technische universität
dortmund

Institute of Energy Systems, Energy
Efficiency and Energy Economics
Prof. Dr.-Ing. C. Rehtanz

Master's thesis

Forecasting the aggregated charging load of electric vehicles at different charging sites

Tim Unterluggauer

Acknowledgments

Abstract

The anticipated large-scale market ramp-up of electric vehicles in the future represents both an opportunity and a major challenge for the power grid. On the one hand, electric vehicles may support the power grid by offering their flexibility on the Control Reserve market. On the other hand, simultaneous charging could lead to severe grid bottlenecks in the low voltage network. In either case, the prediction of the aggregated charging load is of great importance. Within this work a new multivariate multi-step forecasting approach based on Artificial Neural Networks, namely the Long Short-Term Memory, is introduced. Using historical charging data, a prediction of the aggregated charging load in 15-min resolution, clustered according to the charging at home, the charging at work, the charging at public car parks and the charging at shopping centers, is conducted. For each charging site two forecast horizons, a 1-hour and a 1-day charging load forecast, are analyzed. The results indicate that no reliable forecast of the charging load can be accomplished with a forecast horizon of one day and that the prediction accuracy can be significantly increased if the period is shortened to one hour.

Kurzfassung

Die zukünftig zu erwartende starke Verbreitung von Elektrofahrzeugen stellt sowohl eine Chance als auch eine große Herausforderung für das existierende Stromnetz dar. Auf der einen Seite können Elektrofahrzeuge zur Stabilisierung des Stromnetzes beitragen, indem sie ihre Flexibilität auf dem Regelleistungsmarkt anbieten. Auf der anderen Seite könnte das gleichzeitige Laden auch zu gravierenden Netzengpässen im Niederspannungsnetz führen. In beiden Fällen ist die Vorhersage der aggregierten Ladelast von großer Bedeutung. Im Rahmen dieser Arbeit wird eine besondere Form der Künstlich Neuronalen Netzen verwendet, das Long Short-Term Memory, und ein neuwertiger multivariater Prognoseansatz entwickelt, welcher die Ladelast für die nächste Stunde oder den nächsten Tag prognostiziert. Basierend auf realen Ladedaten erfolgt die Prognose der aggregierten Ladelast in 15-minütiger Auflösung für verschiedene Ladeorte – dem Laden zu Hause, dem Laden am Arbeitsplatz, dem Laden auf öffentlichen Parkplätzen und dem Laden in Einkaufszentren. Die Ergebnisse zeigen, dass bei einem Prognosezeitraum von einem Tag keine zuverlässige Lastprognose erzielt werden kann, die Prognosegenauigkeit bei Verkürzung des Prognosezeitraums auf eine Stunde aber deutlich gesteigert wird.

Contents

List of Figures

List of Tables

Nomenclature

Formula Symbols

a	Number of inputs to neuron e
$agPow$	Aggregated power of PEVs available for the provision of Control Reserve
$b_{\tilde{c}}$	Bias of the Long Short-Term Memory candidate cell state
b_e	Bias of neuron e
b_f	Bias of the forget gate
b_i	Bias of the input gate
b_o	Bias of the output gate
$capP$	Capacity price of Control Reserve
$charging\_load_{new,k}$	Modified charging load for charging session k
$cosine\_type\_day_t$	Cosine encoded type of day indicator for timestep t
$cosine\_quarter\_hour_t$	Cosine encoded quarter hour indicator for timestep t
c_t	Cell state at time t
$\tilde{c}_t$	Candidate cell state at time t
c_{t-1}	Cell state at the previous timestep $t-1$
d	Number of aggregators and sub-aggregators
D	Number of daily forecasts
$dmax$	Indicator of the daily peak charging load
dS	Degree of PEV charging simultaneity
dt	Type of day under consideration
e	Indicator of single neuron e
$energy_{orig,k}$	Original amount of energy charged during charging session k
enP	Energy price of Control Reserve
f_t	Forget gate at time t
$f(z_e)$	Activation function for neuron e
h	Hidden state of the recurrent neural network
$hsig$	Hard sigmoid activation function
h_t	Hidden state at time t
h_{t-1}	Hidden state at previous timestep $t-1$

i_t	Input gate at time t
in	Amount of input timesteps for the LSTM
$int\_type\_day_t$	Integer encoded type of day indicator for timestep t
$int\_quarter\_hour_t$	Integer encoded quarter hour indicator for timestep t
k	Indicator of a single charging event/session
l	Total number of charging events/sessions
$load_{dmax}$	True peak load of the day
$\hat{load}_{dmax}$	Predicted peak load of the day
$load_{max}$	Maximal charging load value of the training data
$load_{min}$	Minimal charging load value of the training data
$load_t$	Actual charging load value at time t
$\hat{load}_t$	Predicted charging load value at time t
$load_{t,norm}$	Normalized charging load value at time t
$\hat{load}_{t,norm}$	Normalized predicted charging load value at time t
$L(load_{t,norm}, \hat{load}_{t,norm})$	Mean Squared Error loss function
m	Amount of timesteps in the charging load time series
MAE	Mean Absolute Error
$MAPE\_peak$	Mean Absolute Percentage Error of the peak forecast
mod	Modulo operation
n	Number of features
N	Number of predicted charging load values
$NMAE1$	Normalized Mean Absolute Error 1 (normalized by the mean charging load)
$NMAE2$	Normalized Mean Absolute Error 2 (normalized by the difference between maximum and minimum charging load)
$nslot$	Number of 15-min time slots
$num\_apart(1\ or\ 2)$	Number of new apartments in year $year$ (residential buildings with 1 or 2 apartments)
$num\_apart(3\ or\ more)$	Number of new apartments in year $year$ (residential buildings with 3 or more apartments)
o_t	Output gate at time t
out	Amount of output timesteps for the LSTM
$peak\_load\_dev$	Average deviation between real and predicted peak load
$pl_{residential,\ year}$	Amount of new residential parking lots in year $year$
$pl_{workplace,\ year}$	Amount of new workplace parking lots in year $year$
$power\text{-}to\text{-}zero\_ts_k$	End of charging timestamp of charging session k

q	Cardinality of input feature
$ReLU$	Rectified Linear Unit activation function
$revS1$	Revenue in scenario 1
$revS2$	Revenue in scenario 2
$revS3$	Revenue in scenario 3
rev_ts_k	Reference timestamp for charging session k
s	Type of charging site under consideration
$sine\_type\_day_t$	Sine encoded type of day indicator for timestep t
$sine\_quarter\_hour_t$	Sine encoded quarter hour indicator for timestep t
$slot$	15-min time slot index of the predicted and real peak load
$start\_interval_k$	Indicator of the 15-min interval at which charging starts for charging session k
$start\text{-}of\text{-}charging\_ts_k$	Start of charging timestamp of charging session k
$tanh$	Tanh activation function
tCR	Type of Control Reserve under consideration
$time\_dev$	Average time deviation between time of real and predicted peak load
$time_{new,k}$	Modified charging time of charging session k
$time_{orig,k}$	Original charging time of charging session k
tPr	Type of price used for the calculation
$Useful\_area_{year}$	Useful area of office buildings completed in year $year$
w	Weights of the neural network
$w_{a,e}$	Weight of input a of neuron e
$W_{\tilde{c},h}$	Weight matrix of the hidden state to build the candidate state
$W_{\tilde{c},x}$	Weight matrix of the current input to build the candidate state
$W_{f,h}$	Weight matrix of the hidden state at the forget gate
$W_{f,x}$	Weight matrix of the current input at the forget gate
$W_{i,h}$	Weight matrix of the hidden state at the input gate
$W_{i,x}$	Weight matrix of the current input at the input gate
$W_{o,h}$	Weight matrix of the hidden state at the output gate
$W_{o,x}$	Weight matrix of the current input at the output gate
x	Neural network input(s)
$x_{a,e}$	Input a of neuron e
x_t	Input at time t

y	Neural network output(s)
y_e	Output of the neuron e
y_t	Output at time t
z_e	Network input of neuron e
σ	Sigmoid activation function
$\odot$	Element-wise multiplication (Hadamard product)

Abbreviations

AAE   Average Absolute Error

AC   Alternating current

ANN   Artificial Neural Network

AR   Autoregressive

ARIMA   Autoregressive Integrated Moving Average

ARMA   Autoregressive Moving Average

av   average

BEV   Battery electric vehicle

Bi-LSTM   Bi-directional LSTM

BPNN   Back Propagation Neural Network

BPTT   Back Propagation Through Time

CEE   Commission internationale de réglementation en vue de l'approbation de l'équipement électrique

CNN   Convolutional Neural Network

CR   Control Reserve

CRPS   Continious Ranked Probability Score

CV   Coefficient of Variation

DC   Direct current

DER   Distributed Energy Resources

Destatis Statistisches Bundesamt

dev   Deviation

DIRMO   Combination of DirREC and MIMO

DirREC   Combination of Recursive and Direct Strategy

DNN   Deep Neural Network

EU   European Union

EV   Electric vehicle

FC   Fuzzy Clustering

GA   Genetic Algorithm

GBRT   Gradient Boosted Regression Tree

GRU   Gated Recurrent Unit

GTB   Gradient Tree Boosting

HEV   Hybrid electric vehicle

ICCB   In-Cable Control-Box

LSSVM   Least Squares Support Vector Machine

LSTM   Long Short-Term Memory

MA   Moving Average

MAE   Mean Absolute Error

MAPE   Mean Absolute Percentage Error

max   maximum

MCS   Monte Carlo Simulation

MIMO   Multi-Input Multi-Output

min   Minute

MPSF   Modified Pattern Sequence-based Forecast

MSE   Mean Squared Error

NILA	Lion Algorithm by Niche Immune
NMAE	Normalized Mean Absolute Error
NPE	Nationale Plattform Elektromobilität
NRMSE	Normalized Root Mean Squared Error
PCR	Primary Control Reserve
PEV	Plug-in electric vehicle
PHEV	Plug-in hybrid electric vehicle
PWM	Pulse-width modulation
ReLU	Rectified Linear Unit
RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
R-ANN	Rough Artificial Neural Networks
RR-ANN	Recurrent Rough Artificial Neural Networks
SAE	Stacked Auto-Encoders
SARIMA	Seasonal AutoRegressive Integrated Moving Average
SCR	Secondary Control Reserve
sec	Second
SMAPE	Symmetric Mean Absolute Percentage Error
SVM	Support Vector Machine
TCR	Tertiary Control Reserve
ts	timestamp

TWDP-NN   Time Weighted Dot Product Based Nearest Neighbor

V2G   Vehicle-to-Grid

WPA   Wolf Pack Algorithm

1. Introduction

Global warming, caused by extensive emission of greenhouse gases, is one of the greatest challenges of the 21st century [1]. In 2017, the transport sector was responsible for 27 % of all European greenhouse gas emissions, thus ranking as the second largest emitter of greenhouse gases in the European Union (EU) right behind the energy sector [2], [3]. At 44 % passenger cars attributed to a major part of the transport emissions [2].

The goal of the EU, as laid down in the European Green Deal, is to become climate neutral by 2050 [4]. Consequently, a reduction in emissions by 60 % by 2050 compared to the baseline year of 1990 is being targeted for the transport sector [2]. In order to meet the set targets electric vehicles (EVs) are seen as one of the solutions to reduce carbon emissions in the transport sector. Other related concerns, such as urban air pollution and its impact on health, have also encouraged politicians to promote the adoption of EVs as a replacement for conventional internal combustion engine powered vehicles [5]. Due to high costs, range anxiety and other technical factors such as a lack of public charging stations, the market ramp-up has been rather slow up to this day except in a few countries such as Norway [6]. However, due to strong political support, a substantial market penetration of EVs can be expected in the future.

1.1. Motivation of the thesis

This anticipated large-scale EV rollout represents both a considerable challenge and an opportunity for the existing power system. On the one hand, the large-scale diffusion of EVs is likely to lead to a significant additional load on the network, which could endanger system stability. Especially in low voltage networks, an increasing number of EVs could lead to grid bottlenecks, as the electricity grid is not designed to cope with this additional load. On the other hand, EVs could also benefit the power system by providing flexibility. Both the usage of EVs as energy storages or the implementation of smart charging strategies could offer this flexibility to the power system. Smart charging strategies could not only avoid or minimize the necessity

of grid expansions, but also help to safely integrate a large number of renewable energies sources into the distribution grid [5], [7].

In order to be able to identify potential network constraints, caused by the simultaneous charging of a large number of EVs, or to make use of the available flexibility the accurate forecast of the charging load of EVs is of utmost importance and could benefit different stakeholders. For network operators, the charging load forecast is relevant to identify bottlenecks in the low voltage network at an early stage to be able to derive countermeasures. Moreover, charging load forecasting is also of great importance for balancing group managers in Germany. A balancing group must be balanced every quarter of an hour and an accurate load forecast helps to avoid high balancing energy penalties. Last but not least, sub-aggregators that bundle the flexibility of a large number of EVs could use the charging load forecast to develop suitable smart charging strategies in order to commercialize the resulting flexibility in the energy market.

1.2. Objective and scope of the thesis

Given the previously discussed motivation, this thesis aims at developing an Artificial Neural Network (ANN) for predicting the aggregated charging load of EVs at two different forecast horizons. The analysis is performed for four different charging sites – the charging at home, the charging at work, the charging at car parks and the charging at shopping centers. The objective is to enable a sophisticated assessment of the different charging locations and forecast horizons in order to highlight potential similarities and differences and derive implications for the successful usage of the forecast by both sub-aggregators and network operators.

In the scope of this work only passenger cars are considered – the analysis of the charging behavior of buses or similar vehicles is not part of this research. Furthermore, this thesis focuses on the short-term load forecasting. Hence, medium and long-term forecasts are not covered within the scope of this work. Lastly, it should also be noted that this work does not aim at benchmarking different prediction methods. Rather, the focus is placed on the analysis of different charging locations and forecast horizons, given the large research gap in the recent literature in this field.

1.3. Structure of the thesis

The work is structured in the following eight chapters, illustrated in Figure 1.1. While **Chapter 1** deals with the relevance and objectives of the work, **Chapter 2** addresses the fundamentals of EVs, the role of a sub-aggregator, the term of load forecasting and the current state of research on EV charging load forecasting. **Chapter 3** provides a case study for potential revenues of an EV sub-aggregator when participating in the German Control Reserve Market. **Chapter 4** establishes the foundations necessary for the understanding of the work by covering the essential aspects of ANNs and subsequently focusing on Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, a particular type of RNNs applied in this work. In the following **Chapter 5**, the authors own approach for charging load forecasting is presented, whereby the data pre-processing, the LSTM training and the LSTM forecasting is discussed in more detail. **Chapter 6** presents the training results with initial hyperparameters, the hyperparameter tuning results and charging load forecast results for the various charging locations. Subsequently, an analysis and evaluation of the results is carried out in **Chapter 7**. Finally, the thesis concludes with a summary of the most important findings and an outlook for future research in **Chapter 8**.

Introduction *Chapter 1*	Motivation of the thesis	Objective & scope of the thesis	Structure of the thesis
State of the Art *Chapter 2*	Fundamentals of electric vehicles / The role of aggregators and sub-aggregators	Load forecasting	Related work
Sub-aggregation in Germany – a case study *Chapter 3*	Context and assumptions made	Control Reserve market in Germany	Potential revenues in Germany
Artificial Neural Networks *Chapter 4*	Fundamentals of Artificial Neural Networks	Recurrent Neural Networks	
Methodology: Multivariate multi-step LSTM forecasting *Chapter 5*	Methodology / Data pre-processing	LSTM implementation and training	LSTM forecasting
Results *Chapter 6*	Training results with initial hyperparameter	Hyperparameter tuning results	Forecast results
Discussion of results *Chapter 7*	Analysis of initial training results	Analysis of hyperparameter tuning results	Analysis of forecast results
Conclusions and future work *Chapter 8*	Summary of the work	Findings of the work	Future work

Figure 1.1: Structure of the thesis

2. State of the Art

Now that the motivation and objectives of the work have been outlined in the previous chapter, a number of fundamentals important for understanding the work are dealt with in this chapter. First, a number of key aspects of EVs are discussed with the aim of establishing some boundaries for this thesis. Next, the role of aggregators and sub-aggregator is addressed to understand the importance of load forecasting. Subsequently, the term of load forecasting is discussed in more detail. Last but not least, a comprehensive review of the existing literature deals with the current state of research on EV charging load forecasting in order to identify shortcomings in the existing literature and to highlight the contributions of this work.

2.1. Fundamentals of electric vehicles

Several years before Carl Benz filed a patent for the first gasoline-powered car in 1886, the EV was invented in 1834 [8]. However, due to sinking oil prices, range and cost advantages of combustion engine powered cars compared to EVs, the EV was slowly pushed out of the market [9]. In recent years environmental considerations and different societal, economic and technological factors have led to the revival of the EV [8]. This section covers the following important aspects of EVs – the different types of EVs, the charging technology, the charging sites and the market ramp-up of EVs in Finland and Germany.

2.1.1. Classification of electric vehicles

To start with, as of today EVs can be grouped into three different types of cars [10]:

- **Battery electric vehicle**s (**BEV**s) are solely powered from onboard electrical battery packs. The battery is charged by connecting the car to the power grid and furthermore can be supplied with energy recovered by recuperation.

- **Plug-in hybrid electric vehicle**s (**PHEV**s), on the other hand, have both a combustion engine and an electric motor. While everyday distances can often

be covered solely electrically, the conventional motor is used for longer dis-
tances. Similar to BEVs, the battery is charged when connected to the power
grid and by recuperation.

- **Hybrid electric vehicle**s (**HEV**s) on the contrary, solely consist of a combus-
  tion engine and an electric motor without a battery and can therefore not be
  charged by an external power supply. The electric motor is powered by recu-
  peration energy and supports the combustion engine to achieve a higher fuel
  economy.

Since the aim of this work is to predict the charging load of EVs, HEVs will not be
considered any further in this work. Unless explicitly stated, BEVs and PHEVs are
referred to as **Plug-in electric vehicle**s (**PEV**s) in the further course of this work.

2.1.2. Charging infrastructure

As previously mentioned, the battery of PEVs needs to be charged in order to be
able to power the electric motor. Figure 2.1 illustrates three different possibilities of
PEV charging, although nowadays only one of the three approaches is widely used
[11].
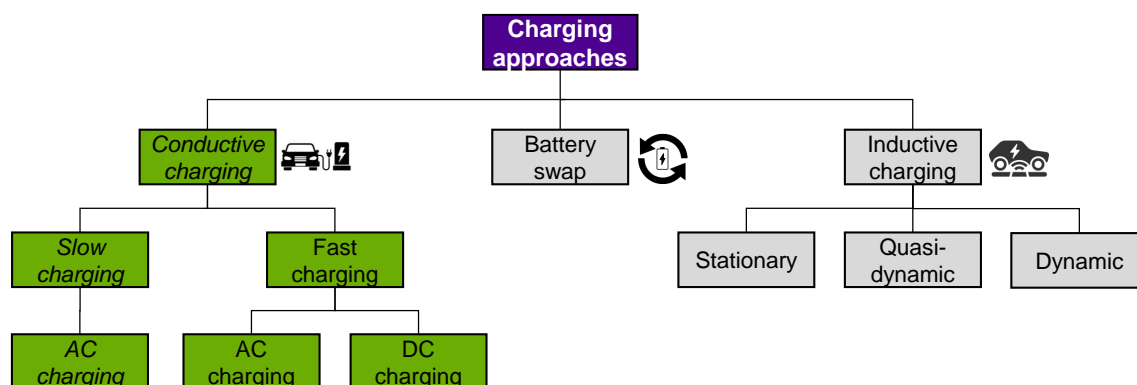
Figure 2.1: Different PEV charging solutions

Battery swapping refers to the automatic or manual replacement of a discharged
battery with a fully charged battery. The major advantage of this approach is its
quickness as the duration is comparable to an ordinary refueling process. However,
due to the high initial investment costs and lack of standardization in batteries, ve-

hicle interfaces and operation platforms, to date, China is one of the few countries with an existing battery swapping station infrastructure [11].

Inductive charging allows power to be transmitted by means of electromagnetic waves without direct contact between the PEV and charging station [11]. Three different approaches can be distinguished – stationary, quasi-dynamic and dynamic charging [12]. Stationary charging refers to inductive charging while the car is parked in a car park or garage. In quasi-dynamic charging the PEV is charged while temporarily stopped during a trip. For instance, buses can be recharged during their stopover at bus stops. Dynamic charging, on the contrary, takes place in-motion while driving the PEV. Therefore, the charging infrastructure needs to be incorporated into the roads. Although inductive charging can both increase comfort and reduce range anxiety, the disadvantages such as high investment costs, limited charging capacity and charging losses still outweigh the benefits at present, thereby hindering market diffusion [12].

Conductive charging remains as the only approach that is currently suitable for large-scale application and hence serves as the foundation for this work. Conductive charging denotes the recharging of the PEV with the help of a charging cable and can be distinguished either based on the charging capacity or the current type. While all charging processes with a charging capacity of up to 22 kW are classified as slow charging, charging events with a nominal charging power of more than 22 kW are referred to as fast charging [13]. Slow charging processes are based on alternating current (AC), whereas fast charging processes can be performed with both AC and direct current (DC). In the following slow charging and fast charging is discussed in more detail.

Slow charging

The ability to charge is one of the most important aspects for the adoption of PEVs. The predominant charging of PEVs takes place at locations where long idle times are achieved, such as overnight parking at home or parking at work. Hence, slow charging is the most commonly used charging possibility so far [14].

Table 2.1 illustrates four possible charging modes for charging the PEV. Mode 1 charging refers to charging the PEV battery with a single-phase AC power source, for instance by connecting the car to the power grid via the Schuko or CEE plug.

Since DC is required to charge the battery, a rectifier is installed in the PEV by default. Moreover, Mode 1 does not enable any communication between the PEV and power source. The maximum charging power is limited to 3.7 kW (16 A, 230 V). However, in order to avoid the overloading of possibly inadequately designed lines or installations, the charging capacity is usually limited to 2.3 kW (10 A, 230 V). Mode 2 enables both single-phase charging (3.7 kW) and three-phase charging with a nominal power of up to 22 kW (32 A, 230 V). An In-Cable Control-Box (ICCB) ensures safety and allows the user to adjust the maximum charging current [15].

Table 2.1: Different charging modes for PEV charging (Source: Own illustration based on [14])

	Slow Charging			Fast Charging
Charging Mode	*Mode 1*	*Mode 2*	*Mode 3*	*Mode 4*
Phase(s)	Single-phase	Single- and three-phase	Three-phase	No phases (DC)
Maximum charging power	3.7 kW (Schuko)	22 kW (three-phases)	43.5 kW	400 kW
Communication between	none	ICCB and onboard charger (PWM)	Charging station and onboard charger	Charging station and vehicle

Fast charging

The long charging time and limited range is a major obstacle hindering the market ramp-up of PEVs. A nationwide network of fast charging stations could reduce the range anxiety and increase the acceptance of PEVs by providing an emergency charging possibility [14]. In contrast to the previously discussed modes, Mode 3 and Mode 4 support fast charging and are mostly designed for public charging stations.

Similar to Mode 2, Mode 3 uses three-phase AC charging and the communication between PEV and charging station is based on pulse-width modulation (PWM). However, the maximum charging power in Mode 3 is significant higher and amounts to 43.5 kW (63 A, 400 V). Mode 4, on the other hand, denotes charging the PEV at a DC charging station. DC charging has the major benefit that rather than using the installed on-board charger, a stationary charger is employed in the charging station, which usually allows much higher charging capacities. Communication likewise works via PWM [15]. At present, most DC charging stations provide 50 kW of power [16]. Recently, however, an increasing number of charging stations with a charging

capacity of 350 kW have been built and it is expected that charging capacities of 400 kW will be realized in the near future [16], [17].

In this thesis solely charging sessions with slow charging events are analyzed. Hence, fast charging is not considered further in the course of this work.

2.1.3. Charging sites

The charging infrastructure can be grouped into three categories – private, semi-public and public. Table 2.2 illustrates the different charging infrastructure, typical charging locations for each category and the denotation of the clustered charging sites within this work.

Table 2.2: Overview of the different charging infrastructures and their typical charging locations (Source: Own illustration based on [18], [19])

Private			Semi-public		Public
Garage/ parking space at own property	Parking lots/ underground car park of dormitory/ multi-family house	Corporate parking lots on own premises	Car park/ highway service station	Shopping centres, car parks, customer parking	roadside, public parking lots
Residential charging		Workplace charging	Public charging	Shopping center charging	

Public charging infrastructure encompasses charging stations in public spaces and on the roadside. Semi-public charging infrastructure refers to charging stations in parking facilities, motorway service stations, petrol stations and supermarkets. Last but not least, the private charging infrastructure includes charging stations in corporate car parks and private parking spaces in privately owned homes or residential complexes [18]. In this work four different charging sites for PEVs are considered, the charging at home, the charging at company premises and the charging at multi-storey car parks and car parks of shopping centers. Thus, this work focuses on the private and semi-public charging infrastructures. In the following residential charging refers to charging at home, workplace charging denotes the charging at company premises, public charging refers to charging the car at car parks in the city and shop-

ping center charging indicates the PEV charging at car parks of a shopping center complex.

2.1.4. Market ramp-up of electric vehicles in Finland and Germany

Up to now, PEVs has lagged behind expectations and self-imposed market ramp-up targets in Germany and Finland. As mentioned before, the main reason for the slow market ramp-up is the lack of charging infrastructure, the range anxiety of customers and the high acquisition costs compared to conventional cars [6]. Table 2.3 illustrates the market ramp-up of PEV passenger cars and the stock growth of all passenger cars within the last ten years in Germany and Finland.

Table 2.3: Number of electric passenger cars and total passenger cars in Germany and Finland within the last ten years (Source: Own illustration based on [20], [21])

Date	BEV (Germany)	PHEV (Germany)	All passenger cars (Germany)	BEV (Finland)	PHEV (Finland)	All passenger cars (Finland)
01.01.2010	1,588	no data	41,738,000	17	6	2,758,291
01.01.2011	2,307	no data	42,300,000	34	19	2,858,244
01.01.2012	4,541	no data	42,928,000	70	20	2,958,568
01.01.2013	7,114	no data	43,431,000	127	145	3,036,618
01.01.2014	12,156	no data	43,851,000	188	253	3,105,834
01.01.2015	18,948	no data	44,403,000	386	455	3,172,735
01.01.2016	25,502	no data	45,071,000	657	889	3,234,860
01.01.2017	34,022	20,975	45,804,000	889	2,230	3,322,672
01.01.2018	53,861	44,419	46,475,000	1,538	5,303	3,398,937
01.01.2019	83,175	66,997	47,096,000	2,500	12,222	3,470,507
01.01.2020	136,617	102,175	47,715,977	4,830	23,003	3,549,803

Considering the number of 238,792 PEVs in Germany, measured at the beginning of 2020, it becomes clear that the target of one million PEVs by the year 2020, set by the Federal Government in 2009, was missed by a huge margin. In last year's progress report, written by the members of the "Nationale Plattform Elektromobilität" (NPE), it is expected that this figure will be reached in 2022 [17]. However, even this assessment remains questionable considering the monitored growth rate. By 2030, the number is expected to reach seven to ten million PEVs [22]. In order to achieve these ambitious goals, the German federal government recently announced an increase in subsidies for PEVs. BEVs and PHEVs with a net list price up to 40,000 € are subsidized with 6,000 € and 4,500 € respectively. If the net list price exceeds

40,000 €, the subsidy drops to 5,000 € and 3,750 € respectively [23].

In Finland, on the contrary, a target of 250,000 PEVs by the year 2030 was set in it's energy and climate strategy. Starting in 2018, the purchase of a BEV is subsidized by 2,000 € [24]. Looking at the current numbers of PEVs in Finland at the start of 2020, a negligibly small market share of around 0.8 % can be observed, slightly higher compared to the share of around 0.5 % in Germany. In contrast to Germany, however, it appears that PHEVs are more popular than BEVs and therefore account for a significantly higher proportion of the PEV stock. Since 2017, there has also been a significantly higher increase of PEVs compared to previous years. This pattern may play an important role later in the analysis of the results of the prediction model, as this trend can be seen in the aggregated charging load at the different charging sites over the course of the year and certainly have an influence on the model performance.

2.2. The role of aggregators and sub-aggregators

As shown in the previous section, the market ramp-up of PEVs slowly starts to take off and countries like Germany are investing a large amount of money to electrify the transport sector. As mentioned earlier the anticipated large-scale diffusion of PEVs poses both threats and new possibilities for the power system. In today's power system it is becoming increasingly difficult to maintain the necessary balance between power consumption and generation due to the shutdown of conventional power plants and the rapid growth of Renewable Energies. Thus, on the one hand, severe intermittent electrical loads caused by the simultaneous charging of a large number of PEVs could aggravate this problem and lead to security risks for the power grid [25]. On the other hand, Distributed Energy Resources (DER), such as PEVs, could help to balance consumption and weather dependent power generation such as wind or solar power by providing ancillary services and benefit the power system [26].

