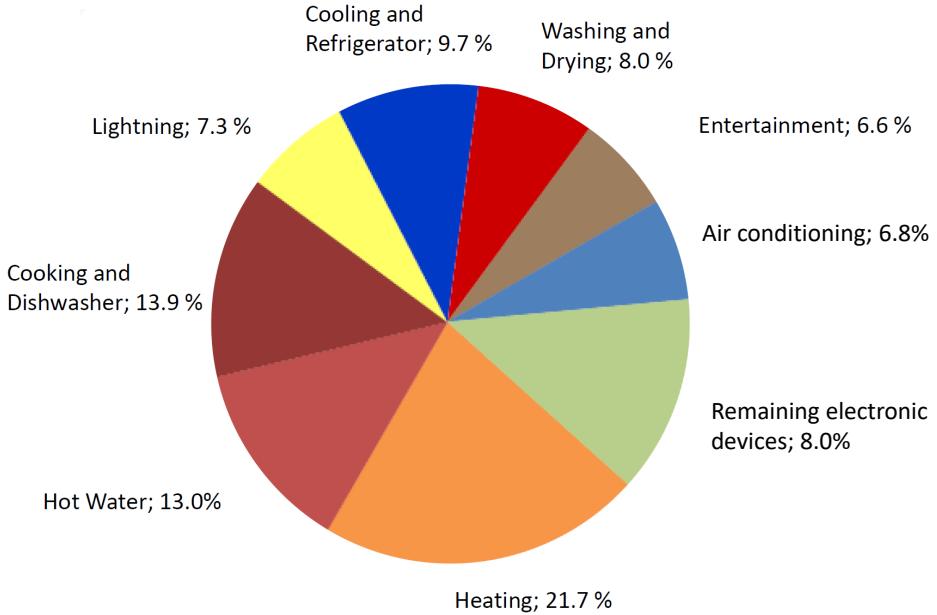


Figure 4.1: Split of electricity demand for a typical household (taken from [4]).

To conclude, the weather plays a key role in determining the electricity demand. Different combinations of lagged and processed temperature and irradiation values were compared based on the choices in [17] and [18]. By means of the mean square error, the best-performing combination is selected and presented next. This metric is chosen as it is widely accepted when dealing with time series and further, the models applied determine the solution vector based on the same metric making it a consistent choice.

Temperature: Selected are the current value, the values with 15 min, 30 min, 1 hr, 1.5 hr, 2 hr and 2.5 hr delay. This ensures to adequately capture the temporal trend. Furthermore, the average, minimum and maximum temperature of the past 24 hours ($t-96 : t$, since data is available in 15 minutes intervals) are chosen as well as the average over 24 hours from the previous two days ($t-288 : t-192$ and $t-192 : t-96$). This provides insights into recent trends of the weather. To capture the slow dynamics as for example the thermal inertia of a house, averages of the past 7 and 28 days are included as well. Finally, a seasonal rolling mean allows to incorporate seasonal variations giving a total of 15 features.

Irradiation: As is shown in Fig 4.1, lightning influences the electricity consumption as well. Additionally, solar irradiation helps heating up buildings, thus, reducing electricity consumption of a heat pump. Thus, selected are the current value, the values with 15 min, 30 min, 1 hr, 1.5 hr and 2 hr delay giving a total of 6 features.

Calendar Effects: As noted in [17], calendar effects such as the day of the week and holidays influence considerably the electricity consumption. Hence, individual data vectors are provided for each day of the week and data vectors for the day before holidays, holidays and the day after holidays giving a total of 10 features.

By fitting a model for each time of the day, the intra-day structure of consumer behaviour is implicitly part of the modelling approach. To sum up, 31 exogenous variables are selected to explain the electricity demand.

4.2 Model Structure

A separate model will be fitted for each quarter-hourly time frame denoted by $p \in [0, 1, \dots, 95]$. It is observed that demand patterns change throughout the day, thus motivating the use of different models for the different times. This statement is supported by Fig 4.2 which shows a scatter plot of the scaled demand (divided by mean consumption before clustering) against the current temperature for two different clusters. The dependency on temperature can change considerably such that for a specific cluster (right plot) the scaled demand is independent of temperature and a few hours later a non-linear dependency is observed. To limit the number of models considered, key times were identified as noon, when PV production is high, and as evenings, when EV charging takes place. Thus, the periods are restricted to $p \in [40, 48, 56, 64, 72, 80]$.

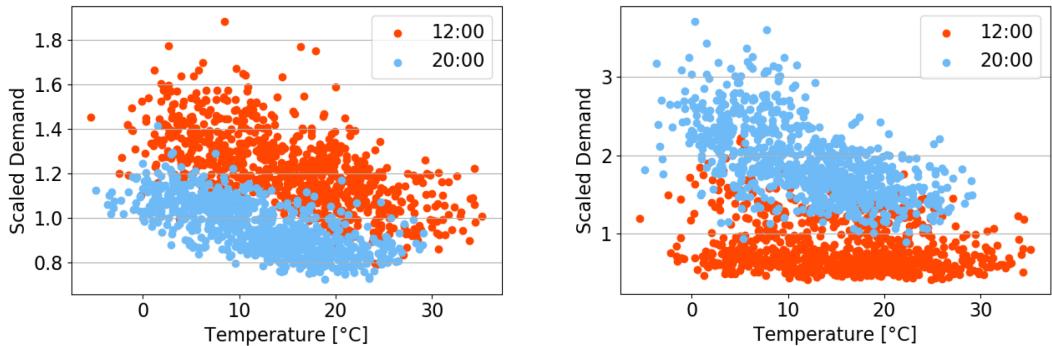


Figure 4.2: Scaled demand versus current temperature for two different clusters.

Based on the feature analysis in the previous section, the model can be written as in Eq 4.1. Note that the natural logarithm is taken before fitting, which is motivated by the increased performance identified during model validation. A second reason is the observation that the variance of the residuals changes throughout the year. However, both models (linear regression and the multi-layer perceptron) assume a constant variance. Using a logarithmic transformation will help to meet this assumption.

$$\log(\bar{y}_k(T)) + \log(y'_{k,p}(t)) = z_k(T) + h_{k,p}(t) + f_{k,p}(w_t) + g_{k,p}(v_t) + \epsilon_{k,p}(t), \quad (4.1)$$

where variable k is used to denote the model for a specific cluster, p is the period during the day, t the day during the year and T the year. Further, w_t is a vector of recent temperatures and v_t a vector of recent values for the irradiation. Hence:

- $\bar{y}_k(T)$ is the mean demand of model k in year T
- $y'_{k,p}(t)$ is the standardized demand of model k in period p at day t
- $z_k(T)$ models the development of the mean demand of model k in year T
- $h_{k,p}(t)$ is the influence of the calendar effects (holidays, day of the week) on the demand of model k in period p at day t
- $f_{k,p}(w_t)$ models the temperature effects of model k in period p at day t
- $g_{k,p}(v_t)$ models the irradiation effects of model k in period p at day t
- $\epsilon_{k,p}(t)$ represents the model error of model k in period p at day t

A few things are pointed out next. First, the logarithm of the product of mean and fluctuating part is taken since the clustering approach required normalization of the demand curves such that all profiles have the same scale. Further, the functions on the right-hand side extract the exogenous variables discussed in Section 4.1. For example, $f_{k,p}(w_t)$ takes the vector of recent temperatures and extracts the 15 features required to model the demand at a specific instance. These functions modify their inputs (by a weight vector for linear regression, by weights and non-linearities for the multi-layer perceptron) such that the output matches the dimensions on the left-hand side.

In Sections 4.3 and 4.4, models will be introduced to fit Eq 4.1 by a linear combination of the input variable in the first and by a non-linear combination thereof in the latter section.

4.3 Linear Regression Model

In Subsection 4.3.1, the general theory for Linear Regression (LR) will be recapitulated. Subsequently, the fitting results are presented for the electricity demand (Subsection 4.3.2) and the PV production (Subsection 4.3.3).

4.3.1 Linear Regression Theory

Given a data set $D = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$ with $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$, the goal is to learn a mapping $f : \mathbf{x}_i \rightarrow y_i$ i.e.:

$$f(\mathbf{x}_i, w_0, \mathbf{w}) = w_0 + \mathbf{w}^T \mathbf{x}_i, \quad (4.2)$$

where w_0 is a bias term and $\mathbf{w} \in \mathbb{R}^d$ the weight vector. For the ease of notation, the bias can be included into the weight vector by augmentation i.e. $\tilde{\mathbf{w}}^T = [\mathbf{w}^T, w_0]$ and $\tilde{\mathbf{x}}_i = [x_{i,1}, x_{i,2}, \dots, x_{i,d}, 1]^T$. Next, the goodness of fit needs to be quantified to have a measure to optimize for. A common choice is the summed square error:

$$\hat{R}(\tilde{\mathbf{w}}) = \sum_{i=1}^n r_i^2, \quad (4.3)$$

where $r_i = y_i - f(\mathbf{x}_i, \tilde{\mathbf{w}})$ is the residual. Note that the dependence of \hat{R} on the data is dropped at it is fixed during optimization. Hence, the weight vector is given by:

$$\tilde{\mathbf{w}}^* = \arg \min_{\tilde{\mathbf{w}} \in \mathbb{R}^{(d+1)}} \hat{R}(\tilde{\mathbf{w}}). \quad (4.4)$$

Note the increased dimension of the vector $\tilde{\mathbf{w}}$ due to the included bias term. It should be noted that this optimization problem has a closed form solution. To discover it, let us define the data matrix $\mathbf{X} \in \mathbb{R}^{n \times (d+1)}$ with the j^{th} column given by $\tilde{\mathbf{x}}_j$ i.e. $\mathbf{X}_{:,j} = \tilde{\mathbf{x}}_j$. Further, the vector $\mathbf{Y} \in \mathbb{R}^n$ can be obtained by stacking the individual y_i 's. By differentiation and setting the gradient to zero, one obtains:

$$w^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (4.5)$$

Hence, Eq 4.4 minimizes the summed square error given the data D .