However, a major share of the so far unused and future increasing DER flexibility potential will be distributed widely throughout the power system. Additionally, the individual capacity of a single DER is too small to be traded on the Control Reserve

(CR) market. As a result, new market roles emerge, namely the aggregator and sub-aggregator [26]. In Figure 2.2 four possible flexible resources for aggregation (PEV fleets, energy storages, smart heating/ cooling and smart factories) and a hypothetical hierarchy between aggregators and sub-aggregators are illustrated.
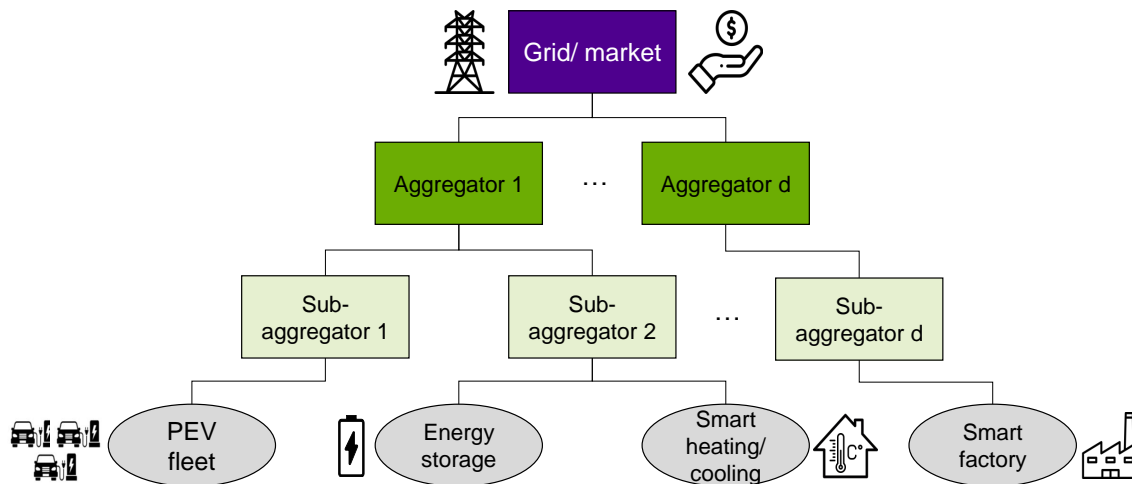
Figure 2.2: Visualization of the market roles of the aggregator and sub-aggregator (Source: Own illustration based on [26] and [27])

The sub-aggregator serves as a link between DER owner and aggregator and therefore occupies an intermediate position in the value chain for marketing the DER flexibility to the CR market. The sub-aggregator can be understood as an entity with the ability to monitor and control DERs and to predict the amount of flexibility for every tradable timeslot in the market. As for PEVs, the charging point operator might be able to take on the role of sub-aggregator, as he has the possibility to carry out smart charging without major adaptations. The aggregator, in contrast, combines the flexibility capacity of multiple sub-aggregators and participates in the market. With this distinction made, it becomes apparent that the expertise and tools needed for both the aggregator and sub-aggregator differ widely. However, in some cases one entity might occupy both roles at the same time [26].

2.3. Load forecasting

To enable the sub-aggregator to market possible flexibility potentials of PEVs, the charging load forecast of PEVs plays a decisive role. Load forecasting refers to the process of predicting future load demands and can be categorized into very short-term (minutes and hours in advance), short-term (one day to weeks in advance), medium-term (months to a year beforehand) and long-term (years ahead) forecasting [28].

Moreover, load forecasting can be seen as a time series problem. A time series is a sequence of values measured over time and arranged chronologically [29]. Time series forecasting can be categorized according to the input and output of the model. With regard to the input data, univariate and multivariate models can be distinguished. Depending on the output, single-step and multi-step models may be differentiated. While univariate approaches make predictions solely based on univariate data such as historical load values, multivariate models use additional time series data like weather or categorical data for achieving the forecasting task. Single-step forecasting refers to models predicting ahead in time for only one timestep, whereas multi-step forecasting performs predictions up to a certain time prediction horizon [30].

Multi-step ahead time series forecasting can be furthermore classified in the following five different methods [31]:

1. A Recursive Strategy,
2. a Direct Strategy,
3. a combination of both the Recursive and Direct Strategy (DirREC),
4. a Multi-Input Multi-Output Strategy (MIMO) and
5. a combination of DirREC and MIMO Strategy (DIRMO).

The **Recursive Strategy** is based on a 1-step ahead prediction and can be seen as the oldest and most intuitive multi-step prediction method. A model is trained for a single-step ahead forecast and the predicted value is then used as part of the inputs in order to forecast the next timestep based on the same model. This process is repeatedly carried out until the desired prediction horizon is generated. The major

drawback of this method can be seen in the possibility of accumulating prediction errors which can lead to quickly degrading forecast accuracy with increasing prediction time horizon. The **Direct Strategy** tries to eliminate the aforementioned problem by developing a distinct model for each prediction timestep. The disadvantages of this method include the higher computational complexity as well as the difficulty of modeling dependencies. The **DirREC Strategy** combines the benefits of both previously discussed approaches. Separate models are constructed for each timestep and the predicted value of the previous timestep is added as an additional input parameter.

While the first three methods can be seen as single output strategies, the **MIMO strategy** follows the approach of developing a model that is able to forecast the entire prediction sequence at once. The MIMO Strategy eliminates both the stochastic independence assumption of the Direct Strategy as well as the accumulation of errors that affect the Recursive Strategy. However, due to the fixed prediction horizon the MIMO Strategy might lack flexibility. This drawback was the motive for the development of a second multiple-output strategy called **DIRMO**. DIRMO aims at maintaining the most desirable aspects of both the DirREC and MIMO Strategy. The forecast is split into different blocks which perform the prediction task based on the MIMO Strategy [31].

2.4. Related Work

Load forecasting plays a vital role in the management of today's power systems and aims at securing the network security of electrical power systems [28]. In contrast to the extensively researched field of load forecasting in power systems, very few publications can be found in the literature that address the issue of charging load forecasting for PEVs. However, with the rising demand for PEVs in recent years, research on that topic has become more popular.

In the following, the most important studies of the last four years will be reviewed. The objective of the short meta-analysis is to reflect the current state of research in order to subsequently summarize the limitations of the considered studies and to emphasize the relevance and contribution of this work. The literature review in this section follows the classification in [29], which categorizes the several forecasting

approaches into linear, non-linear, rule-based, wavelet transform, other and ensemble models. However, only those models which have been applied to PEV charging load forecasting are taken into account. Hence, only linear, non-linear, ensemble and other models are considered.

2.4.1. Charging load forecasting based on linear methods

Linear forecasting approaches use linear functions to model time series behavior. Popular methods include Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) processes. Even though linear models are still applied, they are usually inferior to other models in terms of accuracy, which is why they are often used as baseline models [29]. With regard to the prediction of the charging load of PEVs, ARIMA models are particularly popular among all linear approaches. Three ARIMA models can be found in the literature and are discussed in more detail below.

Seasonal Autoregressive Integrated Moving Average models
To start with, Louie [32] proposes various Seasonal Autoregressive Integrated Moving Average (SARIMAs) models to forecast the aggregated PEV charging station load. Different prediction time frames are considered, including a 1-, 2-, and 24-hour short-term forecasts. The modeling is done by using the aggregated charging load of more than 2,400 public and private charging stations in California, covering a 2-year time span from January 2011 to January 2013. The load forecast is carried out separately for weekdays and weekends as well as San Diego and Washington State. The accuracy is compared to a persistence forecast and a Modified Pattern Sequence-based Forecast (MPSF) algorithm and evaluated on the basis of the Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). It can be demonstrated, that the SARIMA models achieve superior results compared to the two benchmark models in all cases.

Autoregressive Integrated Moving Average models
A plain ARIMA model is introduced in [33] to predict the PEV charging demand of the next 24 hours and assess the impact of PEVs on the distribution network. Compared to [32], the prediction is not based on real charging data but on expected

driving distances and charging times, derived from daily driving patterns in the USA. The paper compares the accuracy of the proposed decoupled forecaster, which independently predicts the charging demand and conventional electrical load, with an integrated forecaster, predicting the overall charging load of the PEV parking lot. The authors conclude that the decoupled forecaster is able to significantly reduce the error by providing more precise demand predictions.

One more study carried out in [34] examines the applicability of an ARIMA model to forecast the uncontrolled day-ahead charging demand of PEVs in 15-minute (min) resolution. Moreover, a storage model for aggregation is introduced to assess the impact of aggregation on the prediction accuracy. The analysis is conducted based on charging data obtained from more than 1,341 non-residential charging stations in California over the full year of 2013. Using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Variation (CV) for analysis the findings hint that aggregation has a positive impact on the forecast results. While absolute errors such as RMSE and MAE rise with increased aggregation, relative errors show the opposite effect and decrease with larger aggregation. Hence, the authors conclude that the prediction error can be significantly lowered with larger aggregation.

2.4.2. Charging load forecasting based on non-linear methods

In contrast to linear approaches, non-linear models rely on non-linear functions to model more complex time series behavior. Non-linear models include ANNs, Genetic Algorithms (GAs) and Support Vector Machines (SVMs) [29]. Looking at the literature, it becomes clear that considerable attention has been paid to ANNs lately. A variety of studies can be found that implement different types of ANNs, which are discussed in more detail below.

Artificial Neural Networks
Jahangir et al. [35] follow a similar approach as in [33] and perform a 24-hour charging load prediction based on historical driving patterns. Contrary to [33], however, three distinct ANNs are implemented to determine the arrival and departure time and correlated trip lengths. Using historical data about travel behavior of conventional cars in the USA first the arrival and departure times are independently forecasted.

Afterwards the length of the trips are estimated in correlation to the previously fore-
casted arrival and departure time to calculate the charging load subsequently. Three
different ANNs are deployed and compared to the results of a Monte Carlo Simu-
lation (MCS). It can be shown that the use of ANNs, Rough Artificial Neural Net-
works (R-ANNs) or Recurrent Rough Artificial Neural Networks (RR-ANNs) leads to
a higher load prediction accuracy compared to the MCS. RR-ANNs hereby gener-
ate the most accurate forecasts, followed by R-ANNs. The paper also addresses
the importance of predicting the load for potential aggregators by implementing an
optimal charging algorithm that aims at minimizing the costs of the aggregator while
considering the network constraints. It can be shown that precise prediction algo-
rithms have a positive impact on the aggregators financial gains.

Zhu et al. [36] suggest a special type of RNN, the LSTM, for single-step PEV charg-
ing load prediction and demonstrate the advantages of their model by comparing
the prediction performance at two different timescales with the results of a simple
ANN. The prediction is based on historical charging data from charging stations in
Shenzhen, covering the period from July 2017 to July 2018. The analysis, based on
RMSE and MAE, reveals two different findings. First, the LSTM model outperforms
the ANN at both time scales. Second, when comparing the prediction errors of the
different time scales, 15-min or 30-min resolution of the charging load, it can be seen
that with decreasing time intervals the prediction error diminishes as well.

Another recently published article by Zhu et al. [37] studies the super short-term
charging load predictability of six different types of ANNs and deep learning meth-
ods. The performance of an ANN, a RNN, a canonical LSTM, a Bi-directional LSTM
(Bi-LSTM), a Gated Recurrent Unit (GRU) network and a Stacked Auto-Encoders
(SAEs) is evaluated on the basis of the RMSE, MAPE and goodness of fit. All mod-
els are trained, evaluated and tested on two different historical datasets from Shen-
zhen, containing charging event data with a minute resolution spanning from June
2017 to July 2018. While the first dataset contains charging data from 24 charg-
ing stations including the charging of electric buses the second dataset belongs to
an aggregator for commercial building chargers. The results demonstrate that the
tested models are suitable for super short-term charging load forecasting. How-
ever, the analysis also finds evidence that the prediction accuracy of the commercial
charging site turns out to be lower than that of the public charging site. Similar to
[36], superior results are also seen for the LSTM model taking into account all three

evaluation metrics.

Finally, [38] provides a further evaluation of different ANN approaches for short-term charging load forecasting. Once again historical charging load data from Shenzhen, measured between April 2017 and June 2018, is used and applied to assess the performance of a Deep Neural Network (DNN), RNN, LSTM and GRU model. The charging load on an hourly basis, the charging time, the real-time electricity prices and holiday marks form the inputs for the charging load forecast. For performance evaluation the Normalized Root Mean Squared Error (NRMSE) and Normalized Mean Absolute Error (NMAE) are applied. Contradictory to [37], the results indicate that GRUs with one hidden layer achieve the most accurate forecast results for the hourly predictions over one day. For all four implemented methods the NRSME does not exceed 4 % and the NMAE never reaches values above 1.5 %. Another main finding is that increasing the number of hidden layers not only deteriorates the accuracy, but also slows down the training process. All four models also struggle to accurately predict peak loads.

2.4.3. Charging load forecasting based on ensemble models

Ensemble models can be understood as a compound of different models with the aim to increase the prediction accuracy by linking individual methods in a meaningful way [29]. The literature research shows that numerous studies follow this approach. In the following, the different hybrid approaches are presented in more detail.

To start with, in a recent paper by Gerossier et al. [39] a bottom-up approach for the short-term forecast of the aggregated PEV fleet load is presented. The dataset used in this paper includes minutely resolved power consumption measurements of 46 privately owned PEVs, measured in 2015 in Austin (Texas). For the day-ahead prediction first the individual PEV power consumption is forecasted with a probabilistic random forest model. In a next step the aggregated load for the PEV fleet is estimated using a deterministic bottom-up approach. Subsequently, the authors use two benchmark models in order to evaluate the performance of the bottom-up model – a persistence and Gradient Tree Boosting model (GTB). It becomes apparent that the developed bottom-up model is superior to the persistence model both in

terms of the MAE and Continious Ranked Probability Score (CRPS). However, the GTB achieves similar results even though the bottom-up approach is slightly more efficient.

Paper [40] follows the approach of forecasting the charging load with the help of SVMs. A hybrid model combining Least Squares Support Vector Machine (LSSVM), Fuzzy Clustering (FC) and Wolf Pack Algorithm (WPA) is proposed and the author aims at forecasting the load for electric bus charging stations. Various influencing factors are taken into account, such as the type of day, the minimum and maximum temperatures, weather conditions and the charging load at the same time in the last three days. The proposed model is trained and tested on two different datasets and compared to the outcomes of a WPA-LSSVM, a regular LSSVM and a Back Propagation Neural Network (BPNN). Both datasets include real load data of an electric bus charging station in Baoding and the necessary meteorological data. It can be shown that in both case studies the FC-WPA-LSSVM model outperforms all three other models in terms of RMSE, MAPE and Average Absolute Error (AAE). However, the author mentions that at certain times FC-WPA-LSSVM leads to less accurate results.

Li et al. [41] implement a hybrid approach, combining a Lion Algorithm by Niche Immune (NILA) and a Convolutional Neural Network (CNN), for short-term PEV charging stations load prediction. With the help of the NILA the optimal weights and thresholds of the CNN are identified in a first step. Taking into account the input parameters seasonal category, minimum and maximum temperature, type of day as well as the charging load at the same time during the former five days the CNN is trained subsequently. In the final analysis the performance of the implemented model is compared to the outcome with a Lion Algorithm CNN, a single CNN and SVM. The NILA-CNN performs superior in all categories considering the RMSE, the MAPE as well as the AAE.

Lastly, an approved Random Forest (RF) algorithm is applied to predict the PEV charging load of single and grouped PEV charging stations in [42]. The study is carried out with historical charging data of Shenzhen (China), measured in the years 2016 to 2018. The authors consider a variety of input factors influencing the load forecast – such as previous day's charging load, year, month, week, day, latitude and longitude, charging capacity indicator and an activity indicator representing the

volatility of the charging load. Evaluating the importance of each input for the fore-
cast accuracy the authors conclude that the input factors previous day's charging
load, activity indicator, week and day are most important. The capacitor indicator
and year, on the contrary, are negligible. Using MAPE and RMSE for error analy-
sis the authors highlight that the proposed RF approach can effectively predict the
charging load on a 15-min basis, both for single and grouped charging stations. This
analysis also reveals that in contrast to the aggregated load forecast, more precise
results are obtained when performing individual forecasts for the various charging
stations.

2.4.4.   Charging load forecasting with different models

In literature two publications can be found which do not link the different models in
one model but implement them individually to compare their prediction accuracy. To
conclude, these two studies are now discussed in more detail below.

Buzna et al. [43] compare different machine learning and time series models to
forecast the daily PEV charging load. To conduct their study, they resort to a large
amount of charging data from 1,700 charging stations. The historical data originates
from charging events in the Netherlands between January 2012 and March 2016.
For time series forecasting two different SARIMAX models are chosen. Whereas the
single SARIMAX model does not distinguish weekdays, the alternative SARIMAX
models predict the charging load for weekdays and weekend separately. Moreover,
a RF and a Gradient Boosted Regression Tree (GBRT) model are implemented,
representing the machine learning algorithms. A 7-day, 14-day and 28-day forecast
is carried out and the accuracy is evaluated using MAPE. For model evaluation a
persistence method serves as a benchmark. The findings reveal the superior per-
formance of all proposed methods compared to the benchmark model. Furthermore,
the authors also note that RF outperforms the GBRT in all scenarios. Comparing the
time series models, the single SARIMAX model indicates a superior performance for
the 7-day forecast whereas the alternative SARIMAX model provides a higher level
of accuracy for 14- and 28-day forecast horizons. The authors conclude that overall
the time series models outperform the machine learning algorithms.

Last but not least, Majidpour et al. [44] apply four different algorithms, namely RF, SVR, Time Weighted Dot Product Based Nearest Neighbor (TWDP-NN) and MPSF, for 24-hour charging load forecasting. The evaluation is based on two different datasets – charging and station measurements recorded between December 2011 to February 2014 from charging stations placed on the UCLA campus in California. While the charging event data includes information about the start and end time of the charging process as well as the charged energy, the station measurement consists of voltage, current, power factor and charging length data. Both speed and accuracy are evaluated for both datasets in order to compare both type of datasets for the feasibility of charging load forecasting. Using the Symmetric Mean Absolute Percentage Error (SMAPE) for evaluation, the authors find that there is no statistical significant difference among the prediction errors. However, the findings indicate that the most accurate predictions based on the charging measurement can be achieved using the TWDP-NN while the MPSF provides the most precise results for the station measurement dataset.

2.4.5. Shortcomings of research and contributions of this work

Table 2.4 summarizes the main findings of the meta-analysis. For each of the presented studies, the proposed model, the baseline model and the model that produces the superior results are displayed. Furthermore, the metrics used for the evaluation are illustrated and key characteristics of the applied dataset and the prediction method are outlined. Finally, the table also provides details on the origin of the data and the time period during which the data was measured. Overall, the meta-analysis of the 13 studies reveals several limitations that will be addressed in this paper.

First of all, most of the papers either use data from the USA [33], [44], [32], [34], [35], [39] or China [40], [41], [42], [36], [37], [38]. Only [43] covers the European market by using historical data from the Netherlands. This paper aims at filling the gap in research on charging load forecasting in the European context by using historical charging load data from Finland.

Additionally, a key limitation of the previous studies is that several datasets con-

tain charging data collected in times with little PEV penetration in the market which may limit the significance of the results. While in [33] and [35] no charging data is used at all, the data in [44], [32], [34], [39] and [43] all originate from years until 2016, characterized by a negligible market share of PEVs. These concerns regarding the validity and applicability of the results to today's situation will be eliminated in this thesis by the use of up-to-date charging data covering the year 2019 and the start of the year 2020.

Table 2.4: Literature review on existing PEV charging load forecasting techniques

Author (Year) [Ref.]	Proposed model(s)	Baseline model(s)	Superior model	Metrics	Characteristics of data and forecast methodology	Data origin [year]
Amini et al. (2016) [33]	ARIMA	–	Decoupled forecast	MAE, MAPE	National household travel survey, 24-hour (single-step) forecast, 1-hour resolution	USA
Majidpour et al. (2016) [44]	SVR, RF, MPSF, TWDP-NN	–	TWDP-NN, MPSF	SMAPE	Charging & station records, 24-hour (single-step) forecast, 1-hour resolution	UCLA (USA) [Dec. 2011 - Feb. 2014]
Louie (2017) [32]	SARIMA	persistence, MPSF	SARIMA	MSE, MAPE	Charging load data (aggregated), 1-, 2-, 24-hour forecast 1-hour resolution	California [Jan. 2011 - Jan. 2013]
Zhang (2018) [40]	FC-WPA-LSSVM	BPNN, LSSVM, WPA-LSSVM	FC-WPA-LSSVM	AAE, MAPE, RMSE	Electric bus charging, multivariate inputs, 24-hour (single-step) forecast, 1-hour resolution	Baoding [2017]
Li et al. (2018) [41]	NILA-CNN	SVM, CNN, Lion CNN	NILA-CNN	AAE, MAPE, RMSE	Public charging (aggregated), multivariate inputs, 24-hour (single-step) forecast, 30-min resolution	Beijing (China) [Jun. 2017 - Nov. 2017]
Lu et al. (2018) [42]	RF	–	–	MAPE, RMSE	Public charging (single & aggr.), multivariate inputs, single-step prediction, 15-min resolution	Shenzhen (China) [2016 - 2018]
Pertl et al. (2019) [34]	ARIMA	–	–	CV, MAE, RMSE	Non-residential charging (aggr.), 24-hour forecast, 15-min resolution	California [2013]
Jahangir et al. (2019) [35]	ANN, R-ANN, RR-ANN	MCS	RR-ANN	R2, MAE, MAPE, RMSE	National household travel survey, 24-hour forecast, 1-hour resolution	USA [2017]
Zhu et al. (2019) [36]	LSTM	ANN	LSTM	MAE, RMSE	Public charging (aggregated), 24-hour (single-step) forecast, 15-/30-min resolution	Shenzhen [Jul. 2017 - Jul. 2018]
Zhu et al. (2019) [37]	ANN, RNN, GRU, SAE, LSTM, Bi-LSTM	–	LSTM	MAPE, RMSE,	Public and commercial charging (aggregated) 1-, 5-, 15-min forecast, 1-min resolution,	Shenzhen (China) [Jun. 2017 - Jul. 2018]
Zhu et al. (2019) [38]	RNN, DNN, GRU, LSTM	–	GRU	NMAE, NRSME	Public charging (aggregated), multivariate inputs, 24-hour (single-step) forecast, 1-hour resolution	Shenzhen (China) [Apr. 2017 - Jun. 2018]
Gerossier et al. (2019) [39]	bottom-up RF	persistence, GTB	bottom-up RF	MAE, CRPS	Residential charging (aggregated), multivariate inputs, 24-hour prediction	Austin (USA) [2015]
Buzna et al. (2019) [43]	SARIMAX, RF, GBRT	persistence model	SARIMAX	MAPE	Charging data (aggregated) 7-, 14-, 28-day forecast, daily resolution	Netherlands [Jan. 2012 - Mar. 2016]

Furthermore, although [37] has indicated that the achievable prediction accuracy may be dependent on the type of charging site, no attention has been given to the analysis and comparison of different charging sites. As stated in [45] and [46] the charging behavior and thus the uncertainty which may have an influence on the forecast accuracy differs between residential, public and commercial charging locations. The shortage of research in this area will be addressed in this work by analyzing the charging load predictability of the before mentioned different charging sites.

Lastly, it can be seen that most studies only perform a single-step prediction [33], [44], [40], [42] or a prediction of the load values with a very coarse resolution [33], [44], [40], [43]. Moreover, all reviewed ANN approaches are limited to single-step predictions [41], [36], [37], [38]. This limitation diminishes the applicability of the presented methods for an aggregator in practice, requiring a short-term load prediction for an extended time period in high resolution. This gap is addressed within this work by developing a short-term forecast with both a 1-hour and 1-day prediction horizon in 15-min resolution.

3. Sub-aggregation in Germany - a case study

Before proceeding to the next chapter dealing with the basics of ANNs, this chapter presents a scenario assessment for potential revenues of a PEV sub-aggregator when participating in the Secondary Control Reserve (SCR) or Tertiary Control Reserve (TCR) market in Germany. The chapter is structured as follows: First, the framework of the scenario analysis is defined and the assumptions made are outlined. Subsequently, the key elements of the CR market in Germany are discussed. Finally, the potential revenues of a sub-aggregator participating in the CR market in Germany are calculated.

3.1. Context and assumptions made

The analysis is carried out from the point of view of the Finnish company Parking Energy Ltd, which provided a majority of the data used in this work. Parking Energy Ltd, is part of the project Smart Otaniemi, which currently investigates possible sub-aggregator concepts for the future in Finland [27]. In Germany the role of the sub-aggregator is also not yet defined. At present, only one research project, called New 4.0 – Norddeutsche Energiewende, is focusing on the development of a sub-aggregator model in Germany, demonstrating that the requirements for sub-aggregators in Germany remain unclear [47].

However, with it's core competence of providing cost-effective cabling systems for the real estate industry to enable PEV charging at all parking bays and charging stations as a service, Parking Energy Ltd might be well equipped to perform the role of a PEV sub-aggregator in the future. It is important to emphasize that the following calculation can only be understood as a rough estimate, due to the lack of data and the numerous assumptions that had to be made, as the multitude of influencing factors could not be modeled. The analysis in this chapter is based on the following key assumptions:

1. All costs arising from the provision of the CR are not considered.
2. All prequalification conditions for participation in the CR market are fulfilled.

3. The PEV owner approves the measures taken for providing the CR.
4. All relevant parking lots will be equipped with the cabling system.
5. Interdependencies between workplace and residential charging are not considered.
6. During the period under consideration, Parking Energy Ltd is contracted to maintain and provide the CR.
7. The revenues generated are not divided between the aggregator and the sub-aggregator but are fully allocated to the sub-aggregator.
8. The goals of the German federal government regarding the market ramp-up of PEVs will be met.

3.2. Control Reserve market in Germany

CR is used to compensate for imbalances between generation and consumption in order to keep the electricity grid frequency within a small range around its nominal frequency of 50 Hz and to correct regional discrepancies in the balances from their benchmark value [48]. Depending on the type of imbalance, either positive or negative CR is needed to stabilize the grid. In the event of a drop in grid frequency, positive CR is required by increasing the power consumption or reducing the network load. Contrarily, negative CR is required if the network frequency exceeds a certain limit [49].

PEVs are capable of providing positive CR by means of a regulated reduction of the charging load or by feeding the energy stored in the battery back into the grid. The concept of providing positive CR by feeding energy back into the grid is called Vehicle-to-Grid (V2G). V2G, however, contains numerous uncertainties regarding the technical implementation, economic feasibility and user acceptance [50]. The provision of negative CR by increasing the charging load also proves to be challenging, implying that the PEV could only use a certain proportion of the maximum charging load throughout the entire charging process. Therefore, only the provision of positive CR through the controlled reduction of the PEV charging load is considered in the following, which imposes the least constraints on the user. As shown in Figure 3.1, three different types of CR must be distinguished with regard to their temporal availability – Primary Control Reserve (PCR), SCR and TCR [48].
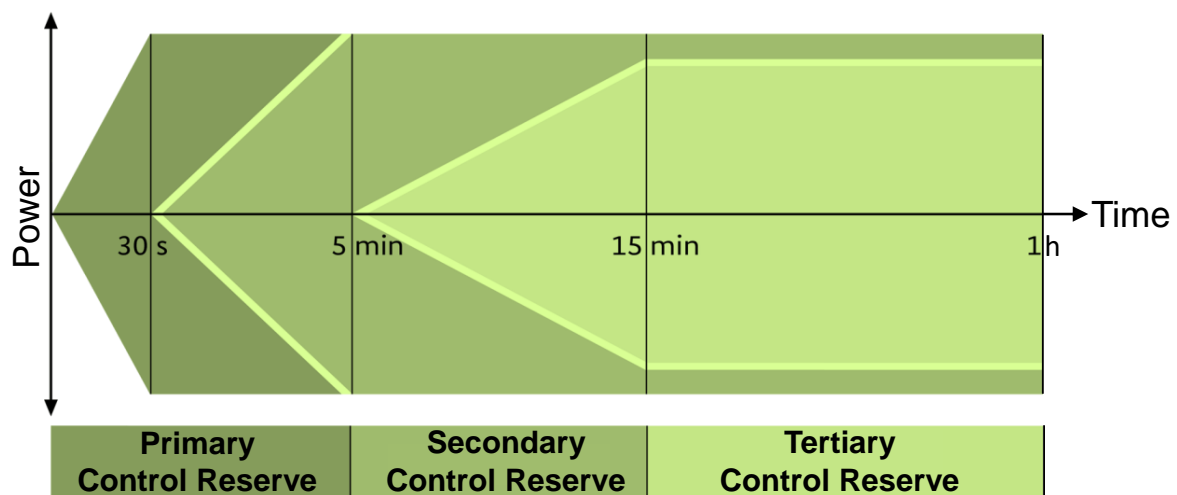
Figure 3.1: Types of Control Reserve with regard to time availability (Source: Own illustration based on [51])

PCR is automatically activated throughout the entire grid control cooperation within a time limit of 30 seconds. Due to the high demands on reaction speed and the symmetrical tendering of positive and negative CR, no further consideration is given to the use of PEVs for the provision of PCR [52]. PEVs may be more suited to participate in the SCR or TCR market for a variety of reasons. First, the response times of SCR and TCR are easier to accomplish. The contractually agreed SCR must be fully provided within five minutes of being automatically requested by one of the four transmission system operators in Germany. The TCR is also requested within one of the four balancing zones in Germany and replaces the SCR after 15 min. Furthermore, the tendering procedure is carried out separately for negative and positive CR, which makes it possible to only offer positive SCR and TCR. In addition, the minimum bidding volume for SCR and TCR was reduced to 1 MW in 2018 in order to facilitate market access for aggregated DERs. The tendering procedure has also been modified, allowing six time blocks of four hours each to be traded every day [53], [54]. Nevertheless, the high prequalification requirements, such as proof of service availability throughout the entire contract period, continue to pose a major challenge for aggregators and sub-aggregators to gain access to the CR market [52].

3.3. Potential revenues for a sub-aggregator in Germany

Having introduced the particularities of the CR market in Germany, an exemplary calculation of potential revenues for a PEV sub-aggregator on the SCR and TCR market in Germany is presented below.

3.3.1. Number of relevant parking lots

To start with, the number of relevant parking lots that can be equipped with the Parking Energy Ltd cabling system must be determined in order to evaluate the potential for PEV aggregation in Germany. Looking at the various parking sites in the context of flexibility and load shifting potential, it becomes clear that there is a much higher flexibility potential for PEV charging at the workplace and at home than in public places such as supermarkets due to the significantly longer idle time [55]. In addition, no data on the number of parking spaces in multi-storey car parks could be found, prompting the focus in the following to be on workplace and residential charging. Furthermore, the analysis is limited to new office and administration buildings and residential buildings built after 1993 due to two reasons. On the one hand, compared to residential buildings, office premises have a significantly shorter service life, which is why there is a trend towards replacing old buildings with new ones and thus making an investment in the cabling system unlikely [56]. On the other hand, recordings for new buildings across Germany began in 1993, which is why 1993 is chosen as the starting date for the analysis of both office and residential buildings.

Table 3.1 portrays the number of constructed office and residential buildings for the years 1993 to 2018, extracted from the database of the "Statistisches Bundesamt" (Destatis) [57], [58]. However, the number of new buildings doesn't indicate the number of corresponding parking bays. Rather, the number of parking lots is estimated based on the usable area (office building) or number of apartments (residential buildings) along with the guideline values for construction projects. The regulation of the construction of parking lots is part of the German "Bauordnungsrecht" and therefore, falls within the legislative competence of the federal states. Since the regulation is handled differently by the federal states, it is difficult to provide a generally accurate statement. Generally, three different procedures can be distinguished [59]:

1. Transfer of the decision-making authority to the municipalities.
2. Specification in the state building regulations.
3. No regulation (Berlin).

While Berlin is the only federal state without regulations, the regulations for office buildings usually range from 30 m² to 40 m² of useful area for one required parking bay [60], [61], [62], [63], [64]. For single-family houses, 1 - 2 parking spaces per apartment are required and residential buildings with more apartments usually need 0.6 - 2 parking lots per apartment. However, some regulations on parking lots for residential buildings are also based on the useful area. For the estimation of the available parking spaces of the relevant buildings which could be equipped with a charging station, the most conservative values are used. For office buildings, one parking space per 40 m² is assumed. For residential buildings, one (1 - 2 apartments) or 0.6 (more than two apartments) parking spaces per apartment are considered. The number of new workplace parking lots for each year is calculated according to Equation 3.1 and the number of new residential parking lots is based on Equation 3.2.

$$pl_{workplace,\ year} = \frac{Useful\_area_{year}\ in\ m^2}{40\ m^2} \tag{3.1}$$

$$pl_{residential,\ year} = num\_apart\ (1\ or\ 2) + 0.6 \times num\_apart\ (3\ or\ more) \tag{3.2}$$

The resulting number of office building parking lots can be seen in column four and residential parking lots in column eight of Table 3.1.

3.3.2. Aggregation potential and Control Reserve market prices

In order to deduce the number of PEVs from the number of parking spaces with the possibility of charging the PEV, the PEV market ramp-up goals of Germany are used as a benchmark [22]. By 2030, seven to ten million PEVs are envisaged on German roads, which in view of the roughly 47 million passenger cars in 2020 in Germany will account for around 15 % to 20 %. Given the still weak market ramp-up of PEVs in Germany, a market penetration of 15 % is assumed for further analysis, resulting in a fleet of about 365,074 PEVs for charging at work and about 885,933 PEVs for

Table 3.1: Summary of the calculated number of workplace and residential parking lots, the useful area and the number of apartments used for the calculation (Own calculation based on [57] - [64]

Con-struc-tion year	Number of office buildings	Useful area in 1,000 m²	Estimated number of parking lots	Number of residential buildings	Number of apartments (1 or 2)	Number of apartments (3 or more)	Estimated number of parking lots
1993	3,743	6,357	158,925	165,828	164,044	221,555	296,977
1994	3,859	6,310	157,750	212,363	212,354	284,309	382,939
1995	3,883	6,896	172,400	207,958	205,165	312,481	392,654
1996	3,477	5,371	134,275	191,577	188,802	292,173	364,106
1997	3,446	5,620	140,500	212,466	211,056	285,586	382,408
1998	3,227	4,874	121,850	215,832	220,611	208,400	345,651
1999	3,398	4,553	113,825	229,014	237,331	167,314	337,719
2000	3,515	4,576	114,400	220,797	229,715	136,445	311,582
2001	3,172	5,069	126,725	177,769	185,372	99,631	245,151
2002	2,897	5,076	126,900	164,838	172,874	79,728	220,711
2003	2,454	4,225	105,625	158,192	165,162	70,354	207,374
2004	2,046	3,421	85,525	170,400	177,204	69,386	218,836
2005	1,827	2,869	71,725	145,604	151,456	61,518	188,367
2006	1,869	3,206	80,150	146,303	150,069	69,616	191,839
2007	1,784	2,312	57,800	120,239	124,040	59,859	159,955
2008	1,859	2,455	61,375	94,415	96,369	54,615	129,138
2009	1,742	2,654	66,350	82,595	83,898	51,463	114,776
2010	1,533	2,227	55,675	84,340	85,367	53,014	117,175
2011	1,643	2,146	53,650	96,549	97,015	61,217	133,745
2012	1,674	2,028	50,700	100,816	100,294	71,041	142,919
2013	1,799	2,512	62,800	103,331	102,246	78,910	149,592
2014	1,714	2,226	55,650	108,908	106,846	101,021	167,459
2015	1,679	2,569	64,225	105,568	102,713	105,095	165,770
2016	1,618	2,513	62,825	109,990	106,301	115,150	175,391
2017	1,769	2,532	63,300	110,051	105,948	122,841	179,653
2018	1,715	2,756	68,900	107,581	103,363	134,954	184,335
Total	**63,342**	**97,352,000**	**2,433,825**	**3,843,324**	**3,885,615**	**3,367,676**	**5,906,222**

charging at home, after multiplying the number of parking lots with the degree of penetration of 15 %.

In a last step, possible revenue scenarios are identified based on the previously determined maximum number of PEVs to be aggregated. To this end, the capacity and energy prices for SCR and TCR are required. However, prices are subject to strong annual, monthly and daily fluctuations and even within a day, the price is highly dependent on the timeslot under consideration. In addition, the capacity and energy price are remunerated according to the pay-as-bid principle. In other words, the price for the provision and call up of the CR is based on the own offer [53], [54].

Within this work it is not possible to cover all dependencies, which is why the calculations are carried out with prices from 06.04.20 to 10.10.20, shown in Table 3.2. For workplace charging the timeslot 8:00 a.m. to 12:00 p.m. (POS_08_12) is chosen and for residential charging the timeslot from 8:00 p.m to 12.00 a.m. (POS_20_24). This selection is based on the assumption that at these times the largest possible number of PEVs are charging simultaneously. Furthermore, since it is not possible to simulate the own bid, the average prices as well as the marginal prices for both the capacity and energy price of the SCR and TCR are taken into account to show two possible scenarios. The demand for SCR and TCR depending on the day and timeslots are also displayed in order to compare them with the aggregated power of the PEVs later on.

Table 3.2: Capacity prices, energy prices and need for Secondary and Tertiary Control Reserve during the week of 06.04.20 to 10.04.20 (Source: Own illustration based on [65])

Day and Date	Type of CR and timeslot	Average capacity price in €/MW (SCR/TCR)	Marginal capacity price in €/MW (SCR/TCR)	Average energy price in €/MWh (SCR/TCR)	Marginal energy price in €/MWh (SCR/TCR)	Demand in MW (SCR/TCR)
Mon.,	POS_08_12	4.13/2.96	7.44/17.06	526.03/1,405.58	9,949/9,999	2,149/1,380
06.04.20	POS_20_24	7.98/2.96	15.33/12.00	590.47/1,148.75	9,949/9,999	2,132/1,311
Tue.,	POS_08_12	1.41/4.64	6.67/11.00	4,147.76/851.76	9,949/9,999	2,153/1,404
07.04.20	POS_20_24	9.12/2.09	14.84/5.36	588.51/549.90	9,949/9,999	2,145/1.364
Wed.,	POS_08_12	3.38/6.76	7.44/9.50	644.80/1,180.84	9,949/9,999	2,153/1,351
08.04.20	POS_20_24	9.23/2.69	24.14/4.50	1,061.30/1,095.73	9,982.90/9,999	2,141/1,340
Thu.,	POS_08_12	4.47/6.22	9.00/8.88	584.60/1,149.51	9,949/9,999	2,133/1,379
09.04.20	POS_20_24	8.45/2.79	14.15/7.14	673.25/1,498.89	9,982.90/9,999	2,142/1,368
Fri.,	POS_08_12	2.36/2.70	4.00/6.90	1,043.51/1,176.41	9,949/9,999	2,114/1,356
10.04.20	POS_20_24	6.77/2.35	15.09/8.00	1,206.02/1,679.51	9,949/9,999	2,120/1,251

3.3.3.   PEV sub-aggregation scenarios

Based on the prices, different scenarios can be developed to demonstrate possible revenue potentials. First, the aggregated power that can be offered on the CR market must be determined. Assuming that not all vehicles of the PEV fleet are charging at the same time, two scenarios are set up where 50 % or 75 % of the PEVs charge simultaneously during the selected timeslots and all charging processes can be reduced by 1 kW in case of a CR request. The resulting capacities amount to 182.54 MW (50 %) and 273.81 MW (75 %) for workplace charging and 442.97 MW (50 %) and 664.45 MW (75 %) for residential charging. Considering the tendered available power for SCR and TCR shown in Table 3.2, a large part of the required capac-

ity could be covered by aggregating the PEVs. In the case of 75-% simultaneous charging, around 7.9 % of the required SCR and 20 % of the TCR from 8:00 a.m. to 12:00 p.m. and as many as 31 % of the SCR and 49 % of the TCR from 8:00 p.m. to 12:00 a.m. could be met.

Table 3.3: Results of the different SCR and TCR revenue scenarios for workplace and residential charging over the period of 06.04.20 - 10.04.20 based on either the average or marginal price and different degrees of simultaneity (Source: Own calculation)

Type of charging site	Share of PEVs charging at the same time	Number of PEVs/ Aggregated power in MW	Type of available positive Control Reserve	Scenario 1: Average and maximal revenues in € for capacity only	Scenario 2: Average and maximal revenues in € for capacity and 15-min call of Control Reserve	Scenario 3: Average and maximal revenues in € for capacity and 1-h call of Control Reserve
Work-place	50 %	182,537/ 182.54	SCR	2,875.00 (av) 6,306.76 (max)	319,887.66 (av) 2,276,419.83 (max)	1,270,925.62 (av) 9,086,759.06 (max)
			TCR	4,249.53 (av) 9,736.68 (max)	267,294.23 (av) 2,291,258.51 (max)	1,056,428.35 (av) 9,135,823.98 (max)
	75 %	273,805/ 273.81	SCR	4,312.51 (av) 9,460.14 (max)	479,831.49 (av) 3,414,629.75 (max)	1,906,388.43 (av) 13,630,138.59 (max)
			TCR	6,374.30 (av) 14,605.03 (max)	400,941.35 (av) 3,436,887.76 (max)	1,584,642.52 (av) 13,703,735.98 (max)
Resi-dential	50 %	442,967/ 442.97	SCR	7,584.54 (av) 15,251.22 (max)	195,580.20 (av) 2,288,458.34 (max)	759,567.19 (av) 9,108,079.73 (max)
			TCR	2,351.12 (av) 6,753.98 (max)	274,918.93 (av) 2,288,275.80 (max)	1,092,622.38 (av) 9,132,841.28 (max)
	75 %	664,450/ 664.45	SCR	11,376.39 (av) 22,875.99 (max)	293,359.59 (av) 3,432,562.15 (max)	1,139,309.18 (av) 13,661,620.63 (max)
			TCR	3,526.54 (av) 10,130.60 (max)	412,363.34 (av) 3,432,288.35 (max)	1,638,873.71 (av) 13,698,761.60 (max)

Table 3.3 displays the average (av) and maximum (max) revenue potentials for three different scenarios, broken down by charging site, percentage of simultaneity and type of CR. In the first case, only the capacity price accounts for the revenue because the CR is not requested. In the second and third scenario, however, the CR is called up for 15 minutes and one hour respectively and for this time is compensated with the energy price. All revenues of the three scenarios are calculated according to the following Equations:

$$revS1_{s,\,tCR,\,dS,\,tPr} = agPow_{s,\,dS} \times \sum_{dt=1}^{5} capP_{dt,\,tCR,\,tPr} \qquad (3.3)$$

$$revS2_{s,\,tCR,\,dS,\,tPr} = agPow_{s,\,dS} \times \left(\sum_{dt=1}^{5} capP_{dt,\,tCR,\,tPr} + \frac{1}{4}h \times \sum_{dt=1}^{5} enP_{dt,\,tCR,\,tPr} \right) \quad (3.4)$$

$$revS3_{s,\,tCR,\,dS,\,tPr} = agPow_{s,\,dS} \times (\sum_{dt=1}^{5} capP_{dt,\,tCR,\,tPr} + 1h \times \sum_{dt=1}^{5} enP_{dt,\,tCR,\,tPr}) \quad (3.5)$$

While $revS1$, $revS2$, $revS3$ represent the revenues in the respective scenarios, $agPow$ indicates the aggregated power of the PEVs and $capP$ and enP correspond to the capacity and energy price. Indices dt, s and tCR denote the type of day, charging site and type of CR under consideration and dS indicates the degree of simultaneity while tPr specifies the type of price (average or marginal) used for the calculation.

When looking at the Scenario 1 results of workplace charging, the provision of TCR seems to be more lucrative than SCR for both degrees of simultaneity and price assumptions. Maximum revenues of 14,605.03 € can be earned. In comparison, maximum earnings of 13,703,735.98 € can be generated in the event that CR is demanded in Scenario 3. When comparing the revenues from SCR and TCR, SCR seems more profitable if reimbursed at the average price. Conversely, if CR is offered at the marginal price, slightly higher revenues can be realized by offering TCR. Contrary to the results for workplace charging, the provision of SCR in Scenario 1 of residential charging yields higher revenues than that of TCR. The maximum revenues amount to 22,875.99 € . In Scenarios 2 and 3, the duration of delivery and amount of compensation determines which type of CR is economically preferable. Assuming compensation at the average price, the provision of TCR is more lucrative in both scenarios for all cases. If, on the other hand, the marginal price is paid, SCR in Scenario 2 and TCR in Scenario 3 offer slight benefits. Maximum revenues of 13,698,761.60 € are possible under the assumptions made.

Overall, it becomes evident that the call-up of control energy significantly increases profits. However, the results should only be seen as an exemplary calculation, as the strongly fluctuating prices are likely to result in significant variations when looking at other weeks.

4. Artificial Neural Networks

As seen in Section 2.4 there are a number of different approaches available that can be used to forecast the charging load of PEVs. In this paper, the ANN approach is chosen for the prediction of the charging load, due to mainly two reasons. Firstly, ANN have shown superior results in non-linear settings, such as complex time series and should therefore present a valid choice for the forecasting of the sophisticated charging load time series used in this paper [66] . Moreover, deep learning with ANN has been a widely studied topic in research in recent years which is also reflected in the literature review in Section 2.4. However, as already shown, the current state of research has considerable limitations due to the lack of data and the complexity of the subject area. Therefore, this thesis tries to make a substantial contribution to the field of ANNs by using unique historical datasets as well as a new methodology which exceeds the previous prognosis horizon of the existing literature.

This chapter first introduces the basic principles of ANNs and deals with the two main types of ANNs – Feedforward Neural Networks and RNN. Subsequently, the RNN is discussed in more detail. In the end, the rationale for choosing the LSTM network and its basic architecture is presented.

4.1. Fundamentals of Artificial Neural Networks

ANNs are computational models inspired by the human nervous system and brain. Dating back to the first mathematical modeling of a neuron by McCulloch and Pitts in 1943, ANNs have been widely employed in several fields of science and engineering [67]. For short-term load forecasting, ANNs have been successfully implemented since the early 1990s [68]. In the following, the functionality and most relevant characteristics and classifications of ANN are described in more detail.

4.1.1. Structure of an artificial neuron

As mentioned before, ANN were developed to understand the biological processes of the human nervous system. They consist of a large amount of simple artificial neurons that send each other information using directed connections [69]. There are three different types of neurons: input, hidden, and output units. Input units are neurons that can receive input signals from the outside world. Hidden neurons are located between input and output units and contain an internal representation of the outside world. Output units are neurons that deliver signals to the outside world.

Figure 4.1 illustrates the schematic layout of a single neuron e. Each artificial neuron consists of a inputs and corresponding a weights, an optional bias, a propagation and activation function and one output [70].
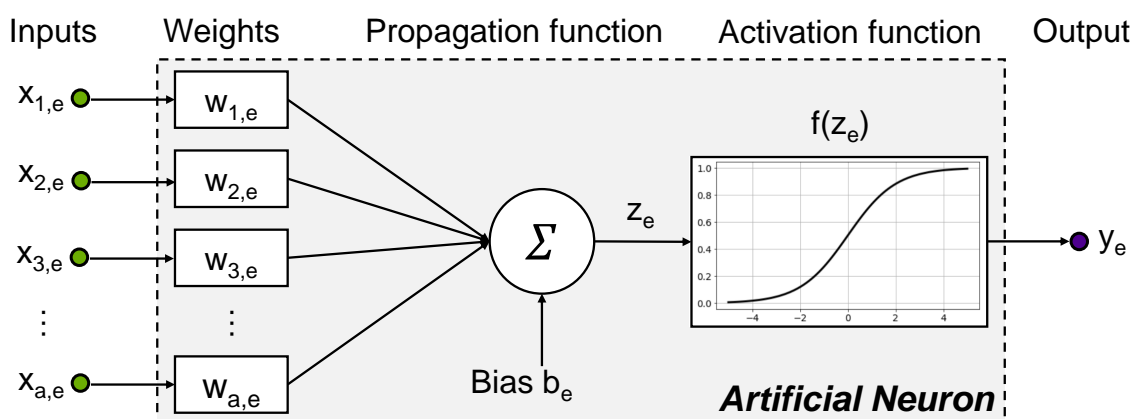
Figure 4.1: Structure of an artificial neuron (Own illustration based on [70])

Neurons can be understood as simple processors that calculate a new output from the receiving inputs. The inputs $x_{1,e}$, $x_{2,e}$, ... to $x_{a,e}$ each posses an unique weighting $w_{1,e}$, $w_{2,e}$, ... to $w_{a,e}$. Depending on their value, the weights exert a different effect. While weights with a positive value have a stimulating effect, connection weights with a negative value inhibit the signal to be communicated. A weight of zero indicates that the neuron currently has no influence on neuron e [69]. Since a neuron only processes a 1-dimensional input signal, the propagation function calculates the network input z_e of neuron e from the multitude of incoming signals. In Figure 4.1 the most frequently used propagation function, the weighted sum of the output values of the predecessor neurons, is illustrated [71].

The network input z_e of neuron e can be calculated according to the formula:

$$z_e = \sum_{i=1}^{a} x_{a,e} \times w_{a,e} + b_e \qquad (4.1)$$

The bias node b_e is optional and enables the output of a neuron to be shifted in order to achieve the optimal fit for the given data. The activation function $f(z_e)$ depicts the relationship between the network input and the activity level of a neuron and determines whether a neuron is active or inactive. If the network input remains below a certain threshold value, the neuron is not activated, it is only activated when it exceeds the previously specified threshold value. Activation functions introduce non-linear characteristics to the network and have a major affect on the model performance [72], [73], [74]. The output y_e of the neuron e is finally calculated by the output function using its activation state. However, in most cases the output signal is equal to the neuron's activation level, which is why the output function in Figure 4.1 is not explicitly portrayed [69].

4.1.2. Overall Artificial Neural Network architectures

Generally, grid layouts can be classified as either feedforward or recurrent architectures. Both variants usually consist of a single input and output layer and a variable number of intermediate hidden layers. A layer is hereby defined as a set of parallel artificial neurons of variable number.

Figure 4.2 depicts both a Feedforward Neural Network and RNN with one hidden layer. As can be seen, the Feedforward Neural Network does not possess recurrent connections and is therefore not able to memorize previous output values or activation states of its neurons [71]. It is assumed that all inputs and outputs are independent from each other, which results in limitations regarding the detection of temporal dependencies. RNNs, on the contrary, possess additional feedback links. Hence, they are specialized in the analysis of sequential/temporal data and have been used successfully in different applications with temporal dependencies [75]. The charging load of PEVs, similar to the electrical load in energy systems, is typically subject to strong time dependencies as the charging load profile generally corresponds to cyclical and seasonal patterns which depend on human activities such as the driv-

ing behavior [75]. Therefore, the RNN architecture is chosen for further proceedings and will be discussed in more detail later on.
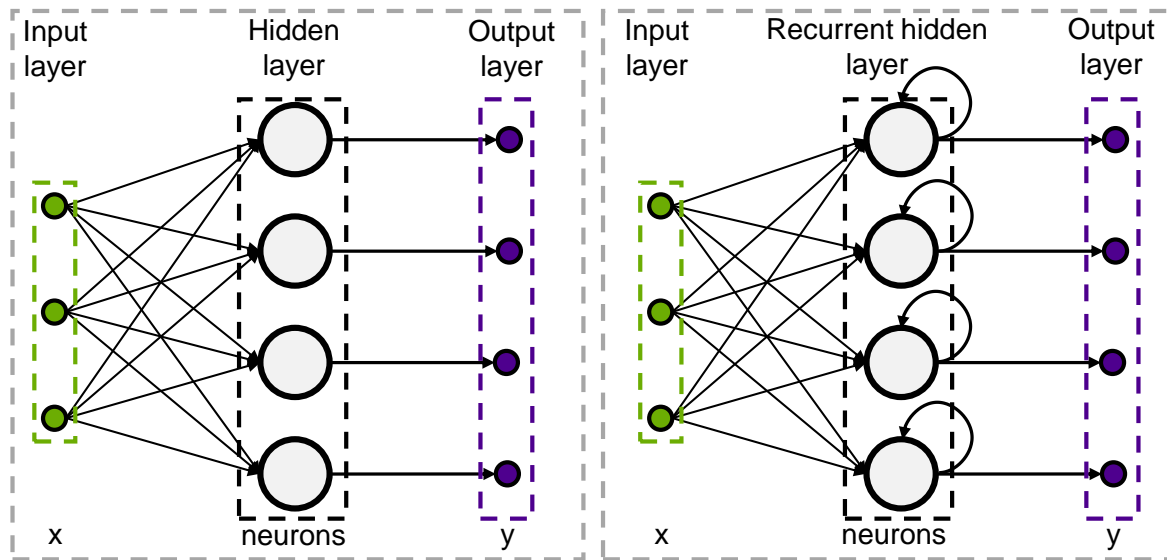
Figure 4.2: Exemplary architecture of a Feedforward (left) and Recurrent (right) Neural Network (Source: Own illustration based on [76])

4.1.3. Artificial Neural Network learning process

One of the key advantages of ANNs is their ability to learn. In theory, the training can involve a number of different approaches, such as [71]:

1. The development of new connections,
2. the removal of existing connections,
3. the modification of the threshold value of neurons,
4. the modification of the propagation or activation function,
5. the creation of new neurons,
6. the deletion of neurons,
7. and the modification of the weights.

While the first three possibilities can be understood and implemented as a modification of the weights, no automated algorithms exist for possibility four in practice, which is why little attention is paid to this particular training method. Approaches five

and six are more important in practice, but these techniques are mostly used only when satisfactory results have not been achieved by optimizing the weights. The central aspect of learning is thus the modification of weights, which is why in the further course of this work learning is understood as the modification of the weights in order to achieve the optimal performance of the ANN after training [71].

The training of the ANN can be primarily classified into three categories - supervised, unsupervised and reinforcement learning. In supervised learning, both the inputs and the desired outputs are provided to the ANN. Based on the inputs, the outputs are predicted and subsequently compared to the desired outputs to calculate the error of the network according to a specific loss function. Based on the loss function, the chosen optimizer updates the weights to reduce the error. Supervised learning usually leads to the fastest training results. In unsupervised training, however, only the input data is supplied to the ANN without the respective target values. Based on the given inputs the ANN subsequently modifies itself by learning about internal features of the presented data. Last but not least, in reinforcement learning inputs are provided to the network along with a given task. Depending on the calculated outputs, the connection weights are either increased (good performance) or decreased (bad performance) [77]. In this work, supervised learning is applied to update the weights during training.

4.2. Recurrent Neural Networks

Now that the basics of ANNs are understood, a more detailed description of RNNs is given below. First of all, the basic features of the training of RNNS are presented. Afterwards, the selection of the LSTM for this work is discussed before its main structure is outlined.

4.2.1. Recurrent Neural Network training

The most popular approach in supervised learning for Feedforward Neural Networks is the so-called backpropagation training algorithm. As the name backpropagation indicates, the weights are modified backwards from output to input layer. First, the

inputs are propagated through the network in order to obtain the output. The predicted outputs are then compared to the actual output and the loss functions is used to determine the corresponding error. Hereafter, the "gradient method" is used to compute the derivatives of the error in relation to each network weight. The derivatives are used to update the weights with the aim to minimize the error. This process is repeated until a certain termination criterion is met [71].

Due to the existence of feedback connections in RNNs, the error at each timestep is dependent on the previous timestep which is why the initial backpropagation algorithm cannot be used to calculate derivatives and modify the network parameter according to the network error. In fact, in order to determine a direct link between the loss function and the network weights, the RNN must be displayed as a directed graph [75]. Therefore, a modified backpropagation variant is used, known as Backpropagation Through Time (BPTT). BPTT operates by unrolling all input timesteps as shown in Figure 4.3. Each timestep consist of the current input x, the previous hidden state h and the current output y. For each timestep the error can be computed and accumulated. Subsequently, the network is restored to its original form and the weights are adjusted. Again, this process is repeated until a termination criterion is fulfilled. However, the RNN can suffer from the vanishing or exploding gradient effect [78]. If the gradient becomes to small or big during backpropagation, the resulting update of the weights will be to small or big as well, hindering the network to deal with long term dependencies [79]. Hence, in this work a RNN variant is required, which does not suffer from the vanishing or exploding gradient problem.
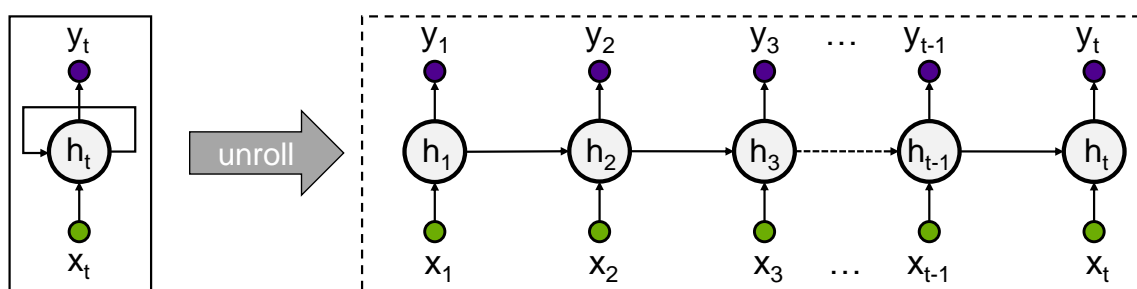
Figure 4.3: Unrolling of a Recurrent Neural Network over time (Source: Own illustration based on [78])

4.2.2. Model selection

There are two different types of RNN that tackle the vanishing and exploding gradient problem in the BPTT learning algorithm – LSTMs and GRUs. The LSTM is one of the most popular approaches of deep learning for load forecasting [80]. LSTM networks have the ability to learn both long-term dependencies as well as short-term features of the temporal data [81]. Numerous studies have shown that LSTM networks can achieve superior results in different domains of electricity load forecasting [80], [81], [82] as well as charging load forecasting as shown in the literature review in Section 2.4 [37], [36], [78].