4.3.2 Electricity Load Profile Approximation

With the framework explained in Subsection 4.3.1, where y_i corresponds to the logarithm of the fluctuating electricity demand and x_i to the feature vector, the electricity load profiles of the individual clusters will be fitted with the features selected in Section 4.1. A separate model is fitted for each cluster and for each specific time considered. This was motivated by the observation that the demand pattern will change throughout the day, which was also noted by [17].

As noted in Subsection 3.2.1, clustering was performed with the profiles scaled by their mean consumption. Furthermore, the resulting centers are log-transformed before fitting since it was observed that better final results are obtained. As a result, the units of the entries of the solution vector will be such that the right-hand-side of Eq 4.1 matches with the units of the logarithm of the electricity demand.

To exploit sparsity, the temperature and calender variables were selected from the provided features by cross-validation for each model separately. In other words, only a subset of the 31 features is used to optimize the performance of the model. Therefore, the data is split into five folds, one fold is left out and the model is fitted on the remaining four. This is repeated for each fold and the Square Error (SE) is calculated for the left out fold. Subsequently, the variable is omitted which led to the largest decrease in SE. This procedure is repeated until the SE does not further decrease.

It was further observed that the error ϵ_t in Eq 4.1 after fitting the model is correlated with the delayed error ϵ_{t-1} . However, linear regression assumes independent and identically distributed errors. In order to remove the serial correlation, the time series can be filtered. Therefore, the new signal is defined by:

$$v_i = y_i - r y_{i-1},$$

$$\mathbf{u}_i = \mathbf{x}_i - r \mathbf{x}_{i-1},$$

where y_i , x_i are as defined in Eq 4.2 and correspond to the fluctuating part of the electricity demand and the feature vector, respectively. To estimate the regression coefficient, an AR(1) model can be fitted to the error ϵ_t which is done by the StatsModels library in Python. Now, the semi-parametric model is refitted with the same features as selected during cross-validation. For the forecasting, the error will be sampled from its underlying distribution (Gaussian) to include the uncertainty in the fitting process. Hence, the outlier corrected histograms of the residuals are reported after the fitting process of the adapted model. As outliers identified are residuals with a distance of larger than three standard deviations to the mean. They are simply projected onto the boundary of the distribution i.e. replaced by the mean values plus or minus three sigma, respectively.

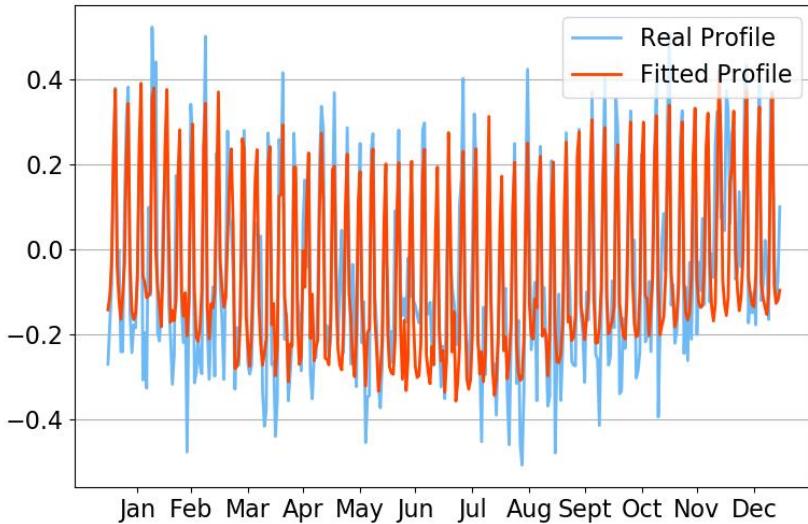


Figure 4.3: Scaled demand for cluster 0.

Fig 4.3 and 4.4 depict the logarithm of the fluctuating part of the load profiles at 12:00 for each day of a complete year for two clusters with superimposed fit in red. The profiles are shown in the logarithmic domain since they serve to get an idea of the fitting process accuracy, which itself operates in the logarithmic domain. For most clusters, the approximations accurately represent the real profiles. With the error resampling, it is accounted for possible inaccuracies in the fitting process

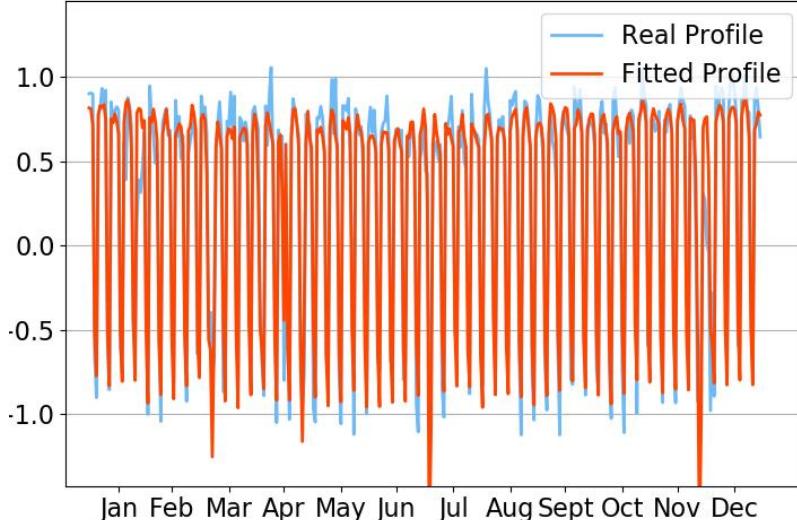


Figure 4.4: Scaled demand for cluster 10.

4.3.3 PV Production Profile Approximation

PV data is available over the same time horizon as smart meter data i.e. from April 2014 until July 2017. In contrast to electricity consumption profiles, photovoltaic production profiles are of similar shape and differ only in the scale. If normalized by the maximum value, the production curves coincide as shown in Fig 4.5. Hence, it is concluded that just a single profile, the average over the scaled profiles, needs to be considered for fitting purposes.

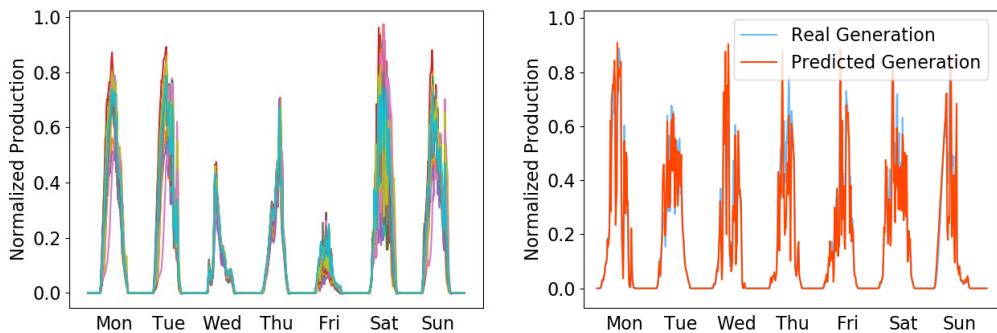


Figure 4.5: Left: PV production profiles normalized by the maximum demand. Right: Average of normalized PV production profiles (blue) and the corresponding approximation (red).

Feature selection is the next step to consider. It is clear that the photovoltaic electricity production depends on the irradiation. To figure out if considering temporal structure in the irradiation leads to an improved estimate, the function fitting procedure is repeated with delayed values included. For the evaluation, the Akaike Information Criterion (AIC) is chosen:

$$\text{AIC} = 2k - 2 \ln(\hat{L}), \quad (4.6)$$

where k is the number of estimated parameters and \hat{L} the maximum value of the likelihood function. The AIC is very similar for all considered cases (time delays up to an hour), thus the most parsimonious model is chosen without any past irradiation values included. A fitted profile for a random week is shown in Fig 4.5. One can conclude that the photovoltaic production is well approximated by the irradiation, as expected.

Next, the extrapolation of PV generated electricity will be considered. It is assumed that today, a single PV array is installed per low-voltage grid area. An analysis of the available data showed that an average cell has a maximum production of around $PV_{\max} = 2000W$. Thus, with the future scenarios discussed in Section 2.2, the production curve in year 20xx is obtained by:

$$PV_{20xx} = \frac{E_{\text{scenario } 20xx}}{E_{\text{scenario } 2015}} PV_{\max} PV_{\text{Approximation}}, \quad (4.7)$$

where $E_{\text{scenario } 2015}$ and $E_{\text{scenario } 20xx}$ are the annual production in Switzerland in 2015 and 20xx from the scenario under consideration. The result is a step-wise increase of the PV production at the end of each year. This avoids the explicit modelling of PV adaptation (for a modelling thereof, it is referred to [39]). $PV_{\text{Approximation}}$ is the scaled production curve obtained by feeding the derived model with irradiation data. To account for future uncertainties, the goal is to obtain a density prediction similarly as for the electricity consumption. Hence, bootstrapping of the irradiation data is suggested. This is done coupled to the temperature bootstrapping discussed in Section 3.3 to capture the correlation.

4.4 Multi Layer Perceptron

The second model used for approximating the load profiles is a multi-layer perceptron. In the first part of the section, some theoretical concepts of a MLP are introduced (Subsection 4.4.1). In the second part, the function approximation is explained for the electricity load profiles (Subsection 4.4.2). Fitting of the PV load profiles is omitted since linear regression was able to accurately approximate the function with only a single explanatory variable (irradiation). Hence, the increased complexity of the model is not justified.

4.4.1 Concepts of Artificial Neural Networks

Artificial Neural Networks (ANN) are inspired by the human brain with its neurons and synapses. Fig 4.6 depicts a general MLP with one hidden layer and serves as illustration to discuss the most important concepts. The inputs x_i are connected to the neurons in the first hidden layer via a bunch of weighted "synapses". Each neuron sums up its inputs, i.e. $\sum_i w_{i,1}x_i$ for the first neuron, and applies a non-linear activation function to it. This non-linearity is essential for the expressiveness of the network. For multiple hidden layers, the outputs of the previous hidden layers are the inputs for the next hidden layer. Finally, the last hidden layer is connected to one or multiple outputs of the network.