In contrast to the LSTM, the GRU network represents a relatively novel alternative to the LSTM network and was first introduced in 2014 by Cho et al. [83]. It attempts to maintain the advantages of LSTM while reducing complexity. Several studies indicate that the GRU accelerates training time and can achieve similar and sometimes superior results in terms of load prediction accuracy [84], [85], [86], [87], [88].

However, the reduced complexity of the GRU might jeopardize the accuracy of the charging load prediction. Therefore, a novel LSTM approach is applied in this thesis. Nonetheless, it should be noted that the results of this work can also be applied to GRU networks for future research as the implementation follows the same design.

4.2.3. Long Short-Term Memory

The LSTM was developed in 1997 by Hochreiter and Schmidhuber [89] and can be seen as an effective tool for time series analysis and prediction [90]. The hidden layers of LSTM consist of three different gate units – namely forget gate f_t, input gate i_t and output gate o_t – and self-connected memory cells which allow the LSTM to learn long-term dependencies when dealing with sequential data [78]. Figure 4.4 offers a thorough insight into the inner architecture of a LSTM cell at a single timestep.
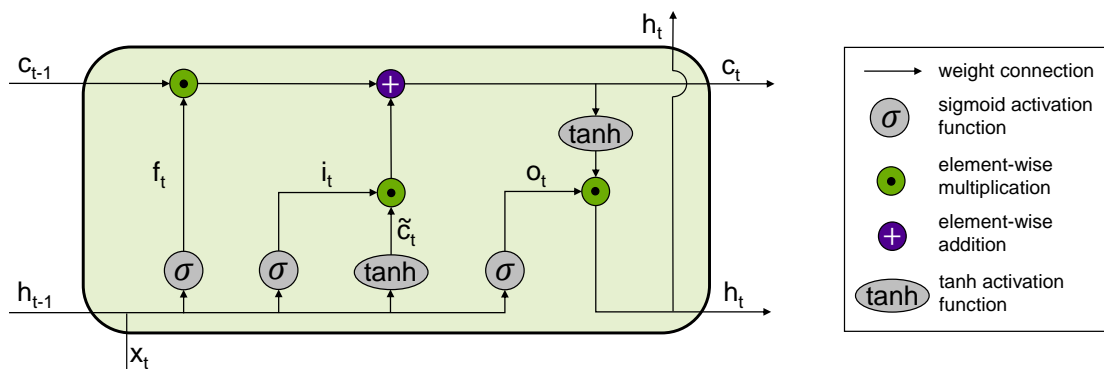
Figure 4.4: Internal architecture of a LSTM cell (Source: Own illustration based on [82])

The functionality of the LSTM cell shown above can be expressed mathematically by the following formulas [91]:

$$f_t = \sigma(W_{f,x}x_t + W_{f,h}h_{t-1} + b_f) \tag{4.2}$$

$$i_t = \sigma(W_{i,x}x_t + W_{i,h}h_{t-1} + b_i) \tag{4.3}$$

$$\tilde{c}_t = \tanh(W_{\tilde{c},x}x_t + W_{\tilde{c},h}h_{t-1}) + b_{\tilde{c}}) \tag{4.4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{4.5}$$

$$o_t = \sigma(W_{o,x}x_t + W_{o,h}h_{t-1} + b_o) \tag{4.6}$$

$$h_t = o_t \odot \tanh(c_t) \tag{4.7}$$

While $W_{f,x}$, $W_{f,h}$, $W_{i,x}$, $W_{i,h}$, $W_{o,x}$, $W_{o,h}$, $W_{\tilde{c},x}$ and $W_{\tilde{c},h}$ label weight matrices, b_f, b_i, b_o and $b_{\tilde{c}}$ are bias vectors. Moreover, x_t denotes the current input and c_{t-1}, c_t, $\tilde{c}_t$, h_{t-1}, h_t represent the previous cell state at time t-1, the new cell state, the candidate cell state, the previous hidden state and new hidden state respectively. The hidden state (output) can be understood as the networks short-term memory, while the cell state is the memory of the LSTM and captures long-term dependencies. The arithmetic operations $\odot$ and $+$ indicate element-wise multiplication (Hadamard product) and element-wise addition. Finally, the sigmoid and tanh activation functions are represented by σ and $\tanh$. While the tanh activation function serves to control the values passing through the network by ensuring that they are in the range between -1 and 1, the sigmoid activation function outputs values between zero and one. The sigmoid function is employed in all three gates and decides on which signals should pass the gates. While a value of zero causes signals to disappear, a value of one

ensures that the signal remains the same and passes the gate.

The forget gate f_t is the first gate of the LSTM unit (see 4.2) and regulates how much information to keep from the previous cell state. Both the current input and previous output are multiplied with their corresponding weights and passed through the sigmoid activation function along with the bias vector. The output values between zero and one controls the amount of information to be deleted from the previous cell state (zero= delete all, one = keep all).

The next stage, which comprises two elements, is to decide what new information to store in the cell state. First, the input gate i_t determines which values to update. As shown in Equation 4.3, the computation of the input gate is similar to the forget gate, the only distinction concerns the different weight matrices and bias vectors of both gates. Next, a candidate cell state $\tilde{c}_t$ is calculated by applying the tanh activation function to the sum of the weighted current inputs, weighted previous outputs and bias. Hereafter, the Hadamard product of input gate and candidate cell state decides which information is important to keep from the candidate cell state, depending on the output values of the input gate (zero=not important, one=important).

With the help of the forget and input gate and the candidate cell state, the new cell state can be calculated according to Equation 4.5. First, the Hadamard product between previous cell state and forget gate vector is calculated. Values of the previous memory are discarded in case of multiplication by values close to zero. The new cell state is then determined by the element-wise addition of the result of the Hadamard product and the result of the element-wise multiplier between input gate and candidate state. This operation ensures that the previous cell state is updated to new values that the LSTM considers relevant.

Lastly, the output gate o_t controls the information to output based on the cell status and can be understood as a filtered version of the cell state. As seen in Equation 4.6 the weighted current inputs and weighted previous output along the bias vector are passed to the sigmoid activation function to determine which part of the current cell state to output. The current cell state is pushed through a tanh function and then element-wise multiplied by the result of the output gate to extract only the relevant information as output. The resulting new hidden state is used for prediction and the new hidden state and new cell state are then transferred to the next time step.

5. Methodology: Multivariate multi-step LSTM forecasting

After having laid the foundations important for the understanding of this work in the previous chapter, this section will present the authors' own approach.

The novel LSTM approach for predicting the PEV charging load introduced in this paper is implemented in the Python programming language, which is interoperable with Version 3 and beyond. Python is chosen due to its reputation in the area for data analysis, its open accessibility and the large variety of libraries available. The primary Python libraries used in this thesis are:

1. Pandas, NumPy, scikit-learn,
2. Keras, TensorFlow and
3. hyperopt.

While pandas and NumPy are mostly used for the data pre-processing and model evaluation, the LSTM is implemented in Keras which uses TensorFlow as a back-end framework. The hyperopt library is used for hyperparameter tuning. Due to performance considerations, training and hyperparameter tuning of the LSTM is performed on the Narvi Cluster – a SLURM managed cluster at Tampere University consisting of 64 CPU and 8 GPU nodes [92].

All libraries were chosen with regard to their popularity, user-friendliness and overall performance. All other packages used in this work are documented in the source code. These additional packages need to be installed before running the different Python scripts. When installed, the software can be run on all major operating systems. In the following, for reasons of clarity, only the name of the function without its required arguments is given when referring to certain implemented python functions. The required arguments can also be taken from the source code under the same function name.

5.1. Methodology

As previously discussed in Section 2.3, different methods for multi-step ahead time series forecasting can be distinguished. In this work a LSTM using the MIMO strategy in a multivariate setting is introduced. The MIMO Strategy is selected based on the results in [31], which conclude that multiple output strategies outperforms single output strategies. Moreover, previous studies indicate that the charging load of electric vehicles is greatly influenced by various factors such as weather, temperature, type of day, hour of day or public holidays [93], [40], [41], [42]. Hence, the multivariate approach is adopted in this thesis. When selecting the multivariate data, the main focus is placed on keeping the complexity of the model low by selecting those which can be easily obtained by the aggregator. Therefore, meteorological and temperature data are not included. All inputs can be derived from the original data, therefore reducing the complexity to a minimum .

The methodology for designing the proposed model in this thesis is shown in Figure 5.1 and can be divided into three main steps – data pre-processing, LSTM training and LSTM forecasting. The data pre-processing will be thoroughly discussed in Chapter 5.2. To start with, the cleaned raw data is used to calculate a time series of aggregated charging load for each charging site, which will be explained in more detail in Section 5.2.2. Based on the calculated load values, an analysis of the different time series is carried out and important features for setting up the model are ultimately selected on the basis of the results (Section 5.2.3). Subsequently, the load and feature data is split into training and test data prior to the encoding of the data (Section 5.2.4). The process of data reshaping according to the requirements of the LSTM network and supervised learning framing will be dealt with in Section 5.2.5.

Following the data pre-processing, the LSTM model is trained, an important process that will be further discussed in Chapter 5.3. In this context, the LSTM hyperparameter (Section 5.3.1), the implementation of training (Section 5.3.2), the overall LSTM architecture and training process (Section 5.3.3) and the hyperparameter tuning (Section 5.3.4) will be addressed.

Once the LSTM has been conditioned, the last step comprises the actual charg-
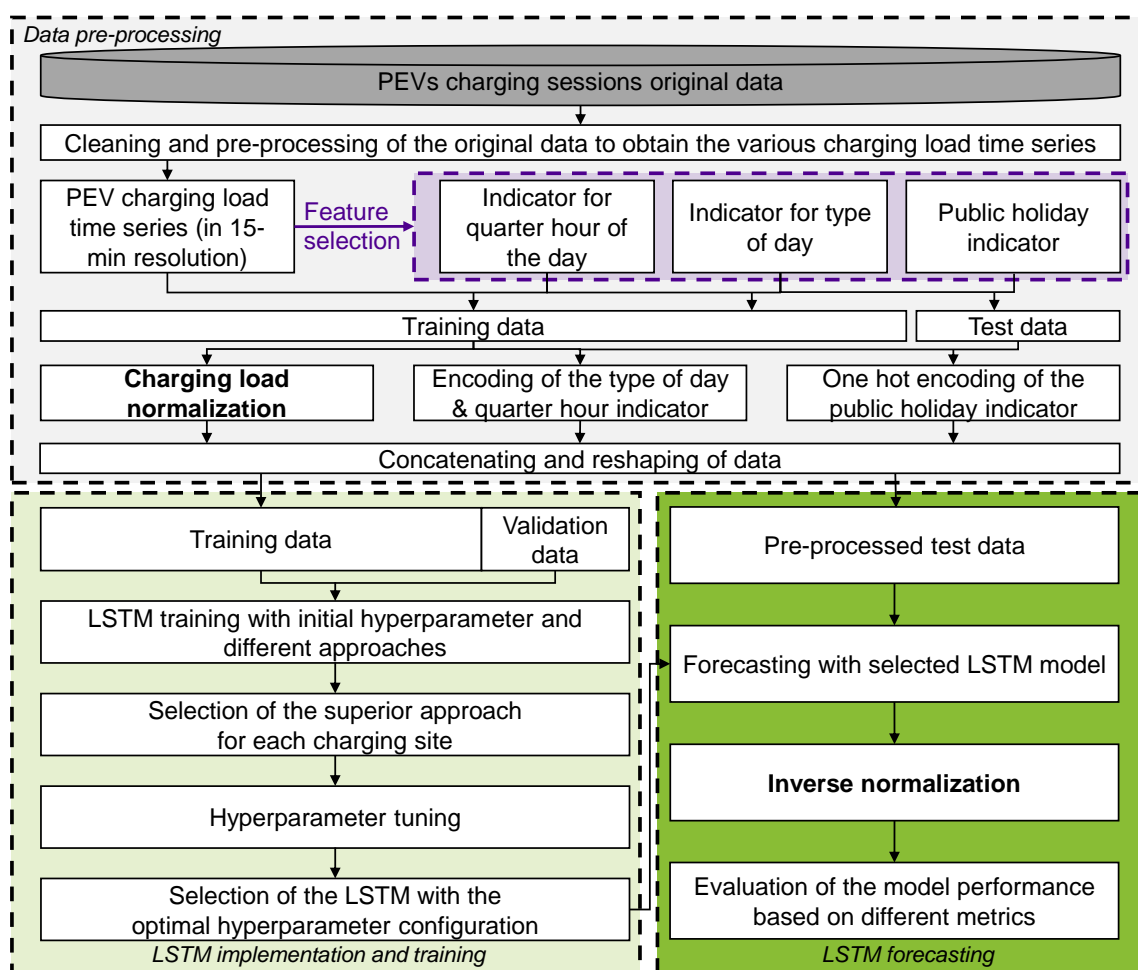
Figure 5.1: Methodical approach of developing the LSTM charging load forecast

ing load prediction (Chapter 5.4). The prediction is carried out by feeding the LSTM with test data that has not been shown to the LSTM before. Based on the prediction, the performance of the model can be evaluated for the different charging sites using a variety of metrics. While the implementation of the forecast is dealt with in Section 5.4.1, the evaluation metrics used in this paper are introduced in Section 5.4.2.

5.2. Data pre-processing

The following paragraph details the process of converting the event based data into an aggregated charging load time series and the subsequent data preparation to enable the processing of the data by the LSTM.

5.2.1. Origin and selection of charging data

The data used in this work originate from two different sources. As mentioned before, a majority of the data has been kindly provided by Parking Energy Ltd, a company headquartered in Helsinki. Further charging data was provided by the in Tampere based software provider IGL-Technologies Oy. IGL operates in the sector of parking management and provides integrated solutions for warming up the car, charging the car and measure electricity consumption.

Table 5.1 depicts the characteristics of the charging data used in this thesis, which have been selected according to two criteria. Firstly, given that ANNs can only achieve good results if a sufficient amount of data is available, sessions of at least one year must be provided for each of the charging sites. The second criterion is that the time periods in which the charging sessions were measured must be the same for each charging site within a cluster. Thus, the sudden increase of the charging load caused by the addition of new charging locations within one year can be avoided. In this context, emphasis was placed on using the latest charging data as far as possible.

Table 5.1: Overview of the original charging session data

Data provider	Type of charging	Number of charging sites	Number of charging sessions	Measured period
IGL-Technologies Oy	Shopping center (REDI)	1	9,283	01.01.2019 – 31.12.2019
Parking Energy Ltd	Residential	21	10,920	21.01.2019 – 20.01.2020
	Public (city car parks)	8	18,785	31.01.2019 – 30.01.2020
	Workplace	7	11,516	31.01.2019 – 30.01.2020

Taking into account the two selection criteria, the charging data of 37 different charging locations, originating from different cities in Finland, can be used in this study. A total number of 21 charging locations are pooled and assigned to residential charging, eight charging sites represent public (car park) charging and the sessions of another seven charging sites form the aggregated load for workplace charging. The dataset available from IGL originates from the shopping center REDI and thus represents the charging in shopping center complexes.

After the selection of the suitable data the datasets contain a total number of 50,504 charging sessions, with every charging site containing measurements over a period of one year. While 10,920 charging events, measured between 21.01.19 and 20.01.20, are allocated to residential charging, 18,785 and 11,516 charging sessions, registered in the period 31.01.19 - 30.01.20, are associated with public and workplace charging respectively. The 9,283 charging sessions of shopping center charging were logged throughout the whole year of 2019.

5.2.2.   Charging load time series generation

In order to generate the suitable input for the LSTM, the event-based data of the accounting system first has to be converted into a time series of aggregated charging load values. The time series is calculated based on the cleaned datasets in Python. Different Python functions are implemented to be able to generate a time series of aggregated charging load in a 15-min resolution for each day over the entire observation period using the different charging events. The data is stored in 2-dimensional array, with the first dimension representing the number of days and the second dimension specified by the number of 15-min charging intervals. While the number of rows is thus determined by the time span between the first and last charging measurement and equals 365, the number of columns totals 96 (4×24). The 15-min interval size is chosen as a compromise between complexity and accuracy. Furthermore, 15 minutes is the smallest time unit to be traded on the Intraday Market of EPEX Spot, which is why the choice of the 15-min interval should be considered reasonable [94].

Figure 5.2 depicts the selected approach to create the charging load time series from the individual charging sessions. While indices k denotes the individual charging sessions, indices l represents the total number of charging sessions in the dataset. In a first step the dataset is read into a pandas dataframe using a SQL connection. The SQL connection allows easy access to the event-based information required to generate the aggregated load time series – the start-of-charging timestamp (ts), the power-to-zero ts and the amount of energy charged for each charging event k. While the IGL data set contains the power-to-zero ts for each charging session k,
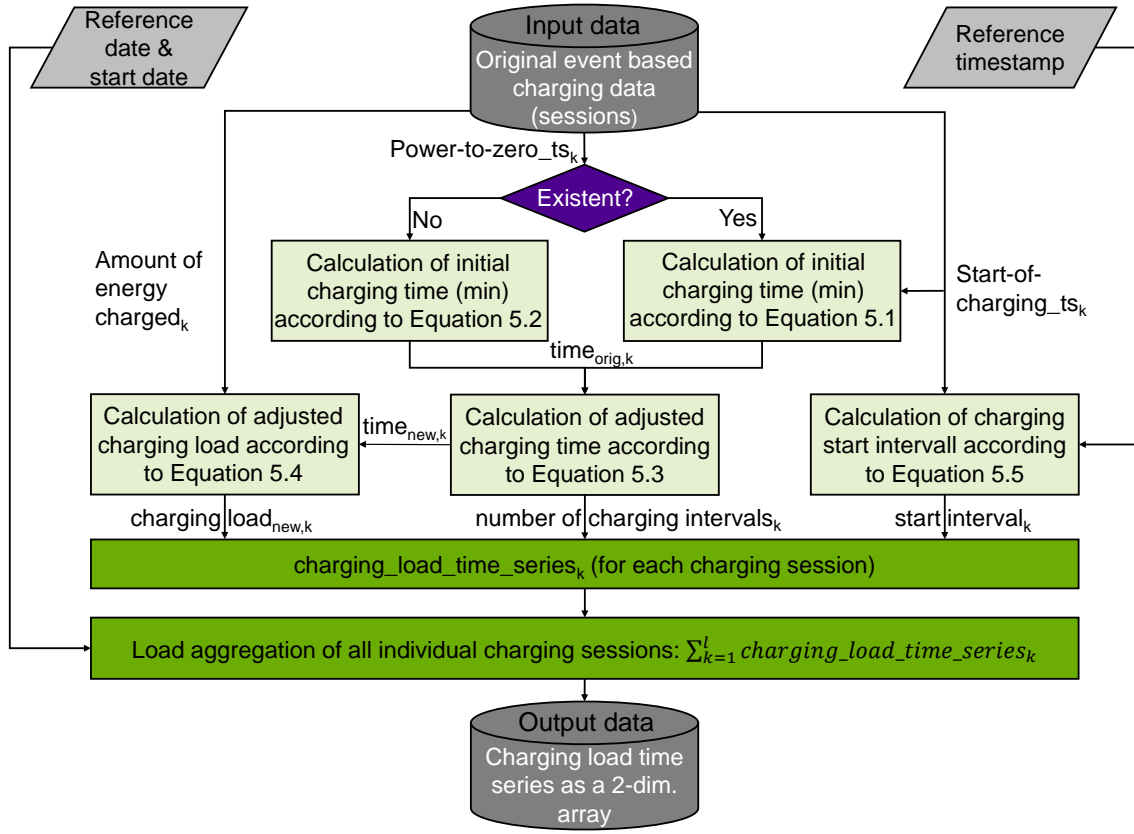
Figure 5.2: Overview of the aggregated charging load time series generation process

not all sessions of the Parking Energy Ltd dataset possess such information. Thus, the initial charging time $time_{orig,k}$ is calculated in two different ways according to Equation 5.1 and Equation 5.2. If the power-to-zero ts is given, the original charging time in minutes is derived by measuring the deviation between start-of-charging ts and power-to-zero ts in seconds (sec) and dividing the result by 60 sec. In case the power-to-zero ts is missing, a nominal charging power of 1.8 kW is assumed and the charging time in minutes is calculated by dividing the amount of energy charged $energy_{orig,k}$ (in kWh) by the charging power and multiplying the result with 60 min per hour.

$$time_{orig,k} = \frac{power\text{-}to\text{-}zero\_ts_k - start\text{-}of\text{-}charging\_ts_k}{60} \tag{5.1}$$

$$time_{orig,k} = \frac{energy_{orig,k} \times 60}{1.8} \tag{5.2}$$

Since the parking and charging times might change due to the analysis interval of 15 minutes, the charging time needs to be adjusted accordingly. Equation 5.3 shows

how the new charging time $time_{new,k}$ is calculated based on the original charging time.

$$
time_{new,k} = \begin{cases} 15 & time_{orig,k} < 15 \ min \\ time_{orig,k} - mod(time_{orig,k}) & mod(time_{orig,k}) < 7.5 \ min \\ time_{orig,k} - mod(time_{orig,k}) + 15 & else \end{cases} \quad (5.3)
$$

The new charging time indicates the number of full 15-min intervals of each charging session. In the event of $time_{orig,k}$ lasting less than 15 minutes, it is rounded up to 15 minutes which equals one 15-min charging interval. Otherwise, a modulo operation is performed which outputs the rest of an integer division by 15. Any time the output value falls below 7.5 minutes, it is subtracted from the original charging time, in other words the amount of 15-min charging intervals is rounded down to the next integer. Otherwise, the deviation between 15 and the output value is added to $time_{orig,k}$, thus the number of charging intervals is rounded up. The new charging time and corresponding number of the full 15-min charging intervals for each charging session is calculated with the help of the python functions $calculate\_min\_charging(...)$ and $calculate\_num\_charging\_interval(...)$.

By means of the adjusted charging time, the modified average charging load of each charging process can be calculated subsequently. Although it must be assumed that the maximum charging load decreases with an increasing state of charge, a constant charging power is assumed in this paper. [95] reveals that the resulting error caused by a steadily assumed charging power during slow charging is relatively small and can thus be neglected for the purpose of this paper. The calculation of the average charging load $charging\_load_{new,k}$, using the $calculate\_charging\_load(...)$ function, can be taken from Formula 5.4 . To determine the charging load, the original amount of energy charged is divided by the modified charging time.

$$
charging\_load_{new,k} = \frac{energy_{orig,k}}{time_{new,k}} \quad (5.4)
$$

In order to determine the aggregated charging load of all charging events for all 15-min intervals a day throughout the observation period, the start interval is of interest in order to accurately schedule the load. All charging events whose start-of-charging ts falls between the start and end of a 15-min interval must be adjusted accordingly.

It is important to ensure that the deviation from the original start-of-charging ts is as small as possible. Therefore, for each charging session a reference ts for the same day of the start-of-charging ts at time 00:00:00 is created. The difference between reference ts and start-of-charging ts enables to determine the start interval, illustrated in Equation 5.5.

$$start\_interval_k = \frac{start\text{-}of\text{-}charging\_ts_k - ref\_ts_k}{60 \times 15} \tag{5.5}$$

The difference between both ts is measured in seconds. By dividing the difference by the amount of seconds and minutes of each interval and subsequently rounding the result, the start of the charging process can be assigned to one of the 96 different 15-min intervals of each day. For this purpose function $calculate\_start\_interval(...)$ is implemented. If the start-of-charging time falls within the first 7.5 minutes of each interval, the start interval is set to the value of the respective interval. Otherwise the charging process starts in the following 15-min interval, thus ensuring the closest match between the original and modified charging process.

To be able to assign the charging load of each session to the correct day, additional information is required to determine the correct row index. Therefore, a reference and actual data datetime object is created by means of functions $create\_start\_date(...)$ and $create\_reference\_date(...)$ which allow to specify where the charging load of each session should be inserted. As a result, all necessary values are available to determine the aggregated charging load in a final step. For each charging session, the computed charging load is stored at the specific day beginning at a given start interval over the determined duration of the charging process. If the charging process spans over two days, this circumstance is taken into account accordingly. Afterwards, the load calculated for each charging session is accumulated iteratively. The resulting output is a 2-dimensional NumPy array containing the aggregated charging load time series of all charging sessions.

5.2.3. Analysis of time series and feature extraction

To be able to determine important features which can be fed to the LSTM as multivariate inputs to support the detection of interrelationships, an analysis of the char-

acteristics of the load series of the different charging sites is carried out in the following. Additionally, the analysis aims at establishing a basic understanding of the characteristics of the load series in order to allow a more thorough assessment of the results in Chapter 7.

Analysis of annual charging load curves

To start with, the course of the aggregated charging load over the entire period of one year is evaluated. Figure 5.3 illustrates the trajectory of the charging load for both public (upper plot) and workplace charging (lower plot) in the period 31.01.2019 to 30.01.2020.
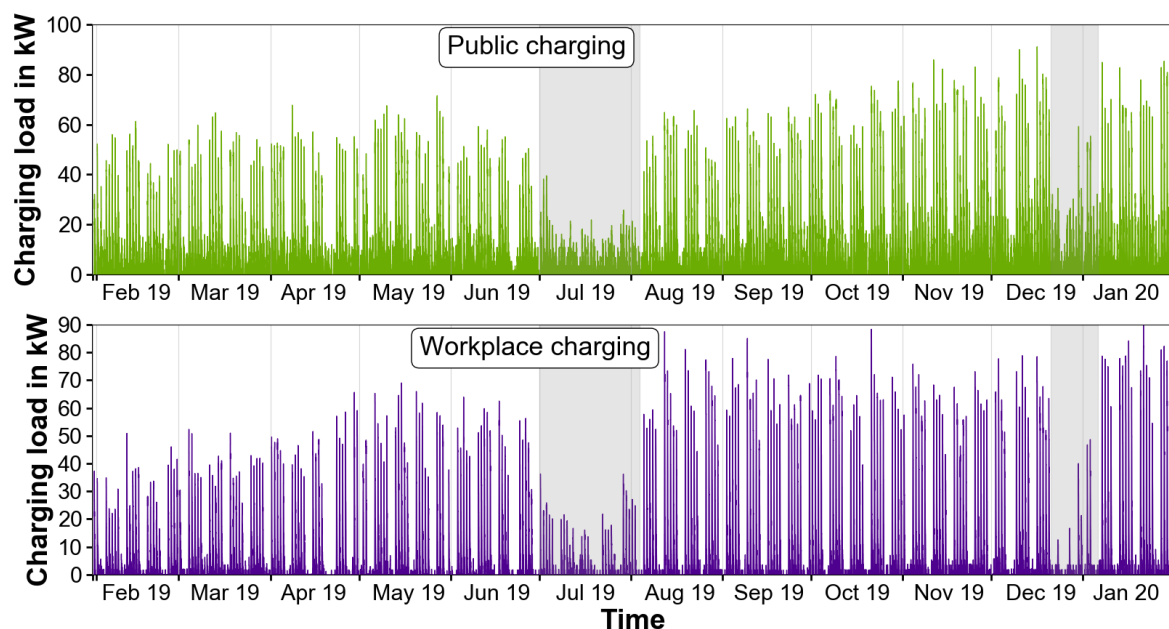
Figure 5.3: Aggregated charging load of workplace and public charging over the period from 31.01.19 to 30.01.20

Three observations can be drawn from the plots. First, an increasing trend of the aggregated load over the course of the year can be observed for both charging sites. Considering the increasing number of PEVs and the associated number of charging processes, this finding is not surprising. The noticeably reduced charging load in the gray-shaded area marks the second conspicuity. This phenomenon is caused by the two main holiday periods in Finland, the summer holidays dating from the end of June to the beginning of August and the Christmas holidays between Christmas and the start of January. Lastly, it also is visible that workplace generally exhibits steep peak loads with a relative low charging load level the rest of the day, while

public charging also displays high peak loads but with a slower decay of the charging load.

By comparison, Figure 5.4 displays the course of the aggregated charging load for shopping center (upper plot) and residential charging (lower plot). Similar to the previous chart, the rising trend over the year and the influence of holiday periods on residential charging is evident, although less significant than before. The load profile of the shopping center also exhibits a slight trend, but the influence of summer and Christmas holidays is not clearly visible. This observation may be due to the fact that the charging stations in the shopping center are used for different purposes, for shopping, during working hours or as a parking space for nearby residential areas.
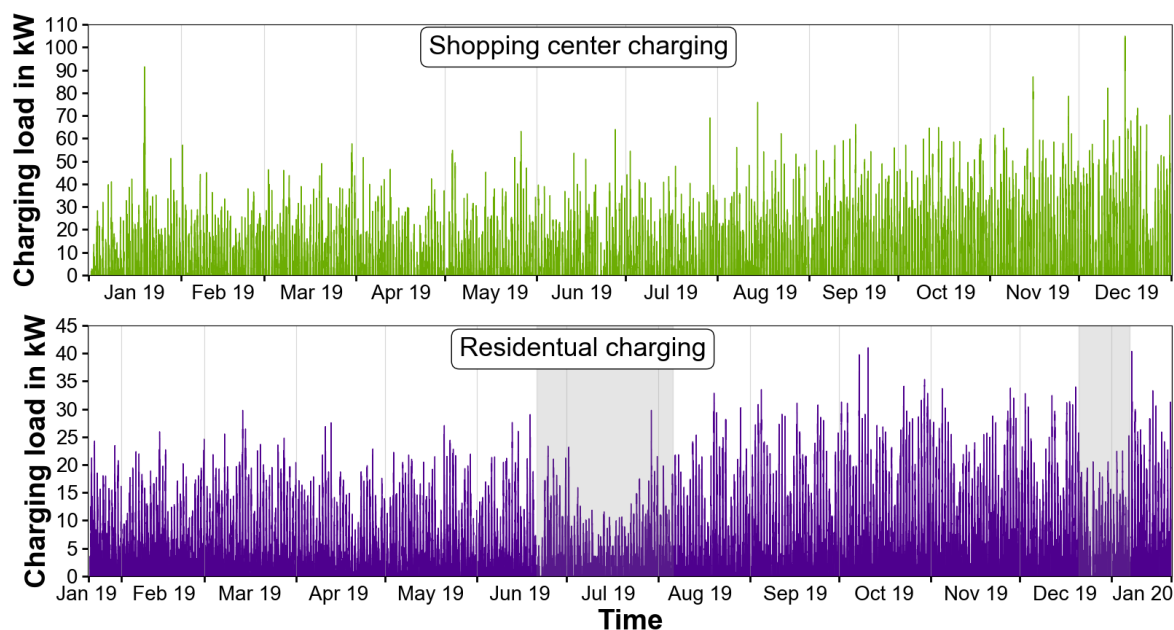
Figure 5.4: Aggregated charging load of shopping center (01.01.19 - 31.01.19) and residential (21.01.19 - 20.01.20) charging

In summary, the rising trend would support the selection of an indicator for the different months as a feature. However, due to the fact that only charging data over one year is available and the train and test split would therefore result in unknown indicators, a feature month is omitted. Nevertheless, with data available for several years in the future, the inclusion of an indicator for the month might be reasonable for later investigations.

Analysis of weekly charging load curves

Following the review of the charging load over the entire time span of one year, a more detailed analysis of individual weeks is carried out with the intention to show further properties. For each charging site the weeks from 04.02.19 to 10.02.19 (green plot), from 25.11.19 to 01.12.19 (gray plot) and the Christmas week from 23.12.19 to 29.12.19 (purple plot) are analyzed. In this way, the influence of the different types of day on the charging load can be demonstrated, as well as how the daily charging load increases over the course of the year due to the previously seen increasing number of PEVs.

Figure 5.5 displays the weekly load curve in kW for public (upper plot) and workplace charging (lower plot).
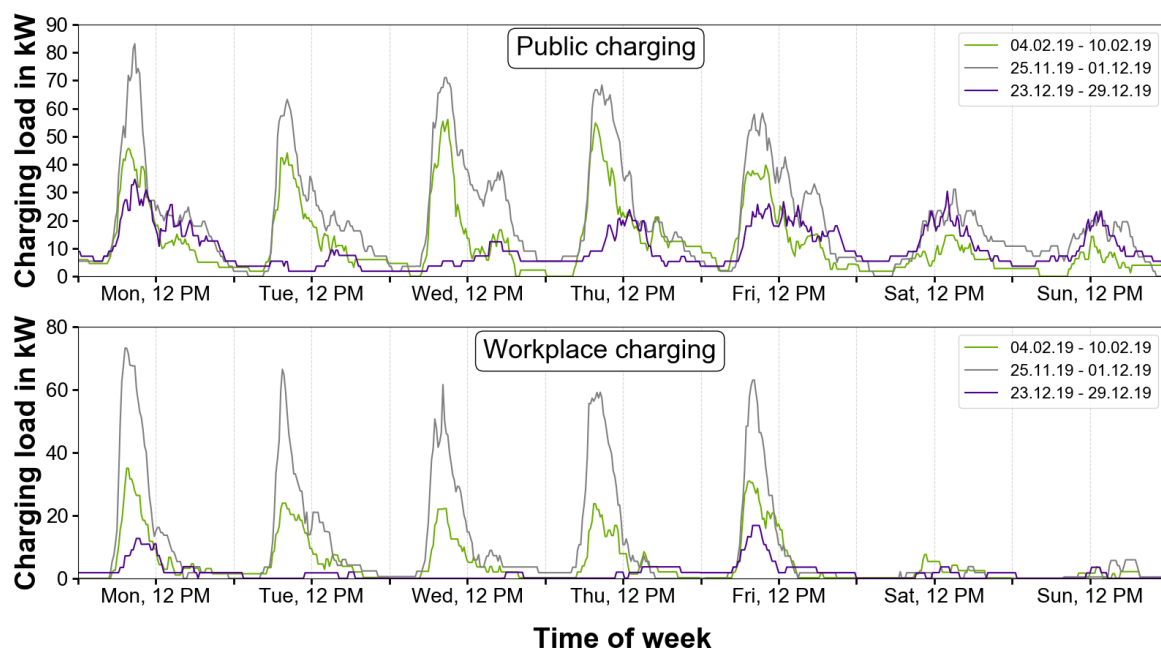
Figure 5.5: Workplace and public charging load exemplified for different weeks

Comparing the load curve in February (green) with that of the end of November (gray), a significant increase can be seen for both charging sites. While the daily peak for workplace charging more than doubles on some days, the surge for charging in public car parks is less pronounced but also clearly visible.

Furthermore, distinct patterns can be identified. As expected, the workplace charging load at weekends is negligible, while weekdays show a similar tendency. The maximum peak arises before noon when the majority of people arrive at work and

start charging. With public charging this pattern is also visible, indicating that many people use the car park for parking their car during working hours. Nevertheless, the curve does not flatten out around midday as in pure workplace charging, but rather has a smaller second peak in the afternoon. On weekends, the load curve drops sharply due to the loss of users during working hours and the peak shifts to after midday.

The impact of public holidays on the load curve becomes clear when observing the purple plot. On public holidays (Tuesday, Wednesday and Thursday), the workplace charging load drops to almost zero. On the remaining days it is also reduced significantly, which can be explained by bridge days and the weekend. The same effect can be seen with public charging, but to a lesser extent.

As shown in Figure 5.6, the rise of the aggregated charging load over the course of the year can also be observed for residential and shopping center charging by comparing the green and gray graphs. In addition, the significant drop of charging load on public holidays is evident when examining the purple load curve for both charging sites. Apart from the peak in the early evening, which can be observed during residential charging on weekdays, less significant patterns can be seen for both charging locations compared to public and workplace charging. Especially the load curves of the shopping center show strong fluctuations with daily changing load profiles.
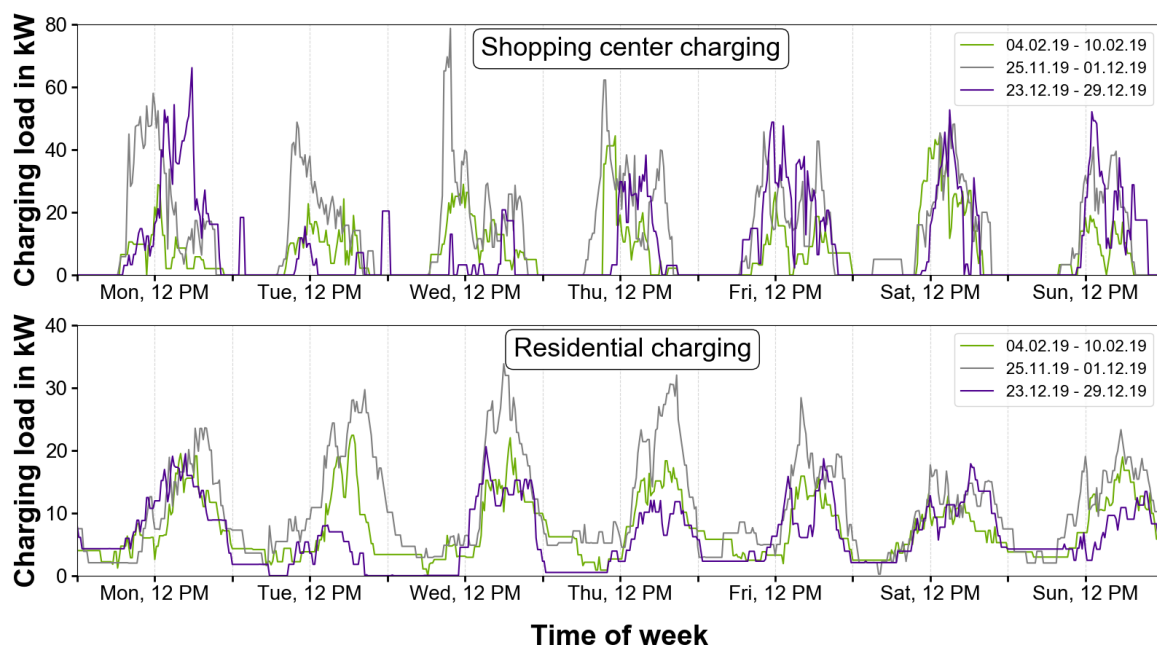
Figure 5.6: Shopping center and residential charging load exemplified for different weeks

Summarizing the findings of the analysis on the weekly load profiles, a clear impact of the type of the day and the time of the day on the charging load can be identified. Consequently, each load value will be linked to an indicator for the quarter of an hour of each day, for the type of day and whether it is a public holiday, collectively representing the multivariate inputs to the LSTM.

5.2.4. Scaling and encoding of data

After completion of the feature selection, the input data for the LSTM is determined. The input data can be divided into numerical and categorical features. While the charging load values are clearly numerical, the three other features selected in this paper can be classified as categorical. Moreover, type of day and quarter hour of the day exhibit cyclic calendrical properties while the feature public holiday possesses binary character (no holiday or is holiday). Due to the fact that ANNs internally work with numeric data they must be converted into numerical values. It is also common practice in machine learning to normalize numeric data in order to accelerate network convergence and ensure that all entries to the LSTM carry the same weight [37], [96].

One of the most common pitfalls in machine learning is the pre-processing and encoding of the data before splitting the available data into training, validation and test sets. This procedure can result in leaking information from the training data to the test data and should be avoided by all means [97]. Therefore, the charging load time series is first split into training and test sets by means of the python function $train\_test\_split\_loadseries(...)$. As the name suggests, the training dataset is used to train the LSTM and update the weights and bias. During model fitting, a certain amount of the training data is used as validation data. The validation set is needed to evaluate the LSTM model during training and be able to tune the hyperparameters of the model to achieve the most appropriate training results. Last but not least, the test data, the data that has never been shown to the LSTM during training, is used to make predictions and provide an unbiased evaluation on the model performance [98]. Due to the limited amount of available data, in this work, the data is split in training and test set with a ratio of 0.9 to 0.1. Moreover, 20 % of the training data is used as evaluation data during model fit, resulting in a 72 % train, 18 % validation

and 10 % test split.

Min-max scaling

After splitting the charging load time series into training and test sets, the next step involves normalizing the charging load values and encoding the categorical data. For pre-processing of the charging load the Python function named $minmax\_scaling\_train\_test(...)$ is implemented, which normalizes the load values in the range [0,1] according to the commonly used min-max scaling method [37], [38], [36], [99]. For scaling, the $MinMaxScaler(...)$ function of the sklearn library is employed, which operates according to the following formula:

$$load_{t,norm} = \frac{load_t - load_{min}}{load_{max} - load_{min}} \tag{5.6}$$

While $load_{t,norm}$ denotes the normalized charging load value at time t, $load_t$ is the actual load value at time t. $Load_{min}$ and $load_{max}$ represent the minimal and maximal load value of the training data. To avoid data leakage, normalization for training and test data is performed separately. At first the training data is normalized and subsequently the test data is normalized by using the same minimum and maximum values of the training data.

One hot encoding

For encoding the categorical variables various encoding techniques can be distinguished, such as ordinal encoding, one hot encoding, sum coding or binary coding. Ordinal encoding is one of the simplest form of encoding categorical variables, assigning a distinct integer to each category. However, ordinal encoding faces the problem of imposing an order on the variable that might not be true [100]. Especially for the type of day and quarter hour indicator this restraint might mislead the LSTM during training. A superior choice might involve using one hot encoding. One hot encoding is one of the most commonly used methods in machine learning to convert categorical values to numerical ones [100], [99], [101], [102]. During one hot coding, the original element from the categorical feature vector with q cardinality is converted into a new vector of ones and zeros with q elements. The corresponding new element is thereby represented by a one, while the rest of the new elements are zeros.

In order to one hot encode the categorical features a correct numerical indicator for quarter hour, type of day and public holiday must be assigned to each value in the load time series. This integer encoding is done both for the training and test data separately. More precisely, for each of the categorical features, a time series is generated which contains the matching indicator for each 15-min load value of the load time series. For this purpose, three different python functions are implemented. Function $get\_quarter\_hour(...)$ assigns an integer (0 to 95) to each 15-min load value, which indicates the current time of each day. Hereby, 0 represents the first quarter of an hour of the day and 95 the final one. Function $get\_type\_of\_day(...)$ likewise links an integer (0 to 6) to each 15-min load value, indicating the type of day. While integer 0 depicts Monday, Sunday is represented by the integer 6. Last but not least, function $get\_holidays\_finland(...)$ derives a public holiday indicator for the different datasets originating from Finland. Integer 0 marks no holiday while integer 1 signifies public holiday.

After integer encoding the one hot encoding is performed on both the training and test set for each feature, calling the functions $one\_hot\_encoding\_quarterhour(...)$, $one\_hot\_encoding\_type\_of\_day(...)$ and $one\_hot\_encoding\_public\_holiday(...)$. All functions implement the $OneHotEncode(...)$ function of the sklearn library. Table 5.2 exemplifies the results of the one hot encoding. It becomes apparent that one hot encoding drastically increases the input dimension for the LSTM.

Table 5.2: Exemplary illustration of categorical feature one hot encoding

Category	Feature	Ordinal encoded	One hot encoded	Dimension
Type of day	Monday	0	[1, 0, 0, 0, 0, 0, 0]	7
	Tuesday	1	[0, 1, 0, 0, 0, 0, 0]	7
	…	…	…	…
	Saturday	5	[0, 0, 0, 0, 0, 1, 0]	7
	Sunday	6	[0, 0, 0, 0, 0, 0, 1]	7
Public holiday	No holiday	0	[1, 0]	2
	Is holiday	1	[0, 1]	2
Quarter hour of the day	00:00:00 – 00:14:59 (first)	0	[1, 0, 0, … , 0, 0, 0]	96
	00:15:00 – 00:29:59 (second)	1	[0, 1, 0, … , 0, 0, 0]	96
	…	…	…	…
	23:45:00 – 23:59:59 (last)	95	[0, 0, 0, … , 0, 0, 1]	96

Sine/cosine encoding

One hot encoding not only increases the dimensionality of the input feature quarter hour and type of day but might also not be the most suitable choice for representing their cyclic characteristic. Several publications therefore use sine and cosine transformation to enhance the ability to recognize the cyclic nature of features like hour of the day, type of day or month and conclude that this type of encoding can significantly improve the neural network performance [103], [104], [105], [106]. While the feature public holiday does not possess a cyclic nature, day of the week and quarter hour of the day do and are therefore encoded with sine and cosine as well. Both one hot encoding and sine/cosine encoding are tested in this paper. The formula for sine and cosine encoding for the quarter hour indicator at time t is illustrated by Equation 5.7 and Equation 5.8 respectively. Equation 5.9 and 5.10 depict the sine and cosine encoding of the type of day indicator at time t. Both training and test data are encoded according to the shown equations using $sin\_cos\_encode\_quarterhour(...)$ and $sin\_cos\_encode\_type\_of\_day(...)$.

$$sine\_quarter\_hour_t = \sin(\frac{2 * \pi * int\_quarter\_hour_t}{96}) \tag{5.7}$$

$$cosine\_quarter\_hour_t = \cos(\frac{2 * \pi * int\_quarter\_hour_t}{96}) \tag{5.8}$$

$$sine\_type\_day_t = \sin(\frac{2 * \pi * int\_type\_day_t}{7}) \tag{5.9}$$

$$cosine\_type\_day_t = \cos(\frac{2 * \pi * int\_type\_day_t}{7}) \tag{5.10}$$

Figure 5.7 illustrates that the feature dimension can be reduced to two for both features and that the sine/cosine encoding is able to represent the cyclic nature of both cyclic variables. In comparison to one hot encoding for instance, the proximity between Monday and Sunday (see left plot) and between the last quarter hour of the previous day and the first quarter hour of the next day (see right plot) can be mapped more precisely. However, both encoding techniques will be implemented to investigate which one is yielding more accurate results and select the superior encoding technique for each charging site.
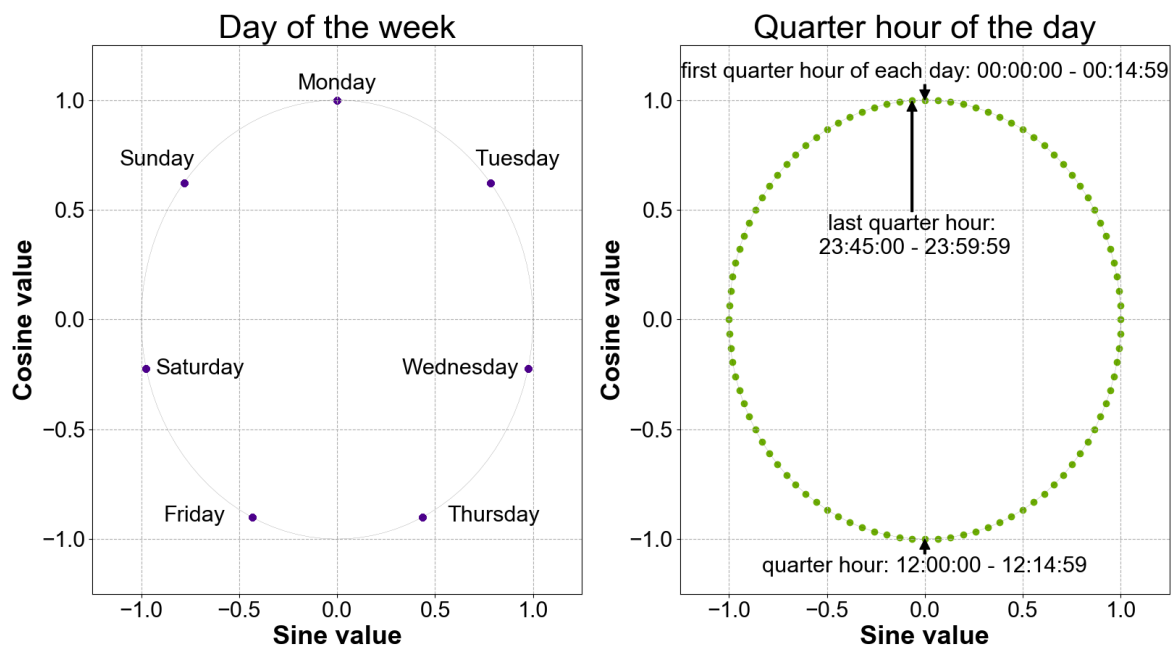
Figure 5.7: Sine/Cosine encoding of temporal features

5.2.5. Reshaping and supervised learning framing

Finally, the input data must be formatted in such a way that the input data shape meets the specific LSTM requirements and that the input data is presented to the LSTM network appropriately so that the supervised learning approach used in this thesis can be successfully applied. The reshaping is done separately for the training and test data, which will not be explicitly mentioned every time in the following for reasons of simplicity. First of all, the features that are still provided separately from each other at this point in time must be merged together. By calling the $concatenate\_all(...)$ function, for every 15-min timestep all features are merged into a single array. The merging is done in the following order: normalized load value, quarter hour indicator of the day, type of day indicator and public holiday indicator. The result is a 2-dimensional array that contains all mentioned features for each timestep. For one hot encoding, each timestep possesses 106 features, while the number is reduced to seven features in the case of sine/cosine encoding.

Furthermore, the data has to be prepared in such a way that the LSTM can be trained and tested afterwards. Within this thesis two approaches are examined, which are illustrated in Figure 5.8 for the month January at the shopping center charging site and 96 timesteps as inputs and outputs. The upper graph depicts the

so-called stateless approach, which uses the sliding window method to artificially increase the amount of training data without altering the results. The LSTM is always given the fixed sequence of input timesteps as input data to determine the load values of the output sequence consecutive to the input sequence. The sliding window enables the creation of a large amount of training data by shifting the input and output sequence with fixed length in each step by one 15-min timestep. Although the stateless sliding window approach has the advantage of drastically increasing initial training data, it also has a significant drawback. During training the internal state of the cells is reset after each batch, therefore implying that data in each batch in not related to other batches [107]. This assumption might be not valid for charging load forecasting, hindering the LSTM to detect long term dependencies outside each batch.

The second approach, illustrated in the lower plot of Figure 5.8, deals with the aforementioned restriction by using the stateful mode. During stateful training, the hidden state calculated for a previous batch of training data serves as the initial state of the following batch of training data. Therefore, the hidden state is maintained across batches and long term dependencies might be more easily understood by the LSTM. Another benefit stems from the fact that the stateful mode can accelerate training time [107]. However, the sliding window cannot be used while using the stateful flag. As seen in Figure 5.8 the fixed input sequences adjoin to each other without overlapping, resulting in a significant decrease of available training data.
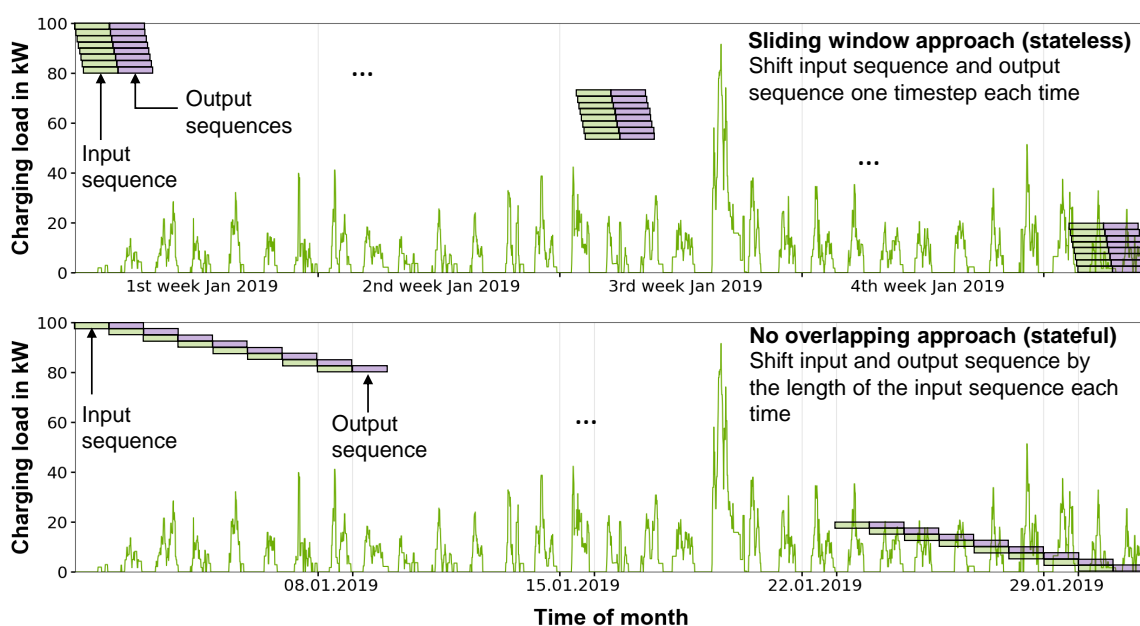
Figure 5.8: Illustration of the stateless and stateful supervised learning framing

The framing of the stateless mode is done by calling $series\_to\_supervised(...)$ which returns a pandas dataframe containing the data framed in a sliding window manner for supervised learning. While the input data x needs to contain both the load value and all encoded features for each timestep, the prediction will only contain the desired charging load values. Therefore, for supervised learning, the real output values shown to the LSTM for comparison with the predicted values may also only contain the load values. This necessity is achieved by calling the function $drop\_columns\_series\_to\_supervised(...)$ which drops all the columns containing the unwanted encoded features for the output y. Finally, the $split\_input\_output(...)$ function is employed to separate the inputs x and corresponding outputs y and store them as an array. Ultimately, the input and output data needs to be reshaped into the 3-dimensional format requested by the LSTM. Using the $reshape\_2d\_to\_3d(...)$ function the data is reshaped into the shape $[number\ of\ samples,\ number\ of\ timesteps,\ number\ of\ features]$.

Figure 5.9 exemplifies this process for the stateless sliding window approach by the example of 96 input and output timesteps. Whereas indice m depicts the number of timesteps of the the whole charging load time series, indice n represents the number of features.
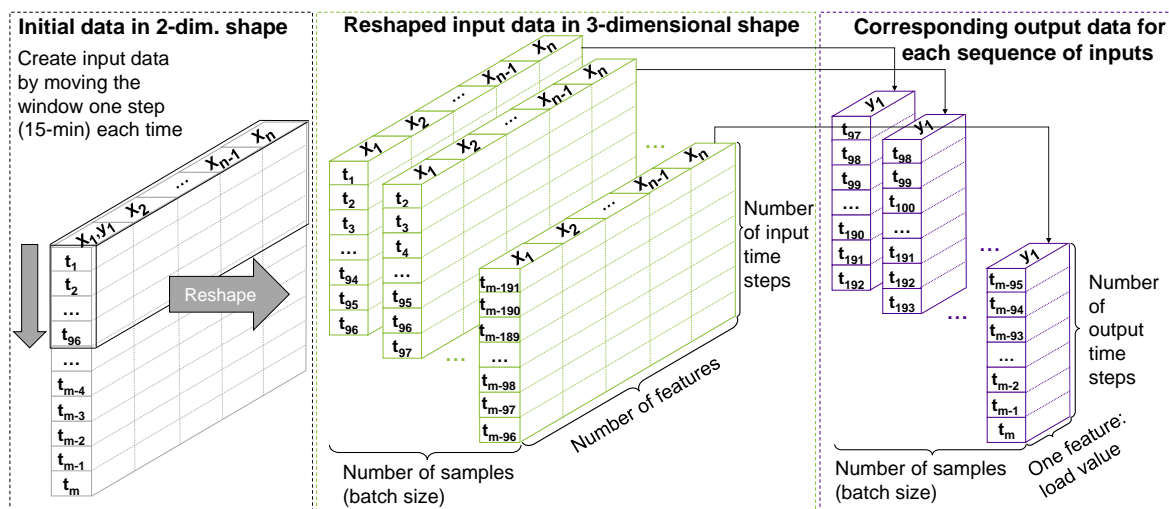
Figure 5.9: Visualization of the reshaping of input and output data in the stateless sliding window approach as demanded by the LSTM

The supervised learning framing for the stateful mode proceeds the same way, with the exception that each sample starts at the next timestep of the last timestep of the

previous sample. The supervised learning framing, the separation of inputs x and outputs y as well as the 2-dimensional to 3-dimensional reshaping is implemented by the $input\_output\_prep\_stateful(...)$ function.

5.3. Long Short-Term Memory implementation and training

Following the data reshaping and supervised learning framing, the LSTM can be implemented and trained. The selection of the LSTM model with the most suitable configuration is carried out in two stages. First, the LSTM is trained with initial carefully selected parameters to choose the most suitable approach for each charging site from the different approaches discussed in the previous chapter. Then, an attempt is made to identify the superior LSTM configuration for the forecast from a number of different hyperparameters.