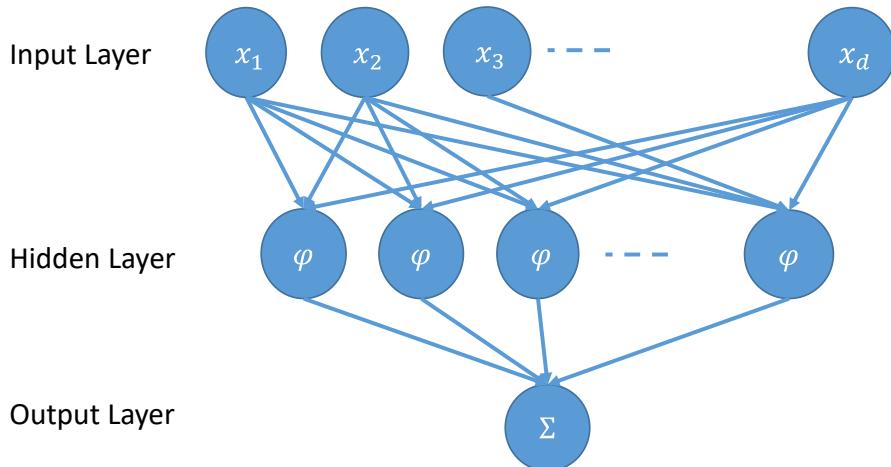


Figure 4.6: Structure of a multi-layer perceptron with one hidden layer.

Learning is performed by comparing the output of the network to the function that one intends to approximate and then back-propagating the error [36]. In the forward

propagation process, the input is propagated through the network to calculate the output. Given this information, the error of the approximation can be calculated. In the backpropagation process, the weights are adapted such that the error is minimized. A very simple adaptation scheme is gradient descent (GD), which updates the weights as follows [36]:

$$w^{(\tau+1)} = w^{(\tau)} - \eta \nabla E(w^{(\tau)}) \quad (4.8)$$

where η is the learning rate, an adjustable parameter to tune convergence, and $E(w^{(\tau)})$ the error evaluated with the old weights. The errors are defined by a loss function, which is characteristic for the given problem at hand. Since ANNs consist of many adaptive parameters, a large number of training examples are necessary to converge. As usually data is limited, multiple passes, called epochs, over the data set are necessary to reach convergence.

With increasing complexity of the model, the chance of overfitting increases. A convenient way to detect overfitting is to plot the training and the test set error as a function of the epoch. The training error should decrease during learning as the model becomes better and better in fitting the seen data. On the other side, the test error might start to increase again when the model starts to fit the noise in the training data thus reducing the generalization capability. One way of preventing overfitting is to introduce dropout regularization. It is a technique in which hidden units are randomly ignored with a probability p during each iteration. Another method would be early stopping, in which the training is stopped before the test error starts increasing. A third method is to enforce an upper bound on the magnitude of the weight vector for every neuron.

4.4.2 Electricity Load Profile Approximation

In the following, the architecture chosen in this specific work will be outlined. For the activation function, ReLu (Rectified Linear unit) is selected as it is a popular choice and shows good performance. Due to the use of a separate model for each specific time, the number of data points to train the model is approximately $N = 850$. To limit the number of trainable parameters, a MLP with a single hidden layer is chosen. The number of hidden units are determined by comparing different models in terms of sum of square test errors and selecting the minimizing one as proposed by [36]. Since the demand curves differ significantly, the procedure is repeated for two distinct profiles as shown on the left of Fig 4.7 and Fig 4.8. Depicted is the logarithm of the fluctuating part of the load profiles at 12:00 for each day of a complete year for two clusters with superimposed fit in red. On the right of the same figures, the RMSE is plotted against the number of hidden units. It is observed that the model shows good performance for relatively small numbers of hidden neurons. Hence, the number of hidden units is set to 10.

For the loss function, the mean square error is selected as the forecasting error should be minimized in terms of square errors. Optimization is carried out by the *adam_decay*

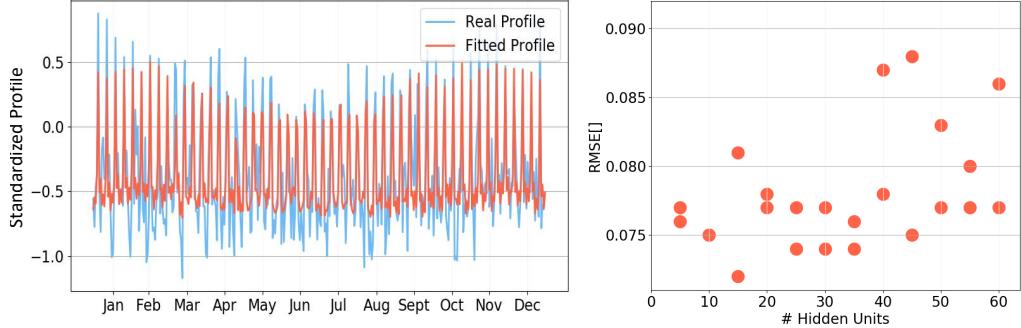


Figure 4.7: Load profile with one particular fit on the left, error versus number of hidden units on the right.

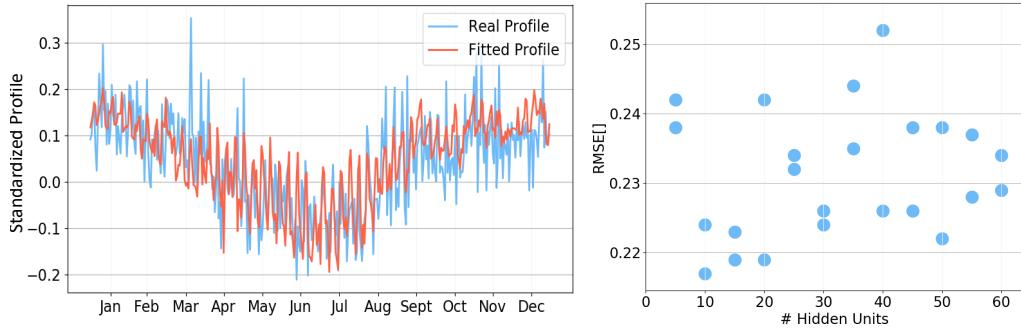


Figure 4.8: Load profile with one particular fit on the left, error versus number of hidden units on the right.

solver of Keras with a learning rate of $2e-4$. Overfitting is avoided by restricting the magnitude of the weight vector for each neuron to a value of 0.1.

The specified model is used to fit the demand curves of the clusters. Similar as in Section 4.3, the residuals of the fitted model are used to extract the AR(1) coefficient before the model is refitted to the filtered signal. As emphasized before, the objective function is not convex, thus, the model can get stuck at a poor local optimum. Therefore, the resulting approximations were analyzed and the fitting process was repeated if necessary.

4.5 Linear Regression versus Multi-Layer Perceptron

After discussing linear regression and multi-layer perceptrons separately, they are compared against each other in this section. In the first part, the obtained accuracy is analyzed by means of two different error measures. Subsequently, the computational efficiency of both models is compared.

4.5.1 Fitting Accuracy

To compare the accuracy between the linear regression model and the multi-layer perceptron, two different metrics are applied. The most well-known of these is the Root-Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_t |\hat{y}_t - y_t|^2}, \quad (4.9)$$

where N is the number of data points, y_t the real signal and \hat{y}_t the approximated signal. It is well suited to compare different models on the same data set but it has shortcomings when comparing models on data sets with different scales. An error measure tackling this problem is the Mean Absolute Scaled Error (MASE) [40]:

$$MASE = \frac{\sum_{t=1}^N |\hat{y}_t - y_t| / N}{\sum_{t=2}^N |y_t - y_{t-1}| / (N - 1)}. \quad (4.10)$$

It is scale independent but a division by zero might occur if the time series is constant. The first year of data is used to fit the model and the second year of data to validate the fit. For the linear regression approximation, the performance is reported once with and once without the exploitation of sparsity. In the next paragraph, the multi-layer perceptron is compared against the linear regression without feature selection. Subsequently, the comparison is repeated for the version with feature selection.

Fig 4.9 shows on the left a scatter plot of the RMSE of the MLP versus the RMSE of the LR and on the right a scatter plot of the MASE of the MLP versus the MASE of the LR. If a point lies below the diagonal this means that the error of the linear regression fit is smaller. Both error measures cannot suggest a superior model. For some clusters and times the linear regression is performing better for others the multi-layer perceptron. However, when analyzing the specific cases when either model outperforms the other, some conclusions can be drawn. Fig 4.10 depicts a case (logarithm of the fluctuations at specific daytime throughout the year) for which the LR fit is more suitable and Fig 4.11 depicts a case (logarithm of the fluctuations at specific daytime throughout the year) for which the MLP performs better. Generally, linear regression performs better for simpler and more obvious patterns. In these cases, the MLP tends to overfit due to its increased complexity. On the other side, for more complex profiles, the MLP is usually better approximating as it has more degrees of freedom.

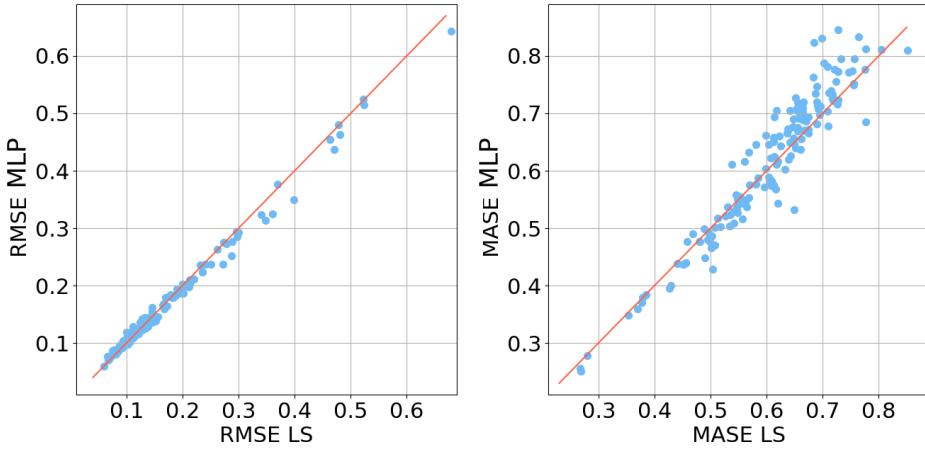


Figure 4.9: Comparison of errors for MLP fit and LS fit (without enforced sparsity).