This chapter is organized in four sections as follows. First, a general understanding of the most important LSTM hyperparameters is provided. Next, the implementation in Python is discussed and the initial selection of hyperparameters is addressed in this context as well. Subsequently, the resulting overall LSTM architecture is illustrated in order to explain the training process in more detail. Finally, hyperparameter tuning is discussed which is used to determine the most suitable configuration of the LSTM.

5.3.1. Long Short-Term Memory hyperparameter

While implementing the LSTM one has to distinguish between model parameter and hyperparameter. Model parameter are those parameters that are internal to the network and modified during training, such as the weights. Contrary, hyperparameters are external and can not be learned during training. Rather, they define the overall design of the model. Therefore, hyperparameters play an important role in the implementation of the model, as even small modifications can have a major impact on the training process and the performance of the model. Depending on the structure of the ANN, the selection of hyperparameters can become very complex. Some of the most important hyperparameters are discussed in the following. They can be

classified into network structure related and training algorithm related hyperparameter.

Network structure related hyperparameter

To start with, the *number of recurrent units* is an decisive hyperparameter for implementing the LSTM design. Both an oversized or undersized number of neurons can lead to difficulties during training. On the one hand the network might not be able to solve the prediction task satisfactorily if the number of units is too small. On the other hand, too many units could lead to overfitting the model, resulting in decreasing test performance [108]. The same behavior also applies to the second hyperparameter, the *number of hidden layers*. Even though in general adding layers can increase model performance too many layers will lead to overfitting and a rapid increase in training time. On the contrary, a too small number of layers can limit the model's ability to learn [109]. Another hyperparameter, the *weight initialization*, uses to have an important impact on the model performance as well. When chosen carefully the right initialization can decrease training time and can avoid bottlenecks during model training. The next hyperparameter, which has already been mentioned in Section 4.1, is the *activation function*. Activation functions form one of the main components of the ANNs and appropriate selection therefore greatly impacts the model's capabilities [110]. Finally, *dropout* also represents a hyperparameter related to the model structure. Dropout is used to reduce overfitting by arbitrarily dropping several neurons during every iteration. In each iteration the model optimizes itself under a slightly different structure, therefore avoiding overfitting and increasing robustness [109].

Training algorithm related hyperparameter

With regard to the training of the LSTM, five important hyperparameters can be distinguished – the loss function, the optimizer, the learning rate, the batch size and the number of epochs. To start with, the *loss function* is the function employed in order to evaluate a candidate solution. The loss function represents a benchmark for how well the predicted values meet the real values [98]. The *optimizer* is closely linked to the loss function. It determines how the network is modified during training based on the loss function [98]. Various optimizers with several advantages and disadvantages are available for updating the network parameters, having a considerable impact on the speed and effectiveness of the training [84], [109], [111]. Another hyperparameter that needs to be mentioned in relation to the optimizer is the *learning*

rate, which is one of the most crucial if not the most important hyperparameter in deep learning. The learning rate controls the increment of each step, in other words it determines how large the weight updates within each iteration may be. In case of an overly high selected learning rate the possibility of skipping a global minimum exists, hindering the ANN to achieve accurate results. However, if the learning rate is selected too low training time can increase drastically with the possibility of stagnation. Therefore, the learning rate must be carefully selected to be able to effectively train the model [75]. The hyperparameter ***batch size*** defines the number of samples that are processed by the ANN during each propagation before updating the weights. While a very small batch size may lead to a problematic increase of training time overfitting can be caused by larger batch sizes [112]. Last but not least, the ***number of epochs*** specifies how often the entire training data is shown to the network. A higher number of epochs naturally increases the training time but might enhance the overall model performance [109].

5.3.2. Implementation of training

Now that the most important hyperparameters have been outlined, the LSTM training implementation and initial selection of hyperparameter will be discussed in more detail. The LSTM model is deployed using the Sequential Model in Keras, which allows easy implementation through simple stacking of different layers by calling the $model.add(...)$ command.

Defining the network structure
The hidden layers are defined by the command $model.add(LSTM(...))$ and possess a number of different arguments, such as the number of units, dropout and recurrent dropout. While in the initial setup no dropout is applied, different number of units are tested – 2, 8, 32 and 128. For weight and bias initialization the standard setting of Keras is employed. Additionally, the input shape is specified through argument passing. While in stateless mode only the number of timesteps and features have to be defined for the input shape, in stateful mode the batch size has to be specified as well and the stateful argument needs to be set to True. Initially, the LSTM is constructed with only one hidden layer. However, in case of multiple hidden layers, the argument return sequences of the previous layers must be set to True as well so that

not only the last output in the output sequence but the entire sequence is passed to the second hidden layer. This argument is important for the later discussed hyperparameter tuning.

As activation functions for the LSTM cells of the hidden layer, the LSTM standard setting in Keras, the Tanh and hard sigmoid function, is adopted. The Tanh activation function outputs values in the range of (-1 to 1) and is calculated according to the following formula:

$$tanh = \frac{e^z - e^{-z}}{e^z + e^{-z}}. \tag{5.11}$$

Similar to the sigmoid activation function used in Section 4.2.3, hard sigmoid outputs values between zero and one but is faster to compute because of the piecewise linear approximation. The hard sigmoid activation function in Keras is computed as follows:

$$hsig = \begin{cases} 0 & z < -2.5 \\ 1 & z > 2.5 \\ 0.2 \times z + 0.5 & else \end{cases} \tag{5.12}$$

A so-called Dense layer is added on top of the hidden layers to specify the number of output values by using the $model.add(Dense(...))$ instruction and argument passing of the number of output timesteps and type of activation function. For the Dense layer, the Rectified Linear Unit (ReLU) activation function is chosen in order to force the outputs to be positive integers, therefore avoiding negative charging load prediction values. Equation 5.13 demonstrates that the ReLU activation function forces all negative values to zero with a linear character otherwise.

$$ReLU = max(0, z) \tag{5.13}$$

Configuring the model training

Once the structure of the LSTM is defined with all network structure related hyperparameter, the $model.compile(...)$ method is used to configure the LSTM for training by specifying both the loss function and the optimizer.

As an optimizer, the popular first-order gradient based Adam optimizer is selected. Numerous studies related to load forecasting or electricity price prediction have shown that the Adam optimizer works well in practice due to its adaptive learning rate and might outperform other optimizer [84], [75], [111], [113], [90]. Given its fast convergence, low memory requirements and little need for tuning Adam is the gradient descent strategy most frequently used in practice [84], [75]. The learning rate is initially set to the default value in Keras of 0.001.

The loss function needs to be chosen according to the prediction task, which can be roughly broken down into regression and classification. Due to the fact that charging load forecasting represents a regression problem, popular loss functions such as MAE and MSE represent potential alternatives. Given that MSE is widely applied in the field of load forecasting [114], [115], [99], [116] and has the advantage of penalizing big outlier predictions, MSE is appointed as the loss function for this work. For charging load forecasting, the goal is to minimize the loss function. The MSE loss is calculated as indicated in Equation 5.14. While N represents the number of predictions, $load_{t,norm}$ and $\hat{load_{t,norm}}$ indicate the normalized true load value respectively normalized predicted load value at timestep t.

$$L(load_{t,norm}, \hat{load_{t,norm}}) = \frac{1}{N} \sum_{t=0}^{N} (load_{t,norm} - \hat{load_{t,norm}})^2 \qquad (5.14)$$

Training the Long Short-Term Memory

To initiate the training process, the $model.fit(...)$ method is called. As arguments, the fit method is given the input and output data x and y, the batch size, the number of epochs and the percentage of training data to be used as validation data. As mentioned earlier, a training to validation ratio of 0.8/0.2 is chosen. The number of epochs is initially set to 500. Moreover, in stateless mode a batch size of 32 and in stateful mode a batch size of one is selected. In stateful operation it must be ensured that the number of samples can be divided by the batch size without remainder. Therefore, a modulo division is implemented to decide which part of the train data can be used to ensure that the number of samples is evenly dividable by the batch size in case of higher batch sizes. The argument shuffle is set to False for both modes to preserve the chronological order of the time series.

In order to select the most appropriate model after training the LSTM with differ-

ent configurations, the training and validation loss for each epoch is recorded and various callbacks are passed to the fit method. Figure 5.10 depicts an illustrative
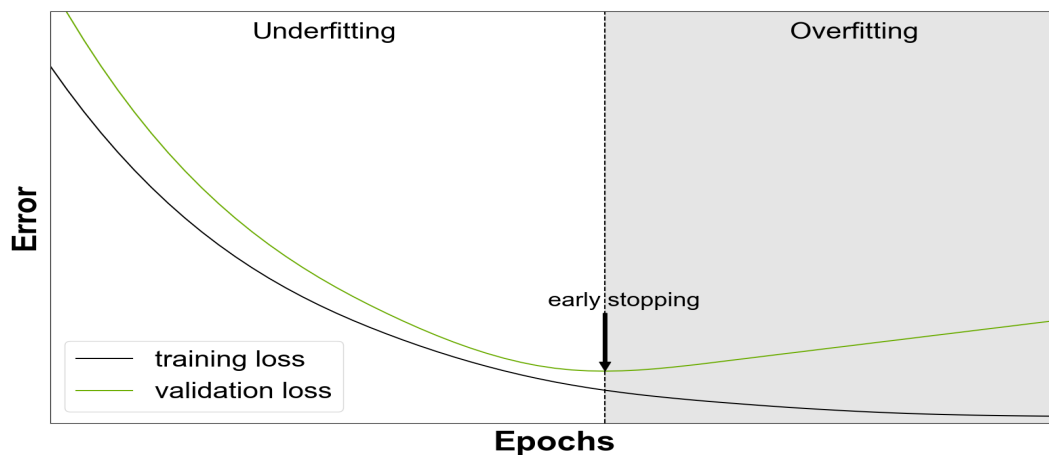
Figure 5.10: Illustration of the early stopping regularization technique (Source: Own illustration based on [117])

course of training and validation loss recorded over a number of epochs and how the regularization technique early stopping can be used to choose the model that generalizes best on unseen data. As shown, the minimum validation loss indicates the model with the greatest ability to generalize to unseen data and the appropriate time to stop the training. A premature stop of the training process, however, leads to underfitting. Underfitting describes a model that is unable to learn the training dataset due to the inability of the model to obtain a sufficiently low training error [118]. If, on the other hand, the training is stopped too late and the validation loss increases again, so-called overfitting occurs. Overfitting refers to a model that has overly memorized the training dataset, resulting in a reduced ability to generalize to unseen data [119].

To prevent the LSTM from continuing training while overfitting, the callback $EarlyStopping$ examines the value of the validation loss after each epoch and aborts the training of the LSTM, if the value does not decrease further over a defined number of epochs. The patience is set to 50 epochs as fluctuations of the validation loss can lead to a temporary increase of the validation loss before dropping again. However, early stopping only works in stateless mode. Two additional callbacks are defined based on the callback $ModelCheckpoint$, storing both the trained LSTM after each epoch and the model with the optimal weights based on the validation loss. The model with the optimal weight is then chosen for the charging load forecast.

5.3.3. Long Short-Term Memory architecture and training process

In the following, the resulting overall architecture of the LSTM and workflow of the training process will be outlined.

Overall architecture

Figure 5.11 illustrates the layout of the proposed LSTM for multivariate multi-step charging load forecasting.
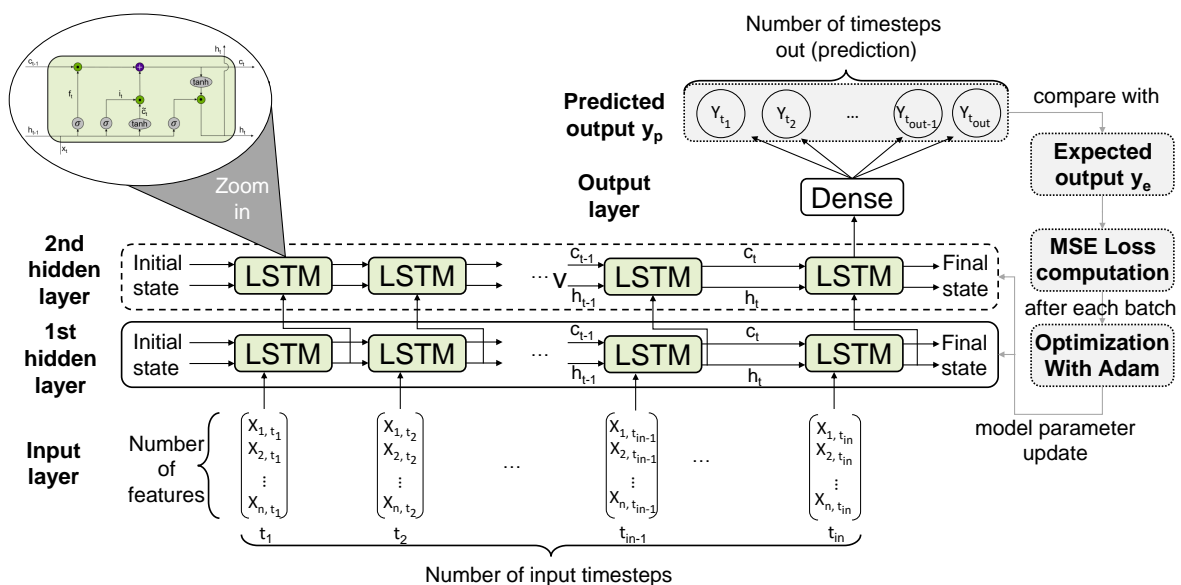
Figure 5.11: Architecture of the proposed LSTM for multivariate multi-step charging load forecasting

Whereas indices n denotes the number of inputs features, indices in represents the amount of input timesteps. Indices out on the other hand depicts the number of output timesteps, in other words the amount of predicted future 15-min charging load values. Within this thesis a load forecast for the next hour (4-timesteps prediction) as well as for the next day (96-timesteps prediction) is provided. A varying number of input timesteps are tested to identify the superior approach to undergo hyperparameter tuning. The variety of timesteps is chosen with regard to the limited amount of training data available and increasing training time when adding a larger amount of timesteps as inputs. For the 1-day prediction in stateless mode both 96 and 192 timesteps as input are analyzed. For the 1-hour predictions in stateless mode the

number of input timesteps is reduced to four, eight, 16 or 32. To allow a consecutive prediction of all timesteps in stateful mode as well, which is desirable in practice, the equal amount of input and output timesteps are applied for the forecast.

Training process workflow

The formatted training data as described in Section 5.2.5 is made available to the LSTM. While the first 80 % of the training dataset is used for training the LSTM, the remaining 20 % is not used for training but to validate the model performance on un-seen data. Both training and validation data consists of multiple samples which are successively processed by the LSTM, each comprising the input sequence and an output sequence. As can be seen in the input layer of Figure 5.11, for each sample the concatenated features of each timestep form the current input for the LSTM cell presented in Section 4.2.3. The dimensionality of the hidden state is defined by the number of units. Frequently, the number of units in LSTM networks is mistakenly in-terpreted as the amount of LSTM cells. However, as shown in the figure, the number of cells is dependent on the count of input timesteps and the number of units only de-fines the output dimension, hence the dimension of the hidden state of each LSTM cell. The cell state and hidden state is passed on from the previous cell to the next cell as discussed in Section 4.2.3. In stateless mode the final state after processing a batch will be removed. In stateful mode, however, the final state for each sample in a batch is provided as the initial state for each sample in the next batch. The state is manually reset after each epoch, as each epoch contains the same time series data.

After the input data has been processed by the cells of the hidden layer the final state of the last LSTM cell in the last hidden layer is passed to a Dense layer which outputs the amount of timesteps to be predicted for each input sequence. After processing a certain amount of samples, defined by the batch size, the forecasted values are compared to the desired actual output values shown to the LSTM to calculate the training and validation loss. Based on the training MSE the model pa-rameter are then modified after each batch size by the Adam optimizer using BPTT. This process is repeated until the maximum number of epochs is met or the training process is terminated prematurely by the early stopping callback due to the valida-tion loss reaching its minimum.

5.3.4. Hyperparameter tuning

As discussed before, the appropriate choice of hyperparameter has a large influence on the learning process of the model and its final performance. Hence, it is unlikely that the initial hyperparameters carefully selected in Section 5.3.2 represent the most suitable configuration for achieving the most accurate prediction results. Selecting the right parameter constitutes an intensive, time-consuming and iterative process. The practice of exploring different values for hyperparameters to enhance the overall model performance is called hyperparameter tuning [109].

Hyperparameter tuning approach
There are three standard approaches to hyperparameter tuning in machine learning, which are widely used due to their ease of application – manual search, grid search and random search. More sophisticated methods exist, such as Bayesian optimization, but are beyond the scope of this paper due to their complex nature [120], [121]. As the name already indicates, the hyperparameter are chosen manually during manual search. This approach demands great experience to derive the optimal parameters with a minimal number of attempts. Using the grid search method, all possible combinations of a defined number and values of hyperparameters are evaluated. The benefit of this approach lies in the fact that it yields the optimal model for the specified grid of hyperparameters. However, the drawback is that it only delivers superior results if the values have been chosen appropriately and the high expenditure of time. Random search eliminates the drawback of a human estimating the optimal hyperparameter value within a certain range by randomization and can reduce computing time. The values are chosen randomly from a certain distribution, thus increasing the likelihood of resulting in a superior performance [109].

In this paper random search is applied for hyperparameter tuning mainly due to two reasons. On the one hand, random search is capable of compensating for the lack of experience in the selection of suitable hyperparameters. On the other hand only a certain amount of time can be spent on the hyperparameter tuning in this thesis due to the time constraint. Given this restriction, random search, as opposed to grid search, is able to cover a wider range of possible hyperparameter combinations.

However, to increase the likelihood of detecting the hyperparameter constellation with superior performance, the dimensionality should be kept to a minimum and only the most appropriate alternatives should be tested during random search. Therefore, not all hyperparameter discussed in the previous paragraph will undergo model tuning with the aim of reducing the dimensionality of hyperparameter tuning to an acceptable level. Moreover, hyperparameter tuning is not applied to all approaches tested in this work. Rather the different model approaches (different encoding techniques: sine/cosine or one hot, different modes: stateful or stateless and different amount of input timesteps) for the 1-hour and 1-day prediction undergo training with the defaults setting. Based on the minimal validation loss after training the resulting superior modeling approach is selected for each charging site to undergo hyperparameter tuning.

Hyperparameter selection

Table 5.3 provides an overview of the initial hyperparameter setting discussed in Section 5.3.2 and the selected parameter for hyperparameter tuning. As can be seen, the loss function, optimizer, activation functions and weight/bias initialization are not subjected to hyperparameter tuning. The number of epochs are increased to 1,000 for the hyperparameter tuning in order to avoid the possibility of underfitting.

Table 5.3: Summary of selected hyperparameters and scope for hyperparameter tuning

	Type of Hyperparameter	Default setting	Hyperparameter tuning
No tuning	Optimizer	Adam	No tuning
	Loss function	MSE	No tuning
	Activation function	Tanh and hard sigmoid (LSTM), ReLU (Dense layer)	No tuning
	Weight/bias initialization	Standart setting Keras	No tuning
	Number of epochs	500 (model approach selection)	1,000 (during tuning)
Tuning	Number of layers	1	[1, 2]
	Number of units	[2, 8, 32, 128]	[2, 4, 8, 16, 32, 64, 128]
	Learning rate	0.001	[0.01, 0.001, 0.0001]
	Batch size	32	[16, 32, 64, 96]
	Dropout	0	[0, 0.2, 0.5]

A total number of five parameters are subject to hyperparameter tuning, leading to

504 possible hyperparameter configurations to choose from. To start with, different values for the number of hidden layers and units are chosen for the tuning process. With respect to the number of layers, in addition to the default setting of one hidden layer, a deeper LSTM with two hidden layers is examined. The number of units varies between two, four, eight, 16, 32, 64 and 128 units with the aim of covering a large spectrum of possibilities. In case of two hidden layers, the same number of units is applied for both hidden layers to reduce the dimensionality of the random search. Moreover, three and four different values are allocated to the learning rate and batch size during hyperparameter tuning. With regard to the learning rate, in addition to the default value of 0.001, both a lower and a higher learning rate amounting to 0.0001 or 0.01 can be selected as possible values. Apart from the default value of 32, possible options for batch size tuning in stateless mode are 16, 64 and 96. Finally, three different dropout variations are examined. The LSTMs can be trained with the default setting of no dropout and a dropout of 20 % or 50 %. If dropout is used, the same dropout mask is applied for each timestep and the recurrent layers as well, as shown in [122] to be a suitable method for LSTMs. Consequently, arguments dropout and recurrent dropout are both set to the same value.

Implementation in hyperopt

Random search for selecting the most appropriate hyperparameter configuration is performed in hyperopt, a popular library for conducting hyperparameter optimization in Python. Random search is conducted for both the hourly and 1-day charging load forecast of all charging sites. In the following, only the most important aspects of the random search implementation in hyperopt will be briefly discussed, a more detailed description is given in [123].

The essential steps of implementing random search in hyperopt can be described as follows:

1. Define the search space.
2. Specify the objective function to minimize.
3. Configure the search algorithm.

To start with, the search space is implemented as a dictionary. The keys of the dictionary define all hyperparameter to undergo optimization. The corresponding value of the dictionary present stochastic expressions. The stochastic expression

$hp.choice(label, options)$ is implemented to define the search alternatives. While options represent a list of the aforementioned selected hyperparameter choices, the label is needed to return the randomly selected parameter choice to the caller.

Next, the objective function needs to be specified. In this case, the validation loss serves as the objective function and indicates the superior hyperparameter configuration. Therefore, the LSTM is implemented as described in Section 5.3.2. However, in order to prevent failures in hyperopt, the training and validation data has to be adjusted so that they are divisible by the different batch sizes without remainder. This condition is met by scaling the training and validation dataset to the closest amount of data divisible by the smallest common divisor of all batch size choices (192). During model fit, the validation loss history is then recorded and the minimum loss for each configuration is returned as the target score.

Lastly, the $fmin$ function is called to perform the random search. Both the objective function and search space are passed as arguments to the function. Moreover, the random search algorithm is defined by specifying $algo = rand.suggest$. The maximum number of evaluations is set to 50 by defining the max_evals argument. A trials object, storing the minimal validation error and hyperparameter setting of each trial, is implemented by configuring the $trials$ argument. To ensure comparability between the different charging sites and reproducibility of the results, the $rstate$ argument is also included in the fmin function, using a seed of one for all charging sites. Both the trials object and the rstate are saved after each evaluation run to ensure that the random search can be continued if the training is interrupted due to the time limit of running each code on the NARVI Cluster.

5.4. Long Short-Term Memory forecasting

After completing the training with the optimal hyperparameter configuration obtained during random search, the LSTM stored with the optimal weights can be used to evaluate the performance of the charging load forecast on previously unseen data. As shown in 5.1, the test data is fed into the LSTM to achieve the prediction task. Subsequently, the predicted values undergo inverse normalization to transform the normalized values to real load quantities. Finally, the model performance is analyzed

on the basis of different metrics. Below, at first the implementation of the prediction will be discussed followed by a review of the metrics used in this thesis to evaluate the model performance.

5.4.1. Implementation of forecast

In a first step, the test data of the stateless and stateful approach needs to be arranged to enable the comparison of different approaches and stateful prediction. Prior to the actual prediction, functions $prepare\_data\_prediction\_stateless(...)$ and $prepare\_data\_prediction\_stateful(...)$ are executed on this behalf.

The stateless test data, which so far includes overlapping sequences due to the sliding window approach, needs to be modified to enable model performance comparison between the different approaches. For the 96-timesteps prediction, the test data of the stateful mode is selected for the stateless approach as well in order to generate the same amount of forecasts to enable the comparison between both approaches. When not using the test data of the stateful mode, the sliding window approach would lead to a much higher number of predictions, thereby falsifying a comparison between the two approaches. For the 4-timesteps forecast, the test data of the stateless approach is prepared in such a way that the forecast results can be directly compared with the 1-day forecast by repeatedly performing the forecast and stringing together the hourly forecast, allowing the results to be displayed graphically as well. The test data prepared by the sliding window approach is therefore filtered in a manner that only every 4th sample of the stateless sliding window approach is used for the prediction.

The stateful test data must be slightly modified to ensure a seamless transition between training and test data, so that the state can be built up during the prediction. For the 1-hour and 1-day forecast, the last four respectively last 96 charging load and feature values of the training data are added to the test data. Although these values were present in the training set, no data leakage occurs. This circumstance is due to the fact that these last values could not be used during the training because the corresponding output values to be predicted are already part of the test data.

In stateless mode the forecast is carried out with a batch size of one, to avoid errors in the prediction process caused by the varying batch sizes during training. To this end, the weights of the trained LSTM are copied with the command $old\_weights = model.get(weights)$ and passed to the modified LSTM model with a batch size of one with the command $lstm\_model.set\_weight(old\_weights)$. The other configuration of the LSTM remains unchanged. In stateful operation it must additionally be guaranteed again that the number of samples is divisible by the batch size without residue. As already practiced during the training of the LSTM, a modulo division is therefore carried out to determine which part of the test data can be used to ensure that the number of samples is evenly dividable by the batch size. Since batch size one is applied for the stateful approach in this paper, the initial data is maintained. Compared to the stateless prediction, the prediction is carried out on both the train and test data. The forecast based on the training data solely serves to establish the state for the prediction on the test data.

The implementation of the LSTM forecast in Keras is handled within function $make\_prediction(...)$, allowing a call of the function with different arguments that indicate the various approaches used within this thesis. The prediction in Keras is performed by the command $model.predict(...)$ or $lstm\_model.predict(...)$ and passing of the inputs of the test data and the number of batch size. In this context, $model$ and $lstm\_model$ indicates the trained LSTM model stored with the optimal weights. For all test inputs passed to the LSTM, the model predicts the corresponding outputs, either four or 96 timesteps into the future.

As result of the load prediction the projected load values are obtained, which can be compared with the real load values for each timestep to assess model performance. However, both the predicted and real output values must be inverse normalized first to obtain the actual charging load values. For this purpose, the initial parameters of the MinMaxScaler used in data pre-processing (see 5.2.4) are retrieved and stored in the variable scaler. By calling $scaler.inverse\_transform(...)$ and separately passing of the predicted and real load values as arguments the real charging load values can be recovered in the following and used for model evaluation.

5.4.2. Evaluation metrics for model performance and comparison

To evaluate the prediction performance of forecasting models a number of different error metrics can be used. Scale-dependent metrics such as MAE, MSE and RMSE and percentage-error metrics like MAPE are some of the most frequently used error metrics for model evaluation [124], [37], [125], [126]. MAE is selected as the first error metric within this thesis due to its simple comprehensibility and calculation, given in Equation 5.15. While N represents the number of predictions, $load_t$ indicates the real load at time t and $\hat{load}_t$ the predicted load value at time t.

$$MAE\,[kW] = \frac{1}{N}\sum_{t=1}^{N}\left|load_t - \hat{load}_t\right| \tag{5.15}$$

However, MAE is scale-dependent, which implies the need for using additional error metrics to be able to interpret the results between the different charging sites. Given its scale-independence, the popular MAPE would provide an easily interpretative error metric. However, the different charging load time series presented within this thesis exhibit a charging load of zero at numerous points in time. For this reason, MAPE cannot be used for overall comparison, as the calculation is based on the division of the error by the true load value at each timestep. To overcome both difficulties two variants of the NMAE are used, as shown in Equation 5.16 and 5.17.

$$NMAE1 = \frac{MAE}{\frac{1}{N}\sum_{t=1}^{N}|load_t|} \tag{5.16}$$

$$NMAE2 = \frac{MAE}{load_{max} - load_{min}} \tag{5.17}$$

While NMAE1 normalizes the MAE using the mean value of the observed real load values as introduced in [127] and [128], NMAE2 is calculated by normalizing the MAE on the basis of the difference between the maximum load $load_{max}$ and minimum load $load_{min}$ of the examined time interval of the time series [38].

For the 96-timesteps forecast three additional metrics are used to be able to better assess the peak load forecast. The average deviation (dev) between the actual daily peak load $load_{dmax}$ and predicted daily peak load $\hat{load}_{dmax}$ is calculated in kW

according to Equation 5.18.

$$peak\_load\_dev\ [kW] = \frac{1}{D} \sum_{d=1}^{D} \left| load_{dmax} - \hat{load}_{dmax} \right| \tag{5.18}$$

While index D specifies the number of predicted days, $dmax$ denotes the daily peak. Likewise, the MAPE is calculated according to Equation 5.19 [129]. Last but not least, Equation 5.20 specifies the average time difference in number of 15-min timeslots separating the real and predicted peak load. While $nslot$ denotes the number of 15-min timeslots indices $slot$ corresponds to the respective 15-min interval of the daily peak loads.

$$MAPE\_peak\ [\%] = \frac{1}{D} \sum_{d=1}^{D} \left| \frac{load_{dmax} - \hat{load}_{dmax}}{load_{dmax}} \right| \times 100 \tag{5.19}$$

$$time\_dev\ [nslot] = \frac{1}{D} \sum_{d=1}^{D} \left| slot(load_{dmax}) - slot(\hat{load}_{dmax}) \right| \tag{5.20}$$

Having thoroughly presented the LSTM proposed in this thesis and having clarified the metrics used for the evaluation of the results, the next chapter will present the findings in full depth.

6. Results

Having covered the essential steps for generating the charging load forecast in the previous chapter, the results are presented in the following. The chapter is structured in three sections. At first, the results of the LSTM training with the initial hyperparameter setting are displayed for each charging site. Subsequently, the findings of the hyperparameter tuning are presented. Lastly, the charging load forecast for each charging site, based on the optimal hyperparameter, will be addressed.

6.1. Training results with initial hyperparameter

To start with, in the following the results of the LSTM training with the default setting of hyperparameters are discussed to determine the mode, encoding variant and number of input timesteps for the hyperparameter tuning. The training is carried out for each charging site and forecast horizon in stateless or stateful mode, with sine/cosine or one hot encoding and a varying number of input timesteps. In stateless mode four, eight, 16, and 32 timesteps are tested as inputs for the hourly forecast, 96 and 192 timesteps for the daily prediction. To allow a consecutive prediction of all timesteps, which is desirable in practice, the equal amount of input and output timesteps are applied for the forecast in stateful mode. For all charging sites, the minimum validation loss and corresponding epoch of each variant are summarized in a table and the average loss calculated from the loss of the different number of units is presented. The variant with the lowest loss is then subjected to the hyperparameter tuning.

6.1.1. Initial training results for shopping center charging

Table 6.1 depicts the MSE results when training the LSTM on the shopping center charging data. Several observations can be drawn from the table. First of all, when comparing the minimal validation loss of the LSTM trained with different number of units, it becomes clear that overfitting occurs much later when using a small number

of units rather than a large number of units. This circumstance is illustrated by the consistently higher epoch number when using two units compared to 128 units.

Table 6.1: Minimal validation losses for shopping center charging when training the LSTM with the initial hyperparameter and different modes, encoding and timestep approaches

Mode	Input & output timesteps	encoding	val. loss* 2 units (epoch)	val. loss* 8 units (epoch)	val. loss* 32 units (epoch)	val. loss* 128 units (epoch)	Average validation loss*
state-ful	4in/ 4out	sine/cosine	12.2641 (109)	5.8727 (248)	5.9513 (69)	33.3763 (1)	14.3661
		one hot	12.8256 (18)	5.9514 (47)	5.4416 (12)	33.3763 (1)	14.3987
	96in/ 96out	sine/cosine	11.8557 (26)	11.5512 (7)	11.0418 (21)	10.2204 (6)	11.1673
		one hot	14.3951 (36)	11.6331 (7)	11.4971 (15)	10.1678 (2)	11.9233
state-less	4in/ 4out	sine/cosine	5.6642 (77)	5.3866 (65)	5.3800 (33)	5.3803 (20)	5.4528
		one hot	5.4362 (135)	5.3237 (33)	5.3670 (19)	5.3473 (18)	5.3686
	8in/ 4out	sine/cosine	5.4578 (151)	5.3930 (83)	5.3505 (22)	5.3513 (21)	5.3881
		one hot	5.3777 (139)	5.3677 (30)	5.3689 (28)	5.3385 (14)	5.3632
	16in/ 4out	sine/cosine	5.4504 (205)	5.3251 (51)	5.3003 (25)	5.2903 (17)	**5.3415**
		one hot	5.4152 (123)	5.3705 (37)	5.3543 (21)	5.3415 (15)	5.3704
	32in/ 4out	sine/cosine	5.4476 (131)	5.4088 (135)	5.3345 (28)	5.3232 (21)	5.3785
		one hot	5.4198 (74)	5.3937 (17)	5.3463 (15)	5.3393 (12)	5.3748
	96in/ 96out	sine/cosine	13.2808 (107)	10.6257 (72)	9.2237 (16)	9.0281 (9)	10.5396
		one hot	11.6514 (25)	9.4725 (27)	9.0637 (8)	8.9356 (7)	**9.7808**
	192in/ 96out	sine/cosine	12.5823 (58)	10.1510 (78)	9.2642 (18)	9.0622 (10)	10.2649
		one hot	11.7118 (36)	9.4719 (20)	9.1177 (7)	8.9634 (7)	9.8162

*All MSE values are given in units of 10^{-3}

Next, it can be seen that the stateless approach for both forecast horizons yields a lower loss when looking at the average validation loss. Looking at the results of the stateful 4-timesteps forecast, it is noticeable that when using 128 units, overfitting starts directly after the first epoch for both encoding variants. The LSTM is not capable of learning, which results in the much higher loss. Lastly, the different input timestep variants and encoding methods yield only very small differences in the validation loss. For the 96-timesteps forecast in stateless mode one hot encoding leads to a lower loss in all variants, whereby 96 input timesteps deliver slightly lower losses than 192 timesteps. Looking at the stateless 1-hour MSE results one hot encoding appears to lead to slightly lower losses for most of the different variants as well with the exception when using 16 input timesteps, where sine/cosine encoding produces slightly superior and the most accurate overall results.

6.1.2. Initial training results for residential charging

Table 6.2 displays the initial training results for the LSTM when using residential charging data. As already seen before, overfitting generally tends to emerge earlier with increasing numbers of units, as indicated by the decreasing number of epochs.

Table 6.2: Minimal validation losses for residential charging when training the LSTM with the initial hyperparameter and different modes, encoding and timestep approaches

Mode	Input & output timesteps	encoding	val. loss* 2 units (epoch)	val. loss* 8 units (epoch)	val. loss* 32 units (epoch)	val. loss* 128 units (epoch)	Average validation loss*
state-ful	4in/ 4out	sine/cosine	2.9682 (240)	2.9772 (25)	3.0110 (16)	3.0122 (14)	2.9922
		one hot	3.0598 (8)	3.0881 (5)	3.1599 (4)	3.1250 (5)	3.1082
	96in/ 96out	sine/cosine	10.5998 (185)	9.7345 (131)	9.1209 (47)	8.8999 (88)	9.5888
		one hot	11.6016 (108)	9.3616 (39)	10.4338 (47)	9.4368 (24)	10.2085
state-less	4in/ 4out	sine/cosine	2.8800 (89)	2.8544 (117)	2.8457 (29)	2.8518 (35)	2.8580
		one hot	2.8931 (116)	2.8951 (42)	2.9033 (22)	2.9132 (26)	2.9012
	8in/ 4out	sine/cosine	2.8691 (189)	2.8530 (75)	2.8338 (57)	2.8340 (38)	2.8475
		one hot	2.8751 (50)	2.8797 (22)	2.9216 (17)	2.8996 (27)	2.8940
	16in/ 4out	sine/cosine	2.8870 (275)	2.8324 (73)	2.8318 (46)	2.8316 (49)	**2.8457**
		one hot	2.8897 (71)	2.8965 (35)	2.8907 (20)	2.9298 (21)	2.9017
	32in/ 4out	sine/cosine	2.8596 (132)	2.8533 (35)	2.8527 (18)	2.8574 (35)	2.8558
		one hot	2.8589 (38)	2.8914 (26)	2.9018 (29)	2.8871 (23)	2.8848
	96in/ 96out	sine/cosine	11.6370 (40)	8.4844 (26)	8.4641 (28)	8.3452 (31)	**9.2327**
		one hot	11.5976 (36)	8.6153 (29)	8.6217 (21)	8.4697 (28)	9.3261
	192in/ 96out	sine/cosine	11.7148 (117)	8.4741 (28)	8.5375 (16)	8.4139 (34)	9.2851
		one hot	11.6693 (53)	8.5207 (23)	8.4856 (43)	8.4829 (18)	9.2896

*All MSE values are given in units of 10^{-3}

Additional similarities can be highlighted as well. Comparing the results of the stateful and stateless mode, the stateless approach provides slightly lower MSE results for both the 4-timesteps and 96-timesteps forecasts once again. Furthermore, the different encoding variants and number of input time steps again lead to only minor differences in the validation loss. However, in contrast to the shopping center charging results, it can be observed that sine/cosine encoding almost exclusively leads to slightly more accurate results for both forecast horizons than one hot encoding. Moreover, for both the hourly and 1-day forecast, the validation losses are lower than for the shopping center charging. For the 1-hour prediction, the lowest average validation loss is generated using 16 timesteps as inputs and sine/cosine encoding. For the 1-day prediction, the lowest average validation loss is also obtained by using the sine/cosine encoding and using 96 timesteps as inputs.

6.1.3. Initial training results for public charging

Table 6.3 lists the LSTM training results on the public charging data. Similar to shopping center and residential charging, the differences in validation loss for the different input and encoding variants are negligible and in most cases a faster overfitting with an increasing number of units can be observed. Additionally, the stateless mode once again delivers slightly more accurate results for both forecast periods based on the average validation loss.

Table 6.3: Minimal validation losses for public charging when training the LSTM with the initial hyperparameter and different modes, encoding and timestep approaches

Mode	Input & output timesteps	encoding	val. loss* 2 units (epoch)	val. loss* 8 units (epoch)	val. loss* 32 units (epoch)	val. loss* 128 units (epoch)	Average validation loss*
state-ful	4in/ 4out	sine/cosine	2.7560 (44)	2.1867 (155)	2.4283 (51)	2.1389 (53)	2.3775
		one hot	3.0056 (167)	2.5535 (500)	2.7790 (145)	2.9909 (96)	2.8322
	96in/ 96out	sine/cosine	15.7161 (467)	13.5583 (430)	11.4747 (334)	12.4940 (10)	13.3108
		one hot	13.9824 (146)	13.3127 (11)	12.3103 (6)	11.1134 (198)	12.6797
state-less	4in/ 4out	sine/cosine	2.4657 (500)	2.4948 (42)	2.1333 (13)	2.0181 (14)	**2.2780**
		one hot	2.6648 (435)	2.6882 (68)	2.7277 (22)	2.6400 (12)	2.6802
	8in/ 4out	sine/cosine	2.6062 (47)	2.2277 (18)	2.2884 (16)	2.3446 (18)	2.3667
		one hot	2.8115 (147)	2.7264 (19)	2.7483 (15)	2.7331 (11)	2.7549
	16in/ 4out	sine/cosine	2.5028 (35)	2.4050 (26)	2.3213 (18)	2.4458 (49)	2.4187
		one hot	2.9130 (52)	2.7274 (45)	2.6435 (14)	2.5853 (9)	2.7173
	32in/ 4out	sine/cosine	2.4622 (100)	2.5497 (24)	2.2842 (15)	2.3475 (16)	2.4109
		one hot	2.7536 (225)	2.7203 (20)	2.6775 (12)	2.5249 (9)	2.6691
	96in/ 96out	sine/cosine	19.5583 (242)	10.6389 (48)	10.2544 (18)	9.7898 (147)	**12.5603**
		one hot	18.3836 (19)	11.5350 (32)	11.0568 (28)	11.0502 (80)	13.0064
	192in/ 96out	sine/cosine	19.2102 (91)	10.7703 (55)	10.3125 (22)	10.4026 (59)	12.6739
		one hot	18.3996 (55)	11.7388 (143)	10.7379 (135)	11.0473 (79)	12.9809

*All MSE values are given in units of 10^{-3}

With regard to the encoding variants, sine/cosine encoding seems to lead to slightly lower losses. Solely in stateful mode one hot encoding might be advantageous for the 96-timesteps forecast. For the hourly forecast the lowest loss can be realised by using 4 input timesteps and sine/cosine encoding. The minimum loss for the 1-day forecast is obtained when applying 96 input timesteps and sine/cosine encoding. Compared to shopping center and residential charging, a lower validation loss is achieved for the 4-timesteps forecast and a higher loss for the 96-timesteps forecast.

6.1.4. Initial training results for workplace charging

Lastly, the initial LSTM training results on the workplace charging data, shown in Table 6.4, reveal a number of previously seen characteristics as well as peculiarities. To start with, overfitting seems to be less of a problem compared to shopping center, residential and public charging. This finding is reflected in the fact that even with an increasing number of units the LSTM can be trained over a large number of epochs before reaching the minimal validation loss. Only in case of the hourly prediction with 128 units in stateful mode overfitting occurs after one or two epochs as already seen with shopping center charging, leading to the extraordinarily high loss.

Table 6.4: Minimal validation losses for workplace charging when training the LSTM with the initial hyperparameter and different modes, encoding and timestep approaches

Mode	Input & output timesteps	encoding	val. loss* 2 units (epoch)	val. loss* 8 units (epoch)	val. loss* 32 units (epoch)	val. loss* 128 units (epoch)	Average validation loss*
state-ful	4in/ 4out	sine/cosine	1.8567 (92)	2.3629 (387)	2.3717 (74)	46.1219 (1)	13.1783
		one hot	2.6490 (367)	2.7348 (147)	2.3402 (482)	46.1219 (1)	13.4615
	96in/ 96out	sine/cosine	9.5155 (12)	10.5475 (436)	9.5745 (62)	9.4166 (29)	9.7635
		one hot	9.4798 (36))	9.0525 (452)	9.3157 (47)	9.3120 (21)	9.2900
state-less	4in/ 4out	sine/cosine	2.0662 (288)	2.2425 (373)	2.1991 (432)	2.2989 (126)	2.2017
		one hot	2.3023 (500)	2.3181 (499)	2.3654 (195)	2.3626 (177)	2.3371
	8in/ 4out	sine/cosine	2.2376 (248)	2.3163 (365)	2.1024 (311)	2.3109 (137)	2.2418
		one hot	2.5764 (142)	2.7807 (23)	2.3358 (242)	2.3332 (140)	2.5065
	16in/ 4out	sine/cosine	2.2033 (276)	2.0447 (277)	2.1184 (120)	1.8531 (221)	**2.0549**
		one hot	2.2995 (496)	2.3754 (165)	2.1553 (190)	1.9572 (132)	2.1968
	32in/ 4out	sine/cosine	2.1253 (226)	2.3055 (294)	2.4355 (19)	2.1572 (11)	2.2559
		one hot	2.4688 (500)	2.1787 (233)	1.9355 (222)	1.9591 (123)	2.1355
	96in/ 96out	sine/cosine	12.0278 (500)	9.6327 (342)	8.9209 (68)	9.2516 (61)	9.9583
		one hot	10.2221 (154)	9.2976 (123)	8.1748 (131)	6.9466 (130)	**8.6603**
	192in/ 96out	sine/cosine	10.6647 (398)	9.4571 (223)	8.3318 (179)	8.2848 (179)	9.1846
		one hot	10.5301 (424)	9.0731 (188)	8.5981 (65)	7.5280 (110)	8.9323

*All MSE values are given in units of 10^{-3}

When comparing the MSE of the two encoding variants, it becomes evident once again that in the majority of applications the sine/cosine encoding represents the superior choice for the hourly load prediction in both stateful and stateless mode. In contrast, one hot encoding yields more accurate results for the 1-day prediction in both modes. Finally, examining the validation loss values for the 4-timesteps and 96-timesteps prediction, it can be seen that in most cases workplace charging achieves lower validation loss values than shopping center, residential and public charging. Once again, for both forecast horizons the stateless mode accounts for the lowest loss. Sine/cosine encoding in combination with 16 input timesteps yields the most

precise results for the hourly prediction. For the 1-day forecast one hot encoding and 96 timesteps as input lead to the minimal validation loss.

6.1.5. Summary of initial training results

Table 6.5 summarizes the findings of the initial training by providing an overview of the LSTM configurations that lead to the lowest validation loss for all charging sites and different forecast horizons. The illustrated LSTM configurations are selected for the subsequent hyperparameter tuning.

Table 6.5: Summary of selected LSTM configuration for each charging site

Charging site	Number of input timesteps	Number of output timesteps	Mode	Encoding
Shopping center	16	4	Stateless	Sine/cosine
	96	96	Stateless	One hot
Residential	16	4	Stateless	Sine/cosine
	96	96	Stateless	Sine/cosine
Public	4	4	Stateless	Sine/cosine
	96	96	Stateless	Sine/cosine
Workplace	16	4	Stateless	Sine/cosine
	96	96	Stateless	One hot

For each charging site the prediction is performed in stateless mode. Moreover, sine/cosine encoding is selected in most cases. Only for the 1-day shopping center and workplace charging load prediction one hot encoding is chosen. The 1-day forecast for all charging sites is based on an input sequence consisting of 96 timesteps. Regarding the hourly load forecast, 16 timesteps are selected as input for shopping center, residential and workplace charging and four input timesteps are selected for public charging.

6.2. Hyperparameter tuning results

With the LSTM configuration selected in the previous section, random search is performed to determine the most appropriate combination of hyperparameters for each charging site and forecast horizon. The results of the 50 iterations of the random

search are displayed by means of a heatmap for the hourly and daily forecast for each charging site. For reasons of readability, the different heatmaps are provided in Appendix A.1. While the y-axis depicts the various batch size, learning rate and dropout combinations, the x-axis represents the different number of hidden layers and units. The minimum validation loss of each training process is presented in the color range between dark purple (high value) and dark green (low value).

6.2.1. Tuning results for shopping center charging

The outcome of the hyperparameter tuning for the 1-hour prediction for shopping center charging is given in Figure A.1. The MSE varies between roughly 0.0053 and 0.0337, which corresponds to a 2.5-fold prediction deterioration between the least and most suitable hyperparameter configuration. While the lowest loss is obtained by selecting one layer and 128 units, a batch size of 16 and a learning rate of 0.001 without using dropout, five distinct hyperparameter combinations lead to the highest loss. Strikingly, the highest MSE scores are always recorded when using the highest learning rate of 0.01. Compared to the lowest loss obtained during initial training (see 6.1) none of the hyperparameter configurations during random search yields an improvement.

The results for the 1-day load prediction can be seen in Figure A.2. Compared to the 1-hour prediction, the spread between the lowest (roughly 0.0088) and the highest validation loss (0.0135) decreases, indicating that the choice of hyperparameter exerts less effect on the prediction results. Moreover, it can be seen that a limited number of units or a large batch size diminishes the learning ability of the LSTM, which is reflected in the overall higher MSE values in those cases. Although the difference is marginal, the minimum loss achieved during random search represents an improvement compared to the previous seen minimum loss during initial training.

6.2.2. Tuning results for residential charging

Figure A.3 provides the tuning results for the hourly load prediction for residential charging. While the maximum MSE equals about 0.0102, the lowest loss amounts

to roughly 0.0028, thus reducing the prediction error by half in relation to real values in kW. The highest scores are generated when using 50 % dropout and 20 % dropout generally also yields higher MSE scores than using no dropout. As already observed with shopping center charging, hyperparameter tuning does not decrease the loss for the 1-hour prediction in residential charging compared to the minimum loss achieved during initial training.

When looking at the hyperparameter tuning results of the 1-day forecast, illustrated in Figure A.4, similar observations can be made compared to shopping center charging. The highest losses are recorded with a low number of units (2 or 4), indicating that a higher number of units is necessary to accomplish the complex prediction task. However, the batch size does not seem to have a significant impact. While the maximum MSE is about 0.021, the minimum loss amounts to about 0.008, marking a minimal improvement from the original minimum loss obtained during initial training.

6.2.3. Tuning results for public charging

The results of the hyperparameter tuning for the public charging hourly forecast, illustrated in Figure A.5, show similar characteristics as the findings for residential charging. The highest losses are caused by the use of 50 % dropout. In addition, the tuning does not result in improvements compared to the minimum loss during the initial training. The lowest loss stands at around 0.002 and rises to a maximum of around 0.0121. This result would imply a 2.4-fold deterioration of the prognosis in actual kilowatt figures, demonstrating that the selection of the correct hyperparameter configuration has a significant impact on the LSTM's performance.

The results of the hyperparameter tuning for the 1-day forecast, which can be seen in Figure A.6, exhibit similar characteristics to those of shopping center and residential charging as well. Compared to the initial training, the lowest loss can only slightly be reduced to around 0.0091. The maximum MSE, however, stands at around 0.0323. The large margin between both values again underlines the importance of hyperparameter tuning for the prediction accuracy.

6.2.4. Tuning results for workplace charging

Lastly, Figure A.7 and Figure A.8 exemplify the findings of the 4-timesteps and 96-timestep forecast respectively for charging at work. Looking at the results of the hourly prediction, similar constraints of the LSTM training using a high learning rate of 0.01 become apparent. Three different variations result in a maximum loss of around 0.0472. Compared to the lowest loss of around 0.0016, the prognosis in kW would be 5.4-times less accurate.

Looking at the results of the 96-timesteps forecast, the gap between minimum and maximum loss decreases significantly. The highest MSE is around 0.0145 and the lowest stands at around 0.0073. High MSE values are again obtained when using 2 units and in this case a small batch size in combination with a small number of units.

6.2.5. Summary of hyperparameter tuning results

Table 6.6 summarizes the selected hyperparameter configuration, the corresponding minimal validation loss during training and if the tuning yields an improvement over the initial training for each charging site and forecast horizon. As can be seen, the hyperparameter tuning only provides improvements for half of the forecasts. Furthermore, the most accurate results are solely achieved with a high number of units (64 or 128). For most cases, a LSTM with only one hidden layer seems sufficient. Only for the 1-day forecast for residential charging and 1-hour forecast for workplace charging the minimum loss is obtained by using two hidden layers. With regard to the batch size, none seems to be superior, all batch sizes are applied for the different forecasts. Moreover, dropout generally does not seem to have a large positive affect on the training, except for the 1-day residential charging load forecast where a 20 % dropout scores the lowest loss. Finally, in most cases the standard learning rate of 0.001 seems to achieve the most satisfactory results.

Table 6.6: Summary of tuning results and hyperparameter selection for the charging load forecast

Charging site	Forecast horizon	Improvement through tuning	Minimal loss in 10^{-3}	Optimal hyperparameter values				
				Number of layers	*Number of units*	*Batch size*	*Drop-out*	*Learning rate*
Shopping center	1 hour	No	5.2903	1	128	32	0	0.001
	1 day	Yes	8.7778	1	128	64	0	0.01
Resi-dential	1 hour	No	2.8316	1	128	32	0	0.001
	1 day	Yes	8.0033	2	128	16	0.2	0.001
Public	1 hour	No	2.0181	1	128	32	0	0.001
	1 day	Yes	9.0543	1	64	96	0	0.0001
Work-place	1 hour	Yes	1.5895	2	64	96	0	0.0001
	1 day	No	6.9466	1	128	32	0	0.001

6.3. Forecast results

For each charging site and forecast horizon, the LSTM trained on the previously outlined selected hyperparameters is used to forecast the charging load for the entire period of 36 days of test data. In the following the results for each charging site are first presented graphically. Both the real and the predicted load curves are displayed for a weekday, a day of the weekend and a public holiday. In addition, the forecasts for the entire period of 36 days are illustrated. In the end, the outcomes are compared by means of the metrics selected for the evaluation.

6.3.1. Forecast results for shopping center charging

To begin with, the prediction results of the three different days are exemplified in Figure 6.1. The upper plot illustrates the results for Monday, 16.12.19, the central plot provides the findings for Sunday, 08.12.19 and the lower plot visualizes the outcome for Finland's independence day (Friday, 06.12.19). Whereas the gray load curve depicts the actual load over the course of the day in kW, the green load curve indicates the forecasted load with the forecast horizon of one hour in kW and the purple load curve corresponds to the load forecast in kW with a forecast horizon of one day. The same layout is also used for the exemplary illustration of the forecast

results at the other charging sites.
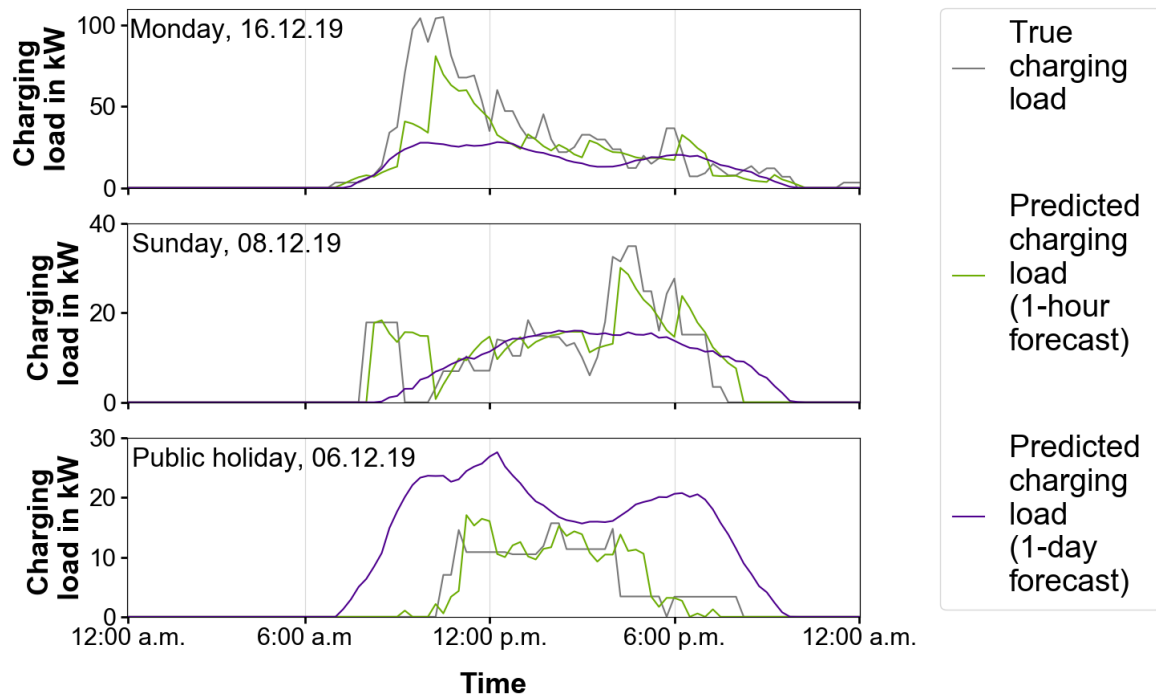
Figure 6.1: Exemplary illustration of the true and predicted charging load on a weekday, day
of the weekend and public holiday for shopping center charging

Looking at the graph, it becomes evident that the 96-timesteps forecast is not ca-
pable of predicting the strongly fluctuating load course in all three cases. While the
peak load for weekdays and weekends cannot be reproduced at all, the load curve
for the public holiday falls too high. In contrast, the load forecast with a forecast hori-
zon of one hour reflects the real load development more accurate. To some extent
a slight offset can be detected. However, the forecast provides a more accurate pic-
ture of the fluctuating structure of the load curve and peak loads are also much more
accurate. Furthermore, the sharp drop in load on public holidays can be predicted
more precisely by means of the hourly forecast.

Figure 6.2 illustrates the hourly and daily prediction results in the period from
26.11.19 to 31.01.19. In both plots the real charging load in kW is represented
in gray. The upper plot also depicts the hourly predicted charging load in kW in
green while in the lower plot the results of the 96-timesteps forecast is represented
in purple. Once again, the same format is also applied when representing the 36
days forecast results at the other charging sites in the subsequent sections.

Looking at the two load forecasts, the findings of the previously presented forecasts for three days appear to prove true. The hourly forecast seems to capture the daily varying load course relatively well. Peak loads are partly well predicted and also the predicted load curve on the public holidays seems to be fairly close to the real charging load. However, the load forecast for the Finnish Independence Day seems to be more accurate than for the Christmas holidays.
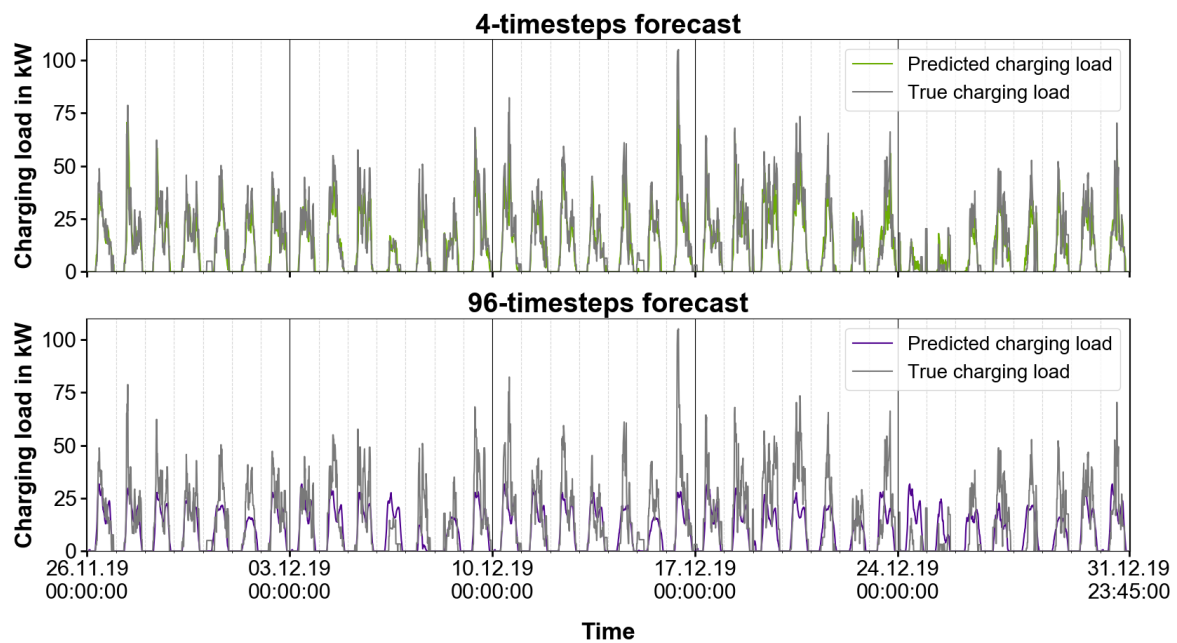
Figure 6.2: Comparison of real and predicted charging load for both forecast horizons for the 36 days of test data for shopping center charging

Moreover, a look at the lower graph confirms that the 96-timesteps forecast is not capable of predicting peak loads and the fluctuating load profile at all. When comparing the forecast results between weekdays and weekends, the forecast on weekdays generally seems to reflect at least the structure of the load curve more closely. The load forecast on weekends, on the other hand, is unable to provide a reliable prediction of the two peak loads that often occur. On most public holidays, the forecasted load is considerably higher than the real load. Only on the second day of Christmas the real load exceeds the predicted load.