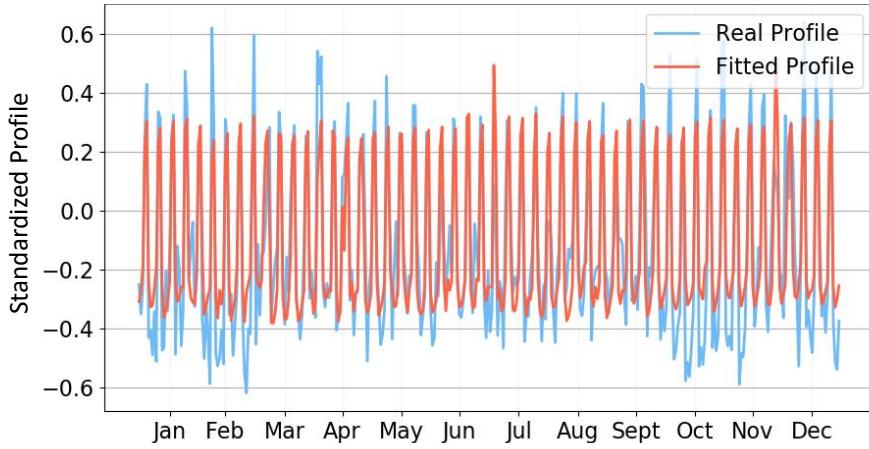


Figure 4.10: Profile with better linear regression approximation and corresponding fit.

In the case of linear regression, sparsity of the solution vector is exploited. This is omitted for the neural network as it would not be computational efficient to refit the model several times with different combinations of the features (refer to the next subsection). The comparison of the RMSE and MASE is illustrated in Fig 4.12 for the MLP and the LR with exploited sparsity. Particularly from the MASE plot it becomes obvious that the linear regression outperforms the neural network in this case. From the 150 approximated profiles (six daytimes for each of the 25 clusters), the LR has for 130 fittings the smaller MASE. This suggests that for some profiles overfitting occurs which can be overcome by just selecting the "best" features. Hence, the use of linear regression is justified for the future predictions.

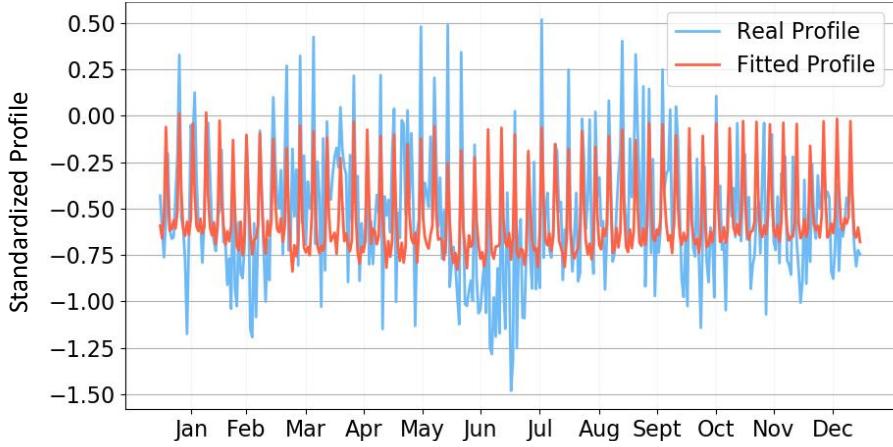


Figure 4.11: Profile with better multi-layer perceptron approximation and corresponding fit.

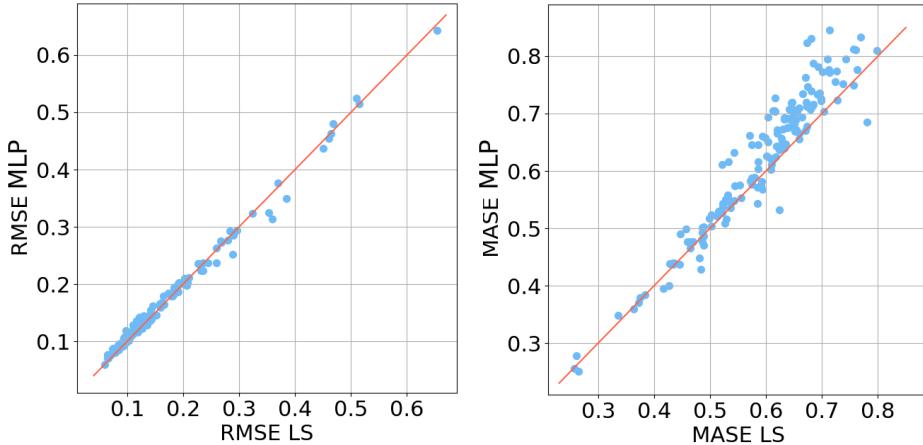


Figure 4.12: Comparison of errors for MLP fit and LS fit (with enforced sparsity).

4.5.2 Computational Efficiency

The run time has been computed for fitting the model, determining the regression coefficient and refitting the model to the training data. For the linear regression model without exploiting sparsity, this process takes in average $2.3ms$ (around $4.5s$ if sparsity is enforced whereby on average 15 out of 31 features are selected) whereas for the multi-layer perceptron an average run time of $217s$ is reported. The most expensive operation in the LR fit is the matrix inversion, which is still very fast with the highly optimized Numpy library in Python (written in Python and C). In contrast, the neural network uses backpropagation of errors with some kind of adaptation scheme (here: *adam_decay*) which is very slow compared to pure matrix operations. Furthermore, multiple passes of the data set need to be performed to achieve convergence. As a result, linear regression is multiple orders of magnitude faster and will be preferred for the rest of the work herein.

Chapter 5

Forecast

The forecasting can be performed with the fitted profiles and the assumptions about the adaptation of the different technologies. To start with, the procedure to obtain the forecasts is outlined in Section 5.1. For the further discussion, a selection of representative low-voltage grid areas is required. It is assumed that around 60 consumers are connected to a single node. Since only a fraction of all households has installed smart meters and only a fraction thereof has been processed due to computational power limitations, nodes with less than 50 consumers are excluded from the analysis. From the remaining nodes, two distinct were selected for the further analysis, denoted node A and node B. In order to justify the proposed approach for forecasting, Section 5.2 validates the estimated density function by comparing it against smart meter measurements from 2015. Finally, long-term forecasts are performed until 2035 in Section 5.3 and the results are discussed.

5.1 Future Prediction

This section discusses the assembly of the individual pieces to obtain information about the future electricity demand in the form of densities. In a second step, a discussion is provided of how each included technology will change the demand pattern.

First, the mean consumption in Eq 4.1 is extrapolated for each cluster by the trend identified in Section 2.1. To interpolate between the different years, a cubic spline is used with natural ends such that a smooth profile is obtained.

Next, the fluctuating part of the demand is reconstructed. Therefore, temperature and irradiation time series are derived from the bootstrapping. Zeros are substituted for the vectors representing the influence of holidays as only the general development in the future is of interest. However, it would be straightforward to include the specific holidays of each year. For convenience, it is assumed that each year starts with a Thursday as 2015 did, which is used to fill the vectors accounting for the day of the week. Using the fitted weight vectors, a profile is obtained for each time of the day, each cluster and

each temperature time series. By the latter one, a distribution over the consumption is obtained for each daytime and each cluster.

The generated EV and PV profiles cannot just be added up since the electrification of the transport and heating sectors are already included in the long term scenario. In the case of HPs, data is only available for the winter. For a smooth transition, the rolling mean is calculated for each profile and the smaller offset of (March, October) is subtracted from the HP profiles. As both offsets are roughly of the same size and the HP profiles were allowed to be shifted by ± 2 weeks, the transition is expected to be reasonably smooth. Nevertheless, the profiles will generally still have a mean, even though it is reduced. To account for that and to not influence the transition between the time with and without HP operation, the mean is subtracted uniformly from the complete year. For the charging of EVs, data is available for all seasons. Hence, the mean of each profile is calculated and subtracted.

In the mid- to long-term, a major change in consumption is caused by the deployment of PV cells. Density functions thereof are obtained based on the bootstrapped irradiation profiles and the fitted weight vector. As outlined in Section 4.5, the PV production is scaled and projected into the future.

To account for the fitting inaccuracy, noise is sampled from histograms extracted in Subsection 4.3.2. The histograms are of Gaussian shape with zero mean and constant variance as required by the assumption for linear regression.

Finally, all the different time series are added up year by year and a discrete distribution is obtained for each cluster. To make it clear, for each of the technologies and the fluctuating part of the electricity demand, 500 samples have been generated for each of the years from 2018 until 2035. These samples are combined together, leading to 500 time series for each year considered. Subsequently, the mean demand is added to each of these profiles depending on the electricity scenario considered. This allows to finally obtain a discrete density for each point in time.

5.2 Validation of Prediction Pipeline

The validation will be carried out for two different nodes. For node A (67 consumers), the two largest positions in terms of cluster count are similar to cluster 4 (price sensitive device) and cluster 3 (household) in Subsection 3.2.1. On the other side, node B's (68 consumers) two largest contributors are both similar to cluster 3 (household). To validate the proposed approach, the actual demand of 2015 measured by smart meters is compared with the box plots obtained from the prediction pipeline outlined in Section 1.3. In contrast to the explanations in the previous subsection, the holidays are included in the prediction of the load profiles. This allows to also capture the effects of Christmas etc. on the consumption pattern necessary when comparing with the real profile. Fig 5.1 (left) and Fig 5.2 (left) show the comparison between real profiles and reconstructed profiles for two different nodes, node A and node B, and two different times of the day. The profile constructed from smart meter measurements is plotted in green and the box plot is plotted in black with its mean in red. From both comparisons it can be concluded that the real profile is well captured by the densities. It proves that the applied model together with the temperature simulations are appropriate for this data set.

The box plots further reveal that the variance for the expected demand is larger during winter. This is the case since heat pump operation depends on the outside temperature and gets more energy intense at lower temperatures. Depending on if the heat pumps in a low-voltage grid area switch on at roughly the same time or in a staggered way, this will lead to large differences in power consumption on a distribution transformer level. If the heat pumps are not controlled actively, a very large uncertainty remains present about the demand at an aggregated level.

One difference between the nodes is pointed out showing the limitation of the used approach. For node B, the densities obtained by varying the temperature are narrow. This indicates that many consumers are not very sensitive to the outside temperature. Hence, the variability cannot be explained only by weather features but is due to the unpredictable behavior of the inhabitants as for example flexible working hours, etc.

For the remainder of this chapter, the analysis of the forecasts is just provided for node A as the discussion of node B is very similar. For the interested reader, the figures corresponding to the second node can be found in the appendix.

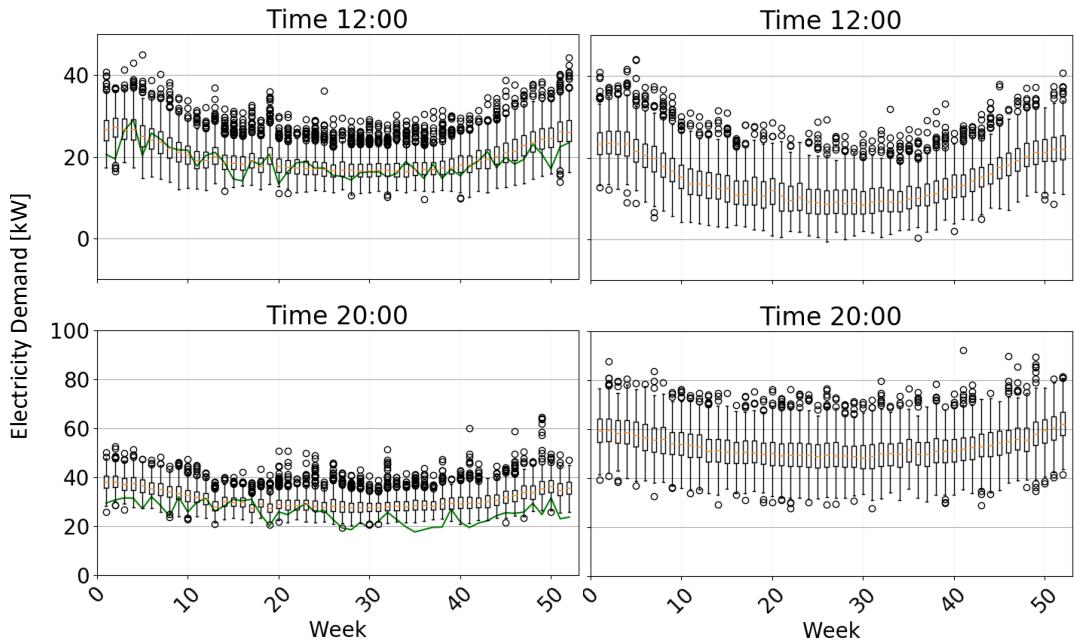


Figure 5.1: Electricity demand at two specific times on a working day over the course of a year for node A with 67 consumers. Left: predicted demand for 2015 with actual smart meter data in green. Right: predicted demand for 2035.

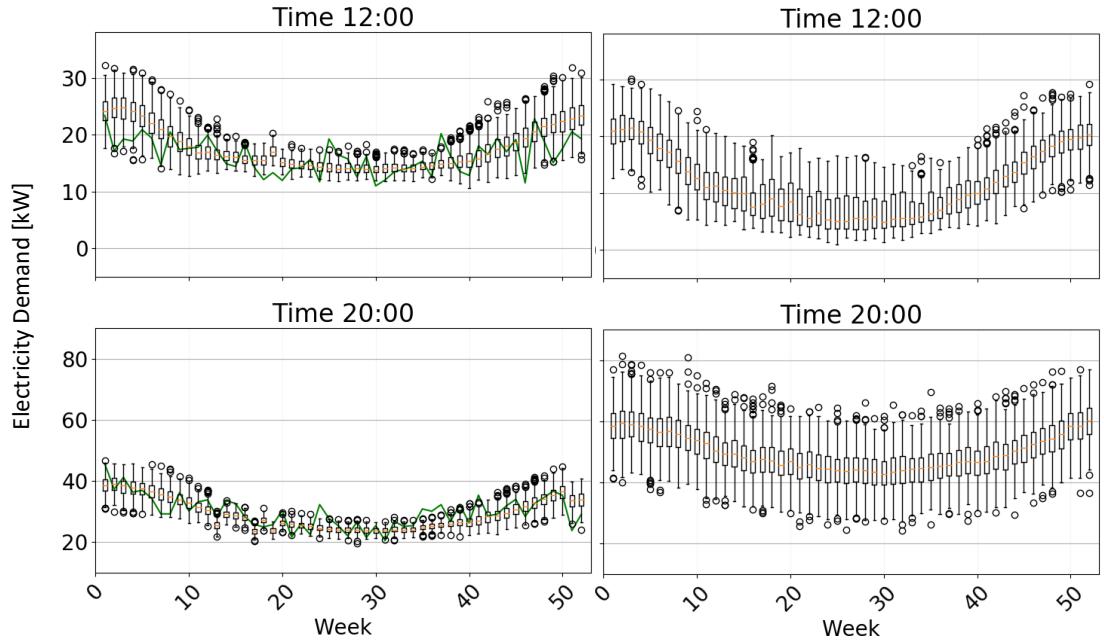


Figure 5.2: Electricity demand at two specific times on a working day over the course of a year for node B with 68 consumers. Left: predicted demand for 2015 with actual smart meter data in green. Right: predicted demand for 2035.

5.3 Electricity Demand until 2035

The forecast of the electricity consumption is subject of this section. To focus on the underlying pattern, the medium scenario for the electricity demand is combined with the high scenario of the PV production in the first three subsections. Further, the analysis is done for a working day as it has been observed that demand is generally higher at these days, and hence, grid overloading is more likely. It is started with the consideration of the demand at a specific time on a working day throughout the year to emphasize the seasonal pattern. Subsequently, the demand is studied during a day and key times are identified when the consumption undergoes the most profound changes in the next years. This knowledge in mind, a closer look is taken at the evolution of the distribution at the previously identified times. In the last subsection, the considered scenario (medium scenario for electricity, high scenario for PV) is put into context with the other considered scenarios.

5.3.1 Variation of Demand during a Year

In a first step, it is analyzed how the consumption pattern changes throughout a year. Fig 5.1 depicts the demand for a working day (e.g. Wednesday) over the course of a year for two different times (top 12:00, bottom 20:00). On the left, the seasonal variations are shown for 2015 and on the right, for 2035. The last year of the forecast horizon is considered since the changes will be the most significant, and hence, the easiest to identify.

The plots show that the same seasonal pattern is observable in the future. When comparing the demands in summer (around week 25) at 12:00, a drop can be recognized in 2035 compared to 2015. This is the influence of the PV production, which is the highest on a sunny summer day around noon. If consumers produce electricity themselves, less is needed from the grid and the demand decreases. In contrast, taking a look at the demand in winter (around week 5) at 20:00, one detects a large increase in power flow from the grid to the consumers between the two years. As the number of heat pumps, and hence their consumption, stays constant (see Section 2.3), this is the effect of the general increase in demand and the electric vehicle charging. Note that it has been argued that most people will charge their car when they arrive home after work. This puts additional stress on the grid and needs to be accounted for.

Finally, an explanation is provided for the changing variance between the years. At noon, the main driver is solar power, the share of which rises in the coming years. The larger the capacity of PV is in the grid, the larger are the fluctuations caused by alternating weather conditions. In the evenings, a massive increase in uncertainty is expected for the consumption. The charging behaviour of electric cars will profoundly change the demand pattern depending on if the EVs in a low-voltage grid area charge simultaneously or in a staggered way.

5.3.2 Variation of Demand during a Day

After considering a complete year, a daily profile will be analyzed to identify key times and motivate the choice in the previous subsection. The demand over a specific working day in summer is depicted in Fig 5.3. The densities with a green frame correspond to year 2015 and the densities with a red frame to 2035. Furthermore, the dashed green line represents the demand constructed from actual smart meter data at a specific working day in summer. First, it is observed that the applied method predicts the demand for 2015 accurately. For each time considered, the true demand lies within the boundaries of the density. Next, the demand is considered at the different times of the day. Around noon, the demand most probably decreases due to self-consumption of electricity produced from photovoltaic cells. Further, the width of the distribution increases, which was previously explained by the intermittency of solar power. In the evenings, power flow from the grid to the households will increase significant. As mentioned in the last subsection, electric vehicles will be charged in the evenings explaining this observation.

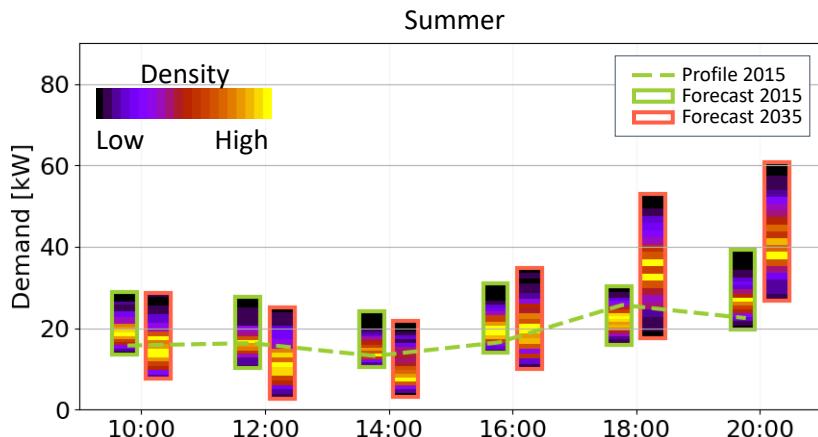


Figure 5.3: Electricity demand for different times during a working day in summer for node A with 67 consumers. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.

Fig 5.4 shows the variation of consumption during a working day in winter. For this season, the demand around noon is not altered much by PV production since the irradiation is much weaker in winter than in summer. For hours after 18:00 the increased demand due to charging of EVs is pronounced. Finally, when comparing the demand of the two seasons, it becomes obvious that more power will be consumed during the cold period of the year. Moreover, the variance of the densities is larger particularly in the evenings when heat pump operation is expected.

At this point, a discussion is provided about the impact of neglecting a smart charging of vehicles. This is done by considering the density at 16:00 and 18:00 in winter. First it is noted that by assumption the increase of the mean electricity demand is the same

for every hour of the day. Further, the influence of PV production for both daytimes is neglected (motivated by the early sunset during winter), the number of heat pumps per low-voltage grid area does not change and their energy intensity is roughly the same (see Fig 3.5). From the densities at 16:00 the change due to the mean electricity demand is visible since EV charging is still negligible at this time (see Fig 3.8). Therefore, the densities at 18:00 provide an insight into how large the impact of the charging behaviour is on the increase of demand and on the uncertainty.

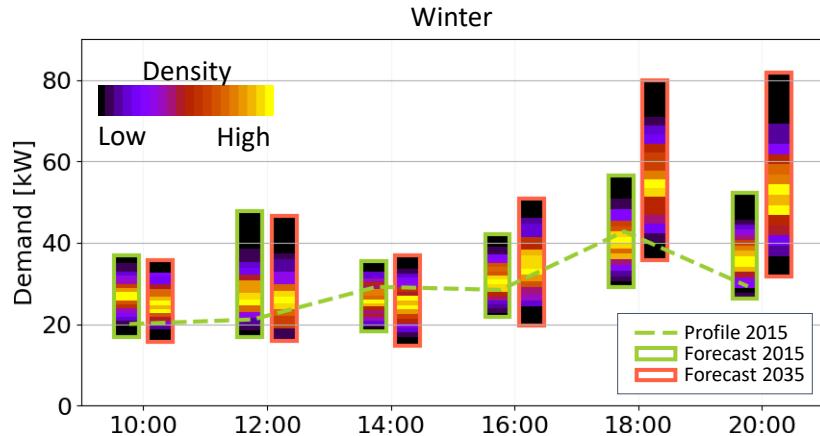


Figure 5.4: Electricity demand for different times during a working day in winter for node A with 67 consumers. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.

As a result of the analysis provided in this subsection, two daytimes are selected worth discussing in more detail: 12:00 during summer to emphasize the self-consumption of solar generated power and for 20:00 during winter to study the impact of the charging of EVs.

5.3.3 Variation of Demand over the Years

In the previous two subsections, the demand has been compared for the year 2015 and 2035. Now, the attempt is made to understand the evolution of the distribution over the years. As mentioned, this is done for a working day at 12:00 in summer and for a working day at 20:00 in winter.

Fig 5.5 shows the density for a single working day at 12:00 in summer. Initially, the mean increases slightly because the medium electricity scenario increases faster than the adaptation of solar power generation. However, after 2025, PV cell deployment accelerates leading to a decrease of electricity demand. In 2035, the mean of the density ends up considerably lower than in 2015. Furthermore, the variance is steadily increasing accounting for the uncertainty coming along with PV production. The self-production

depends mainly on the irradiation, thus the power flow from the grid shows large differences depending on the weather.

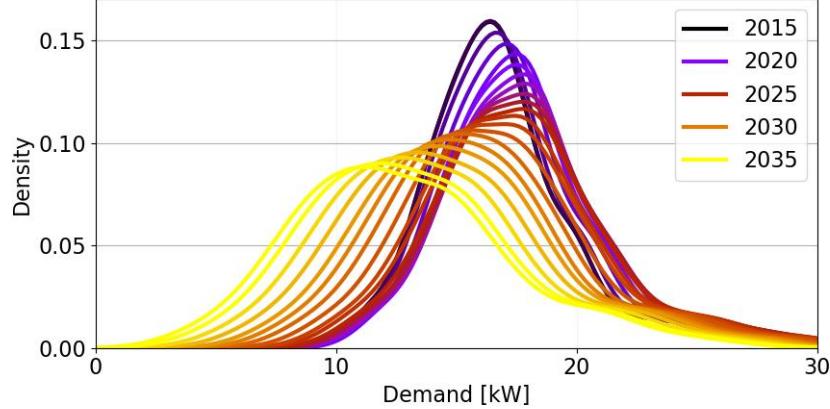


Figure 5.5: Evolution of the density on a working day in summer at 12:00 for node A with 67 consumers.

Next, to study the influence of charging EVs, the demand is analyzed for a working day at 20:00 in winter (see Fig 5.6). In this case, the demand monotonically increases when looking further into the future. On one side, the mean consumption of electricity increases and on the other side, the rising number of electric vehicles puts additional stress on the grid. Similarly as before, the variance keeps increasing because if all EVs charge at the same time, a large power flow is needed compared to the scenario of staggered charging.

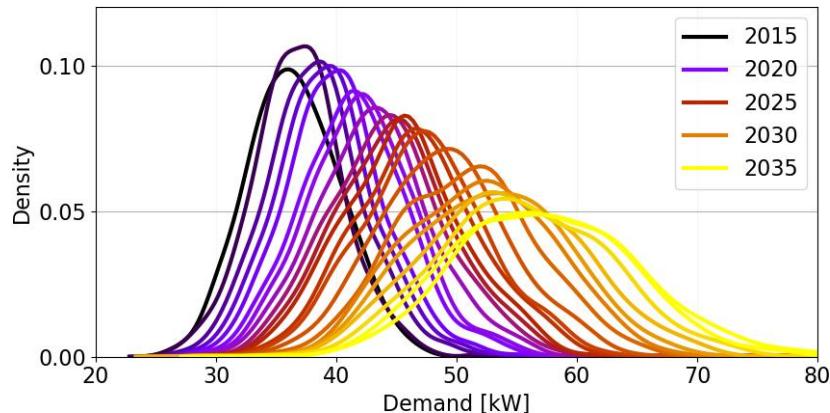


Figure 5.6: Evolution of the density on a working day in winter at 20:00 for node A with 67 consumers.

5.3.4 Development of Demand for different Scenarios

Now, context is given to the previously considered combination of medium electricity scenario and high PV scenario. Fig 5.7 illustrates the expected mean demand for a single working day in summer at 12:00 over the years. To stress this point, the density profiles are taken and the mean for every year is plotted here (note, this is not necessarily the region of the highest density in e.g. Fig 5.7). Depicted are the three electricity scenarios (blue: low, green: medium, red: high) with the lower bound given by the high PV scenario and the upper bound by the low PV scenario. Hence, the previous subsections analyzed the densities leading to the lower bound of the green area. Next, the observable trends will be examined. As has been mentioned in the discussion of Fig 5.5, the demand increases first for all but the low electricity scenario because the general demand increases considerably for all but the lowest scenario. However, when PV deployment reaches a certain rate, the demand will finally decrease at noon. For different combinations of the scenarios, the demand in 2035 may vary from around 8 kW up to 14 kW.

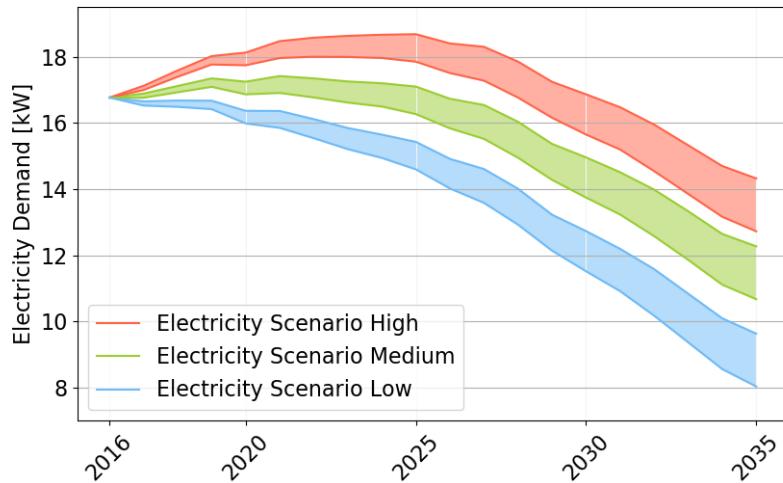


Figure 5.7: Demand for a working day in summer at 12:00 for node A with 67 consumers. The lower bounds are given by the high PV production scenario and the upper bounds by the low PV production scenario.

Fig 5.8 shows the demand for a working day in winter at 20:00. Plotted are the three different electricity scenarios combined with the two PV scenarios. However, since sunset is around 17:00 in January, there will be a negligible PV production, and hence, the curves coincide. For all scenarios, a large increase is observable caused by the fast adaptation of electric vehicles. Comparing the forecasted consumptions in 2035, a difference exists as large as 16% between the low and the high electricity scenario. For the medium scenario, the mean of the demand jumps by a factor of 1.5 between 2018 and 2035.

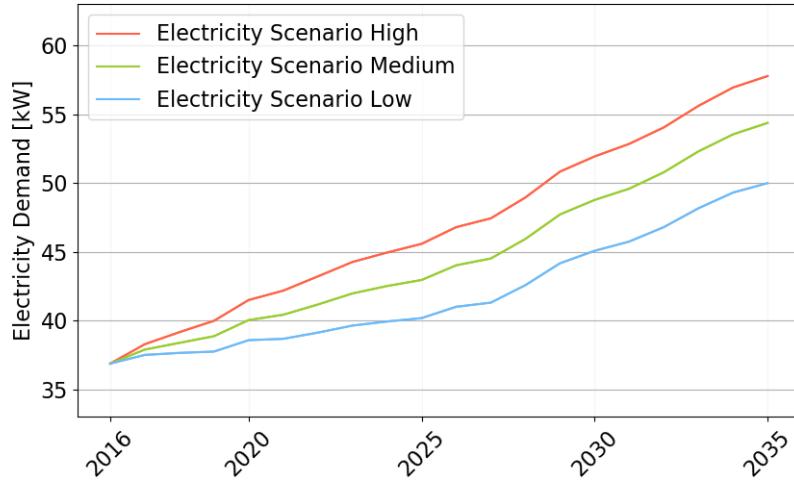


Figure 5.8: Demand for a working day in winter at 20:00 for node A with 67 consumers.

5.4 Overloading of critical Grid Infrastructure

This final section aims as giving a short qualitative discussion of the effects caused by the charging behaviour of EVs. The two grid components that will be addressed are the LV distribution lines and the distribution transformers.

In the last subsection, it has been observed that the demand increases by a factor of 1.5 to around 55 kW for the medium electricity scenario at 20:00 at a working day in winter. Keeping in mind that the region of not negligible density itself has a width of around 30 kW (Fig 5.6), this seems to be very critical for the grid. In the worst case, the grid has to deal with around 75 kW compared to 37 kW today.

In the introduction, it has been discussed that the radial line in the LV distribution grid are loaded with less than 50 % of their capacity. If one assumes that grid operators try to minimize margins in order to maximize the use of the infrastructure, the loading of these lines is expected to be close to 50 % already today. A doubling of the power flow would not lead to the substantially higher degradation of the lines but the supply reliability is no longer guaranteed. Hence, either many of the lines have to be upgraded or the peak load has to be limited to ensure security. Possible actions to achieve the latter are discussed in the last paragraph of this section.

It is more difficult to estimate a reasonable loading factor for distribution transformers in use. Considerable variations are expected depending on the grid operator, the location of the transformer and how many new consumers were connected since the installation of the transformer. Therefore, a discussion is provided with a similar starting point as for the grid lines. Assuming a current loading factor of slightly below 0.5 would allow to barely withstand the expected growth in power flow. However, since the peak load will occur in winter, outside temperatures are lower allowing the transformer to handle higher power flows. If, furthermore, the transformer provides some cooling as

e.g. A.F. (Air Forced), this can be exploited particularly during peak times to cope with the increased load and to postpone the replacement of the device.

In any case, measures explicitly targeting 20:00 in winter time are required to avoid overloading of the transformers. Besides incentivizing the holders of EVs to charge during night, it might be necessary to actively control the charging of EVs to reduce the peak. Finally, it needs to be emphasized that the forecasted demand depends heavily on the assumed scenarios which may vary substantially between cities.

Chapter 6

Conclusion

6.1 Summary

This work considered the low-voltage distribution grid by aggregating a small number of around 60 consumers to perform long-term forecasting. Smart meter data was provided for the city of Basel allowing to implement and test the model with real-world data. The purpose was to investigate the possible influences of different scenarios on the overloading of critical grid infrastructure. With the outcome, the distribution grid operator can plan for the time horizon of necessary grid reinforcements. Further, it provides indication which technologies are crucial and builds a foundation for active grid management by for example smart charging of electric vehicles.

For long-term predictions, it is necessary to consider different scenarios for the development of electricity consumption. Therefore, three scenarios have been considered for the general increase of the demand combined with two scenarios for uptake rates of photovoltaic cells. Furthermore, electric vehicles and heat pumps are incorporated by a single scenario for each.

In order to extract the underlying pattern of the consumers, K-Means clustering is proposed. The extracted cluster centers are then modelled by a term accounting for the mean demand and functions of temperature, irradiation and calendar effects. For the fitting process, linear regression and a multi-layer perceptron are considered. Since the model relies on temperature and irradiation, bootstrapping thereof has been used to obtain a density distribution. The same procedure was repeated for the PV production but with a dependence only on the irradiation. Heat pump and electric vehicle profiles were available and a richer set of profiles was obtained by resampling.

Comparison of the linear regression and the multi-layer perceptron revealed that the first performs as good as the neural network in terms of root mean square error but the fitting process is considerably faster. Hence, only linear regression was used for forecasting. A time horizon until 2035 was chosen to consider the long term trends. The changes in load pattern were analyzed and the key daytimes were identified as 12:00 in summer and 20:00 in winter. The first due to power production by photovoltaic cells

lowering the demand and the latter because peaks are expected due to the charging of electric vehicles and the operation of the heat pumps.

The key findings are that the expected mean electricity demand at 20:00 in winter increases by a factor of 1.5 between 2018 and 2035. Furthermore, in 2035 the expected mean electricity demand at 20:00 in winter varies by around 20% depending on the electricity scenario. This becomes even more problematic as the densities itself show variations of about 30 kW at an expected mean of 55 kW. Since electric vehicles are the main driver for this load peaks, in the mid-term, it is proposed to incentivize consumers to charge vehicles during night. In the long-term, most likely only active managing of the additional consumers will help to alleviate the problem.

6.2 Outlook

This final section addresses possible extensions based on the work in this project. As has been emphasized multiple times, the assumptions of the roll-out of the different technologies is very general and might vary significantly between regions. In a future work, one could study the influence of different adaptation rates of for example PV systems and conduct a sensitivity analysis. This would provide a more solid basis for regulations and for the planning activities of a DSO.

In this work, the influence of batteries was neglected since it likely does not play an important role on a national level before 2030. However, it was discussed that certain cities require consumers installing a PV system to also install a battery. Therefore, in certain regions this technology might play a role in load balancing, which could be studied in more detail. It could be done by for example generating artificial time series for the loading and unloading behaviour of batteries and combining it with the other time series. This can either be done as in the case of heat pumps, where actual consumption profiles were available, or by defining an underlying probability distribution, which assigns a probability of consuming or supplying energy to each time.

Another point to consider is the smart charging of electric vehicles and the effects on the peak demand. This could also include the consideration of vehicle-to-grid opportunities for load shifting. From the key times identified in this work, one could specify a probability distribution for the charging behaviour of EVs. While generating the charging profiles, one could check if the accumulated electricity demand exceeds a certain threshold, in which case charging at the given time for further vehicles is prohibited. Furthermore, one could consider how the possibility of charging at work influences the demand pattern.

In the fitting process of the load profiles, it was noted that the consumption pattern of some consumers is less dependent on weather than for others. An analysis of other influencing factors could help to increase the accuracy of the model further. For doing so, the consumption split can be more thoroughly analyzed to find other drivers for the demand of a household.

A rather obvious extension is to conduct forecasting beyond 2035. This would require more uncertain assumptions about the roll-out of technologies and the development of the general electricity scenarios. Hence, it becomes even more important to conduct a sensitivity analysis of the consumption on the different assumptions. An emphasize should be put on political scenarios and consequences on the adaptation of PV cells, electric vehicles, etc. Also the acceptance of a technology by the public should be beared in mind.

Finally, all the gained insights can be combined together in an accurate analysis of the consequences for the LV grid as for example the grid lines or transformers. It can be for example studied how different increases in mean and peak loading and resulting loading cycles influence the degradation of a transformer. This would allow to plan for updates or for measures counteracting the development of increasing electricity demand.

Bibliography

- [1] G. Andersson, K. Boulouchos, L. Bretschger, “Energiezukunft Schweiz,” 2011. ETH Zürich.
- [2] “Referenzentwicklung Wärmepumpenmarkt - Schlussbericht,” 2008. Bundesamt für Energie.
- [3] P. Haan, R. Bianchetti, “Szenarien der Elektromobilität in der Schweiz - Update 2016,” 2016. Ernst Basler + Partner AG.
- [4] A. Kemmler, S. Koziel, “Der Energieverbrauch der Privaten Haushalte 2000 - 2016,” 2017. Prognos AG im Auftrag vom Bundesamt für Energie.
- [5] “List of major power outages.” https://en.wikipedia.org/wiki/List_of_major_power_outages. Accessed on 5th June 2018.
- [6] K. Bullis, “The big smart grid challenges.” <https://www.technologyreview.com/s/414386/the-big-smart-grid-challenges/>, 2009. Accessed on 5th July 2018.
- [7] S. Tomic, “A study of the impact of load forecasting errors on trading and balancing in a microgrid,” *IEEE Green Technologies Conference*, 2013.
- [8] R. Bacher, “Liberalized electric power systems and smart grids,” 2018. ETH Zürich.
- [9] A. Ebner (BKW), “Consequences of the energy strategy 2050 for the grid infrastructure of a rural distribution system operator,” 2018. IEEE PES Seminar.
- [10] G. Hug, A. Ulbig, “Power system dynamics, control and operation,” 2018. ETH Zürich.
- [11] “IWB Geschäftsbericht - Fokus 2017,” 2018. IWB.
- [12] H. Hamedmoghadam, N. Joorabloo and M. Jalili, “Australia’s long-term electricity demand forecasting using deep neural networks,” *Arxiv*, 2018.
- [13] “Forecasting long-term electricity demand in the residential sector,” *Procedia Computer Science*, 2015.

- [14] N. Citroen, M. Ouassaid and M. Maaroufi, "Long term electricity demand forecasting using autoregressive integrated moving average model: Case study of Morocco," *ICEIT*, 2015.
- [15] H. Liu and J. Shi, "Applying ARMA-GARCH approaches to forecasting short-term electricity prices," *Energy Economics*, 2015.
- [16] S. Khuntia, J. Rueda and M. van der Meijden, "Long-term load forecasting using a multiplicative error model," *arxiv*, 2017.
- [17] R. Hyndman, S. Fan, "Density forecasting for long-term peak electricity demand," *Transactions on Power Systems*, vol. 25, no. 2, pp. 1141 – 1153, 2010.
- [18] T. Zufferey, A. Ulbig, S. Koch, G. Hug, "Forecasting of smart meter time series based on neural networks," *Data Analytics for Renewable Energy Integration*, 2016.
- [19] E. Mocanu, P. Nguyen and M. Gibescu and W. Kling, "Comparison of machine learning methods for estimating energy consumption in buildings," *PMAPS*, 2014.
- [20] E. Mocanu, P. Nguyen, M. Gibescu, "Big data application in power systems," *Deep Learning for Power System Data Analysis*, 2018.
- [21] V. S. Elektrizitätsunternehmen, "Photovoltaik und solarthermische Stromerzeugung," 2015.
- [22] "Schweizer Forschungs- und Technologielandschaft im Bereich Photovoltaik," 2017. Bundesamt für Energie.
- [23] D. Stickelberger, "Schweizer Photovoltaik-Markt im raschen Wandel," 2017. Swissolar.
- [24] "Energie-Spiegel: Facts für die Energiepolitik von morgen," 2012. PSI.
- [25] A. Gunzinger, "Power grid simulation for Switzerland," 2018. IEEE PES Seminar.
- [26] H. M. D. Fischer, "On heat pumps in smart-grids: A review," *Renewable and Sustainable Energy Reviews*, 2017.
- [27] "Bau- und Wohnungswesen 2016," 2016. Bundesamt für Statistik.
- [28] C. Schreyer, "Bedeutung der Elektromobilität aus Sicht der Energie- und Klimapolitik." Presenation at SWISS eSALON, 2017.
- [29] V. Fröse, S. Hobi, C. Jahn, F. Reutimann, M. Sager, P. Schild, S. Walter, "Bericht in Erfüllung der Motion 12.3652 - Elektromobilität. Masterplan für eine sinnvolle Entwicklung," 2015. ASTRA, ARE, BAFU, BFE.
- [30] "Neue Inverkehrsetzungen von Strassenmotorfahrzeugen," 2018. Bundesamt für Statistik.

- [31] “The e-mobility revolution: impact of electric vehicles on the GB power system and emerging utility business models,” 2018. AURORA Energy Research.
- [32] “The future role and challenges of energy storage,” 2012. European Commission - Directorate-General for Energy.
- [33] J. Love, A. Smith, S. Watson, E. Oikonomou, A. Summerfield, C. Gleeson, P. Biddulph, L. Chiu, J. Wingfield, C. Martin, A. Stone, R. Lowe, “The addition of heat pump electricity load profiles to GB electricity demand: Evidence from a heat pump field trial,” *Applied Energy*, 2017.
- [34] A. Slaughter, “Electricity storage - technologies, impacts and prospects,” 2015. Deloitte Center for Energy Solutions.
- [35] T. Zufferey, A. Ulbig, S. Koch, G. Hug, “Unsupervised learning methods for power system data analysis,” *Big Data Application in Power Systems*, 2018.
- [36] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [37] R. Krause, “Lecture notes in introduction to machine learning,” 2018.
- [38] B. Schaule, “Disaggregation of smart meter data into specific load components,” 2018. ETH Zürich.
- [39] R. Bernhards, J. Morren , H. Slootweg, “Evaluating impact of new technologies on low voltage grids using probabilistic data enriched scenarios,” *2016 IEEE 16th International Conference on Environment and Electrical Engineering*, 2016.
- [40] R. Hyndman, “another look at the forecast-accuracy metrics for intermittent demand,” *Foresight*, 2006.

Appendices

.1 Variation of Demand during a Day

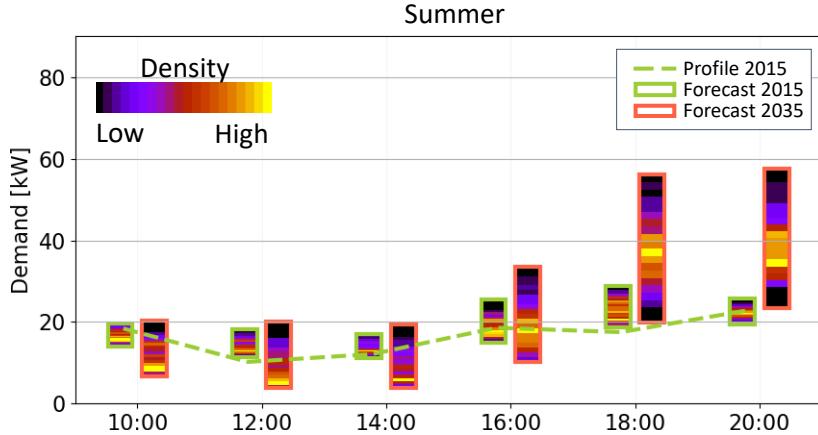


Figure 1: Electricity demand for different times during a working day in summer for node B. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.

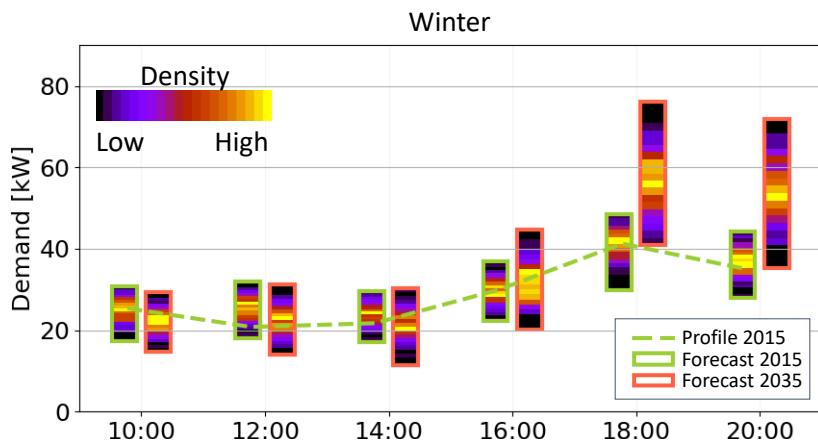


Figure 2: Electricity demand for different times during a working day in winter for node B. The green dotted line represents smart meter measurements for 2015, densities with a green frame represent the predicted demand in 2015 and densities with a red frame represent the predicted demand for 2035.

.2 Variation of Demand over the Years

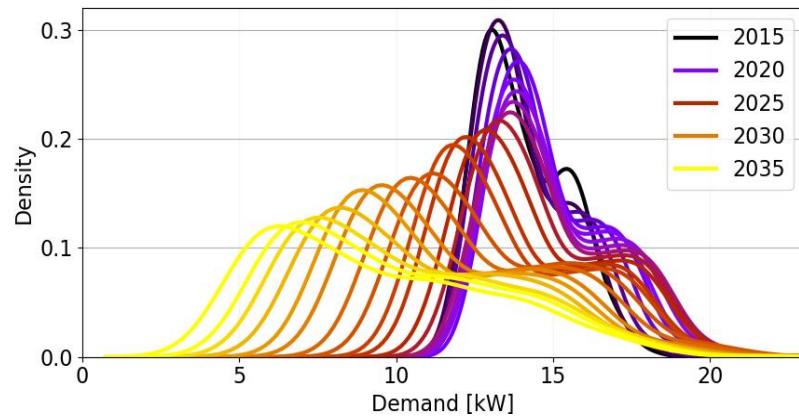


Figure 3: Evolution of the density on a working day in summer at 12:00 for node B.

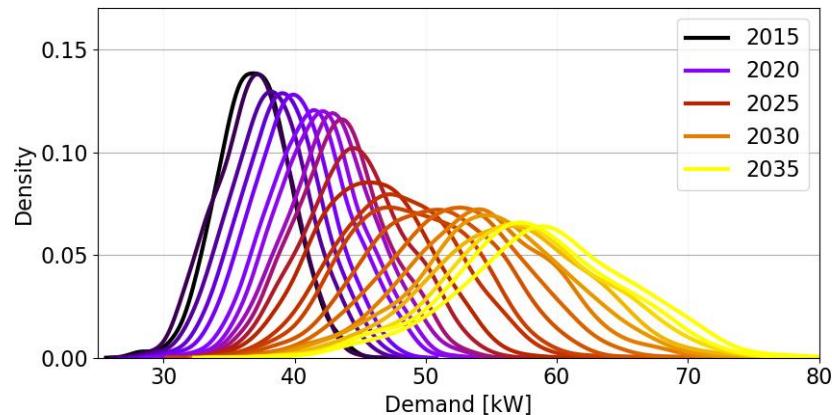


Figure 4: Evolution of the density on a working day in winter at 20:00 for node B.

.3 Development of Demand for different Scenarios

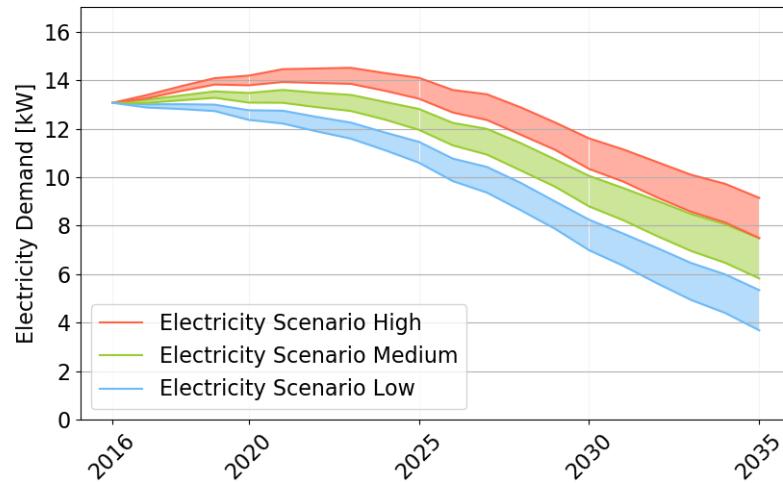


Figure 5: Demand for a working day in summer at 12:00 for node B. The lower bounds are given by the high PV production scenario and the upper bounds by the low PV production scenario.

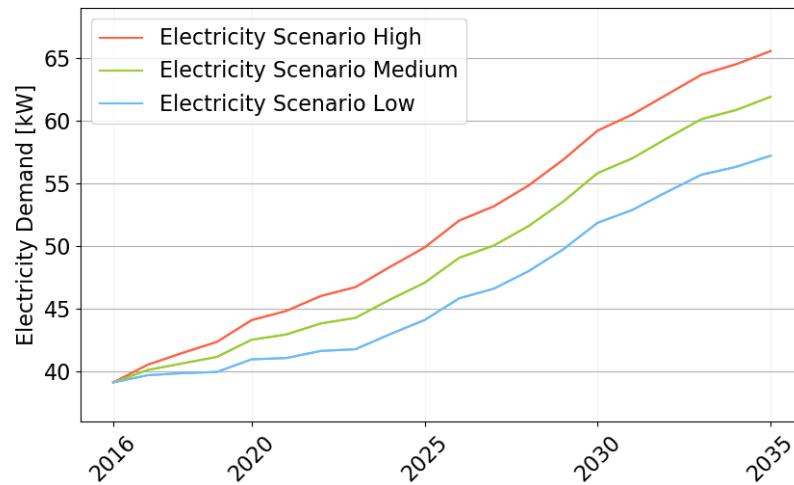


Figure 6: Demand for a working day in winter at 20:00 for node B.