6.3.2. Forecast results for residential charging

For residential charging, the forecasts are exemplified for Monday, 13.01.20, for Saturday, 04.01.20 and for Epiphany (Monday, 06.01.20) in Figure 6.3. Again, the hourly forecast leads to a more accurate prediction of the true load course than the 1-day forecast. The fluctuating structure can be represented relatively well by the hourly forecast close to the true load and the peak load is also predicted reasonably well for the Monday and Saturday in question. For the holiday, the predicted peak load is slightly higher. Compared to shopping center charging, however, the 96-timesteps forecast for residential charging is able to anticipate the general trend of the load profile in the three days considered. The forecast for the weekday and public holiday falls closer to the real load than on the weekend. Additionally, it is evident that the considered holiday does not strongly influence the regular load and thus the user behavior on this day. However, the fluctuations in the load curve can not be modeled by the daily forecast.

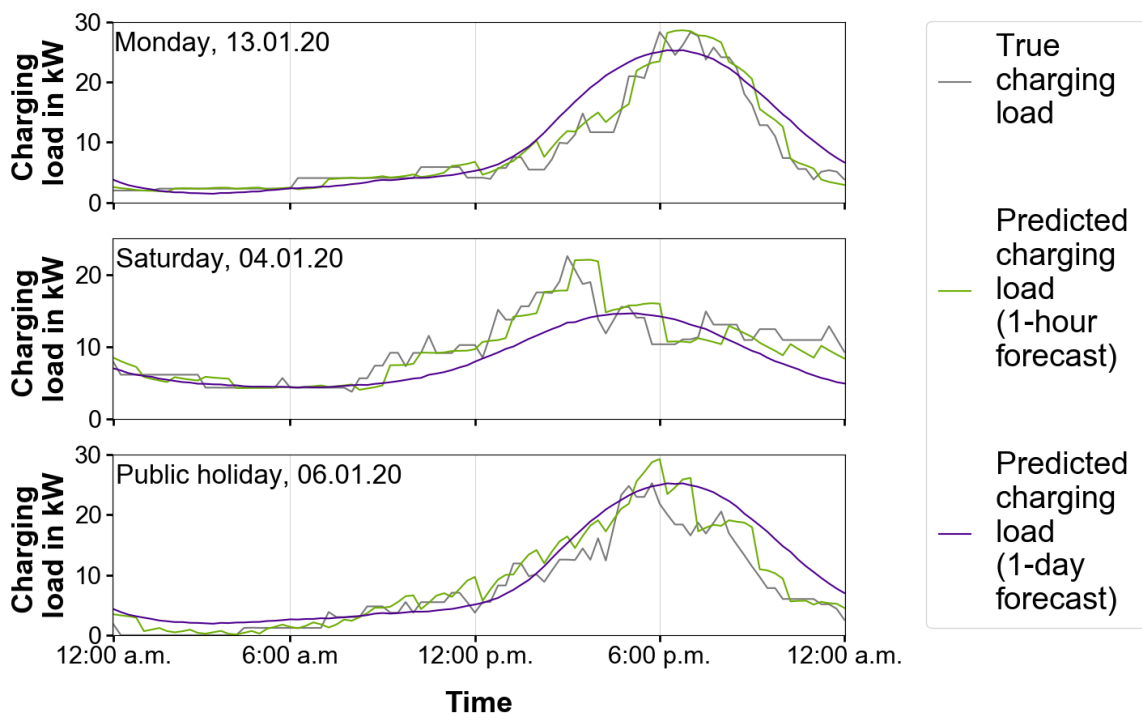Figure 6.3: Exemplary illustration of the true and predicted charging load on a weekday, day of the weekend and public holiday for residential charging

The outcomes of the charging load prediction in the period from 16.12.19 to 20.01.20 are visualized in Figure 6.4. Looking at the upper plot, it can be seen that the hourly load forecast, as indicated, provides a fairly thorough prediction of the true

load curve. The largest deviations are visible during the Christmas holidays. Fur-
thermore, difficulties are seen in forecasting the peak load especially in the period
between Christmas and Epiphany. More significant shortfalls can be seen within
the same period for the 96-timesteps forecast. Between Christmas and Epiphany,
the projected charging load is often significantly higher than the real charging load.
Furthermore, the tendency can be seen that the forecast on Saturdays tends to be
lower than the real load, while on Sundays the predicted load profile is overly high.
On the remaining days, the trend of the actual load can be predicted more precisely,
even if the predicted peak load is often lower than the real peak load and fluctuations
of the load can not be reproduced.
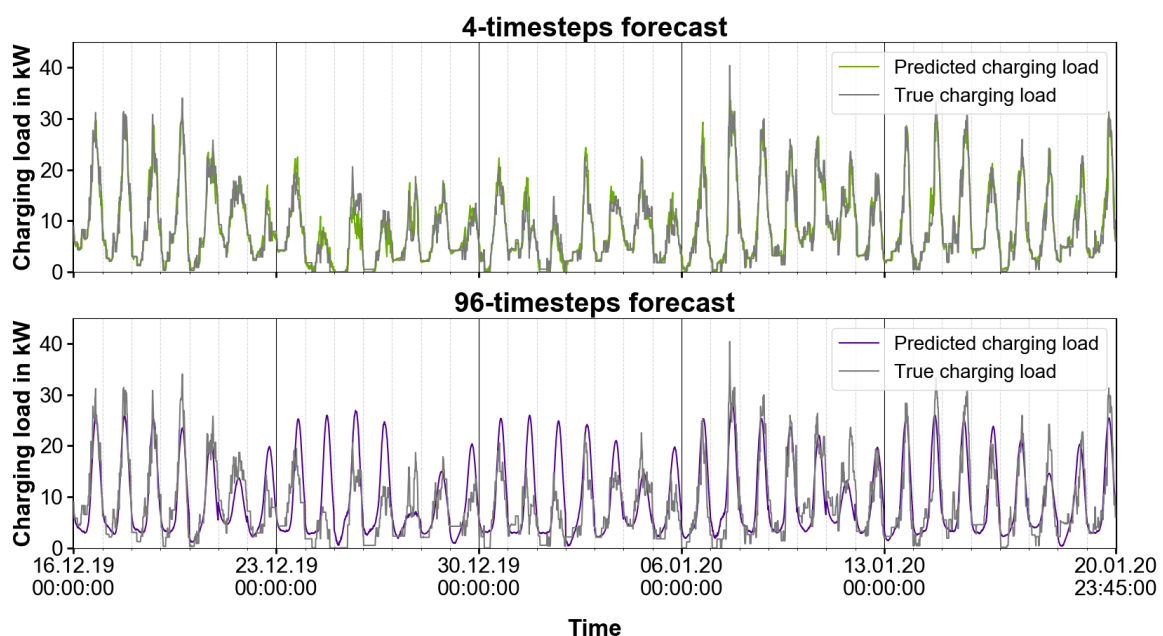
Figure 6.4: Comparison of real and predicted charging load for both forecast horizons for the
36 days of test data for residential charging

6.3.3. Forecast results for public charging

For public charging Figure 6.5 displays the forecast results for Thursday, 16.01.20,
for Saturday, 25.01.20 and for the second Christmas holiday (Thursday, 26.12.19).
Comparing the forecast results for public charging, it becomes evident once again
that reducing the forecast horizon from 96 to 4 timesteps enhances the accuracy of
the forecast. This outcome becomes particularly clear when considering the lower
plot, where the hourly prediction fails to provide an adequate prediction of the load
curve, while the hourly forecast leads to a significant reduction of the prediction error.

For the exemplary day of the week though, the 1-day forecast provides an remarkably realistic picture of the actual load profile and, compared with the results shown for shopping center and residential charging, achieves the most convincing results to date. In addition, the 1-day forecast provides a relatively realistic representation of the real load profile on Saturday as well.

Figure 6.5: Exemplary illustration of the true and predicted charging load on a weekday, day of the weekend and public holiday for public charging

The forecast results over the entire period of 36 days, starting from 26.12.19 and ending on 30.01.20, are portrayed in Figure 6.6. The hourly load forecast seems to be able to predict the load profile for weekdays, weekends and holidays in general fairly accurately. No major outliers are discernible, and the daily peak loads are well captured by the daily forecast. The only striking observation is that for public holidays and partly weekends the forecast does not correctly predict the charging load during the first hours of the day. In such events, the predicted load curve drops to zero while the real load curve remains at a certain positive level most of the time. Looking at the 1-day forecast results, a number of observations can be made. First, as shown in the daily plot before, the 96-timesteps forecast is not able to predict the charging load of the public holidays. For each of the three public holidays, the predicted load

curve significantly exceeds the real load curve. Furthermore, the LSTM struggles to predict the load pattern on weekends. Although, as depicted above, the predicted load profile reflects the real load profile relatively well in a few cases on Saturday, the 1-day forecast is not capable of accurately forecasting the load profile on Sunday. Moreover, it can be seen that on most days the predicted and real peak load diverge significantly.

Figure 6.6: Comparison of real and predicted charging load for both forecast horizons for the 36 days of test data for public charging

6.3.4. Forecast results for workplace charging

Finally, the prediction results for workplace charging are examined as well. Figure 6.7 presents the exemplary results for Friday, 24.01.20, for Saturday, 25.01.20 and the second Christmas holiday (Thursday, 26.12.19). Looking at the upper plot, it can be seen that the hourly forecast provides a very good forecast of the real load profile. The 1-day forecast also shows the structure of the load curve well, but prematurely. The 1-hour forecast is also capable of reproducing the Saturday load pattern well, albeit with a small offset. Conversely, the 1-day forecast cannot be used to forecast the load on Saturday, where the forecast for the entire period amounts to 0 kW. Furthermore, only the hourly forecast is able to provide an appropriate prediction of the load pattern for the public holiday at hand.

These exemplary observations are reinforced by looking at the entire forecast in the period from 26.12.19 to 30.01.20, which is shown in Figure 6.8. The prediction with a horizon of one hour once again provides a good estimate of the real load curve. Difficulties mainly concern the correct prediction of the peak load.
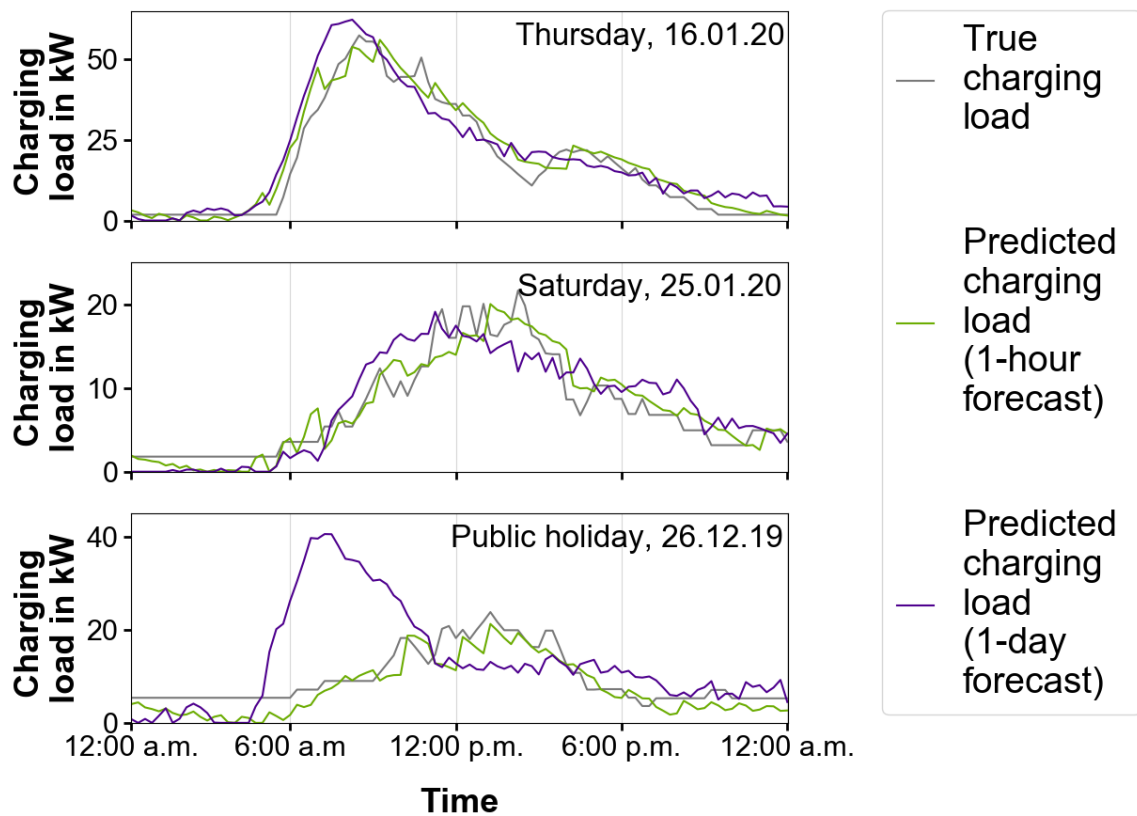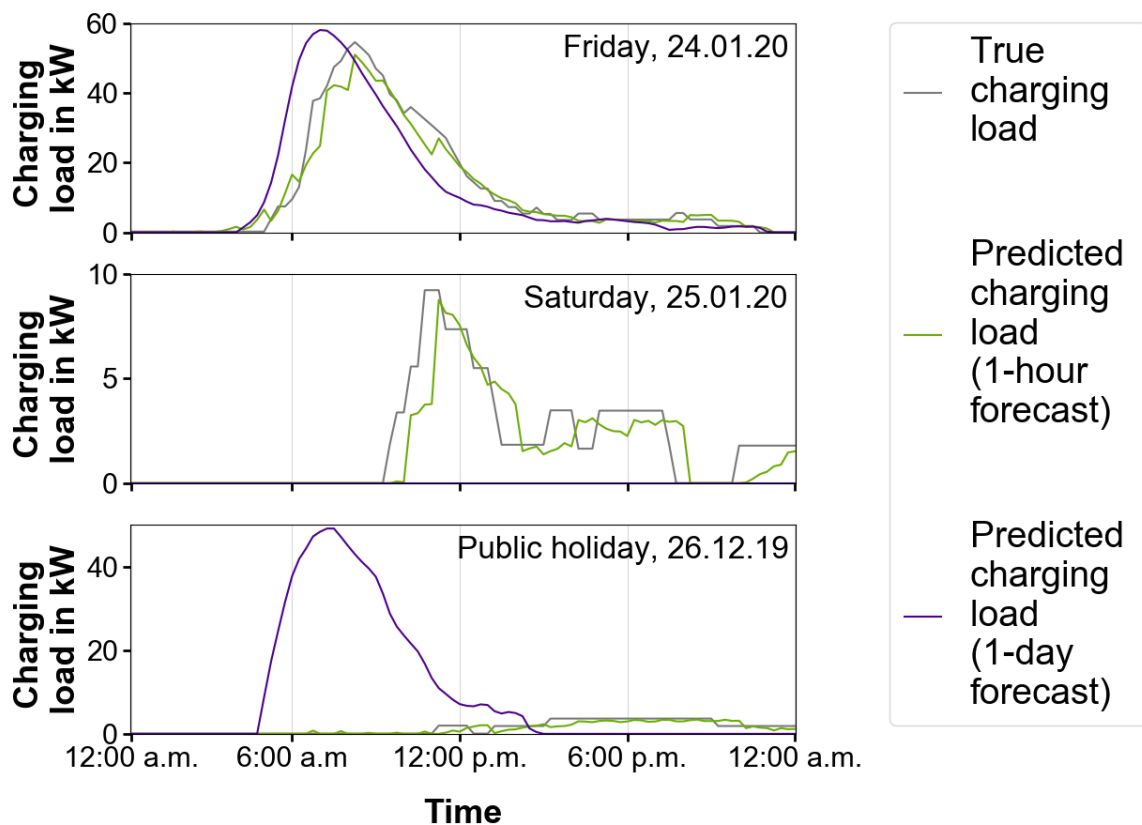

Figure 6.7: Exemplary illustration of the true and predicted charging load on a weekday, day of the weekend and public holiday for workplace charging

When looking at the lower plot, it is evident that the previously shown results of the 1-day charging load forecast on Saturday do not represent an exceptional event. A charging load of 0 kW is predicted for all weekends over the entire period. The poor prediction results on holidays are also confirmed when looking at the forecasts for the 26.12.19, 01.01.20 and 06.01.20. In all cases, the predicted load significantly exceeds the true load curve and the pattern of the load curve cannot be reproduced at all. For all other days, the predicted load curve provides a fairly realistic representation of the structure of the real course of events. However, the true height of the peak load can only be predicted properly in a limited number of cases. A particularly large spread between predicted and real peak loads can be seen on Tuesday and Wednesday of each week.
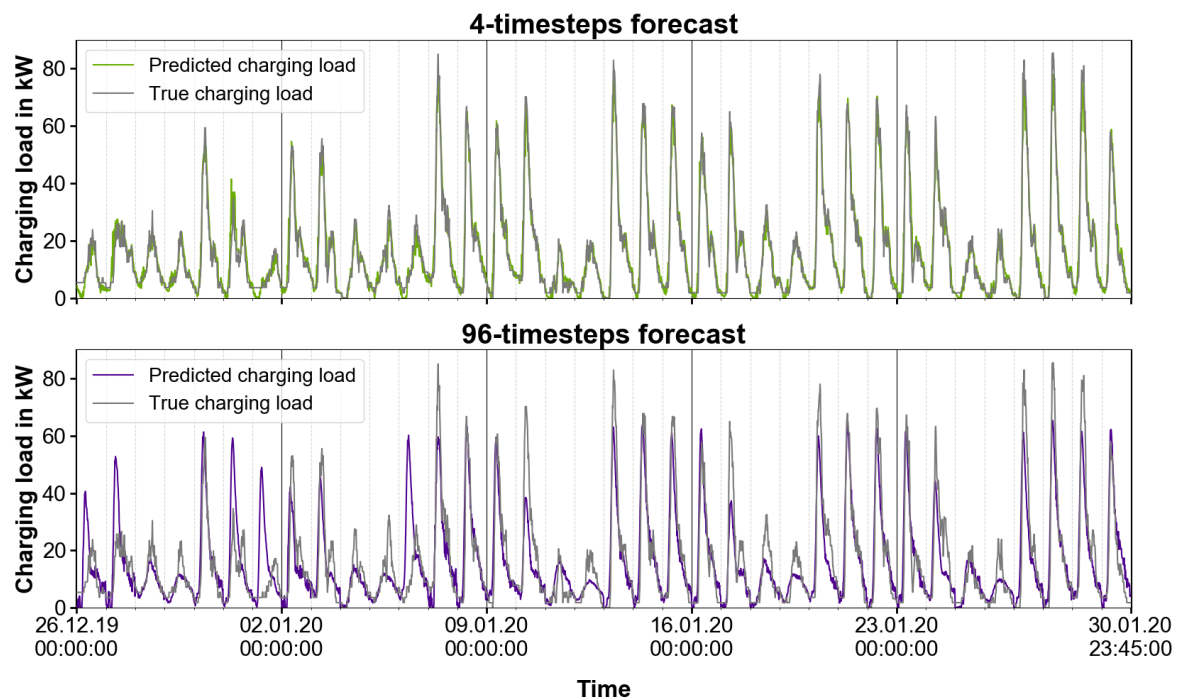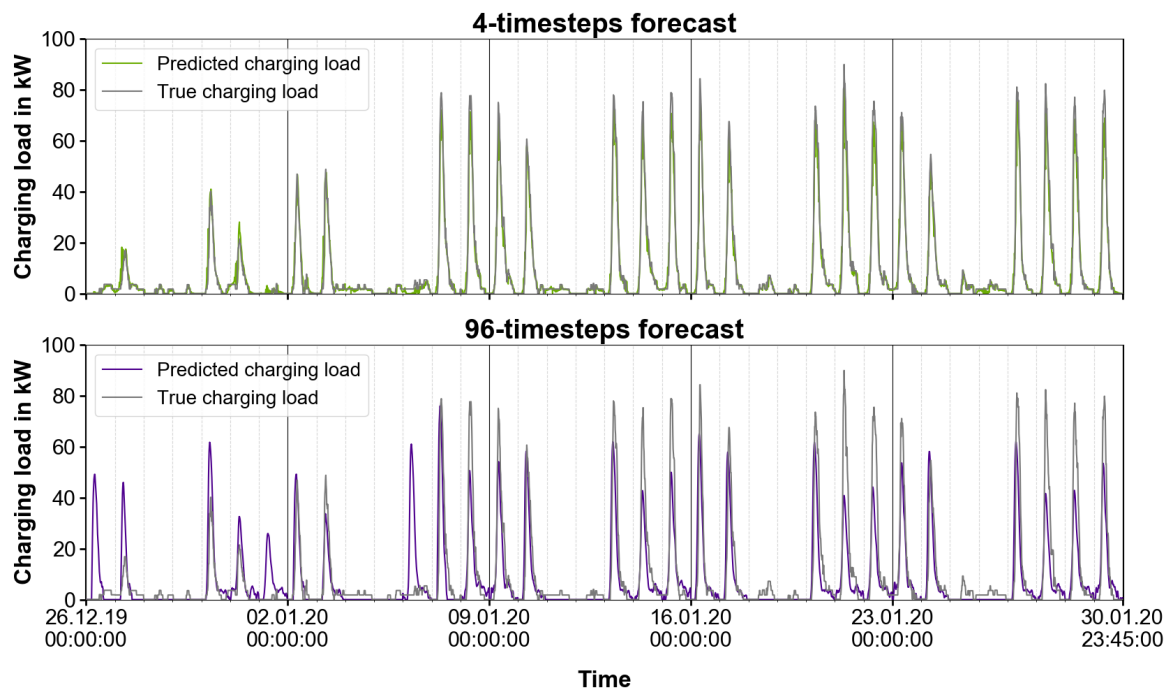
Figure 6.8: Comparison of real and predicted charging load for both forecast horizons for the 36 days of test data for workplace charging

6.3.5. Comparison of forecast results based on error metrics

Having graphically analyzed the prediction results for each charging site separately, the prediction outcomes will now be compared and evaluated according to the selected metrics. First, the results of the MAE, NMAE1 and NMAE2 are displayed for the entire forecast period and broken down by categories of weekday, weekend and public holiday. Subsequently, the peak and corresponding time deviation and the MAPE of the peak load prediction is analyzed for the 1-day load forecast. Finally, the MAE is also calculated separately for each 15-min interval of a day.

Overall, weekday, weekend and public holiday MAE, NMAE1 and NMAE2 results
Table 6.7 summarizes the MAE and different NMAE for all charging sites and the respective forecast horizon. As previously evidenced during the graphical analysis, reducing the forecast horizon from one day to one hour considerably reduces the prediction error. When considering the total MAE as well as the MAE on weekdays and weekends for residential, public and workplace charging, a reduction by at least half can be observed in almost all cases. The error reduction is even more pronounced when looking at the MAE on public holidays, where the MAE is reduced

by more than four times for public charging and by almost 11 times for workplace charging. With shopping center charging, the MAE is reduced in all cases by more than 2 kW. The lowest MAE for both forecast horizons is recorded for residential charging and amounts to around 1.53 kW and 3.37 kW for 1-hour and 1-day forecasting respectively. Workplace charging generates the second lowest MAE values, followed by public charging. The highest MAE values are recorded for shopping center charging.

Table 6.7: Summary of MAE, NMAE1 and NMAE2 results for all charging sites

Charging site	Shopping center		Residential		Public		Workplace	
Forecast horizon	1 hour	1 day	1 hour	1 day	1 hour	1 day	1 hour	1 day
MAE [kW] (all)	4.35	6.78	1.53	3.37	2.70	5.84	1.85	5.35
NMAE1 (all)	0.3521	0.5490	0.1775	0.3910	0.1657	0.3591	0.2038	0.5902
NMAE2 (all)	0.0414	0.0645	0.0379	0.0834	0.0316	0.0684	0.0206	0.0595
MAE [kW] (weekday)	4.41	6.75	1.52	3.33	3.10	6.37	2.56	6.56
NMAE1 (weekday)	0.3686	0.5639	0.1745	0.3810	0.1519	0.3120	0.1908	0.4889
NMAE2 (weekday)	0.0420	0.0643	0.0377	0.0823	0.0363	0.0745	0.0285	0.0730
MAE [kW] (weekend)	4.39	7.15	1.60	3.62	1.83	3.20	0.49	1.31
NMAE1 (weekend)	0.3289	0.5356	0.1894	0.4283	0.2084	0.3634	0.3767	1.00
NMAE2 (weekend)	0.0598	0.0974	0.0470	0.1064	0.0563	0.0982	0.0536	0.1422
MAE [kW] (public holiday)	3.71	5.74	1.36	2.89	2.48	10.64	0.88	9.51
NMAE1 (public holiday)	0.3122	0.4835	0.1619	0.3454	0.2618	1.1240	0.5952	6.4538
NMAE2 (public holiday)	0.0729	0.1128	0.0466	0.0994	0.0915	0.3929	0.1604	1.7389

A different picture emerges when looking at the NMAE1 and NMAE2 results. For the hourly prediction, according to the NMAE1, the most accurate forecast is obtained for public charging, followed by residential, workplace and shopping center charging. Conversely, using the NMAE2, the most precise results are seen for workplace charging, followed by public and residential charging. Shopping center charging again achieves the least accurate ratings. For the 96-timesteps forecast the most accurate forecast results are achieved for public and residential charging again when using NMAE1, however, the least accurate prognosis is given for workplace charging this time. In contrast, using the NMAE2, the most precise results are achieved for workplace charging, followed by shopping center and public charging. Residential charging comes last.

When comparing the clustered NMAE1 results, it can be shown that on weekdays,

weekends and public holidays the 1-hour forecast yields the most accurate results for public and residential charging. Shopping center and workplace charging score the poorest results. Conversely, when using the NMAE2 as an evaluation criterion, varying results are visible depending on the type of day. For weekdays the most accurate results are seen for workplace charging followed by public and residential charging. Shopping center charging generates the least accurate results. While residential charging provides the most exact forecast for weekends, workplace, public and shopping center charging take second, third and fourth place. For public holidays residential charging ranks first again, followed by shopping center and public charging. Workplace charging comes last. Similar findings to the NMAE1 results of the 1-hour forecast can be seen for the NMAE1 results of the 1-day forecast. Residential and public charging achieve the most precise scores in most of the cases. However, for public holidays, shopping center charging outperforms public charging and comes in second place behind residential charging. Looking at the NMAE2 results, the poorest performance for weekdays can be observed for residential charging while shopping center charging yields the most accurate results. On weekends shopping center charging ranks first again, however, the least accurate score is obtained for workplace charging. For public holidays, workplace charging performs particularly poor again. Residential charging obtains the most favorable score.

MAE results for the different 15-min timeslots of a day
Next, the prediction results are analyzed separately for each of the 96 quarter hour intervals of a day. Table A.1 details the MAE for each 15-min interval at each charging site and for both prediction horizons. The first value corresponds to the MAE of the 1-hour forecast in kW and the second value represents the MAE of the 1-day forecast in kW. In addition, six 4-hour averages of the MAE are provided.

To start with, it can be observed once again that the 1-hour forecast can significantly reduce the MAE compared to the 1-day forecast. A higher MAE is recorded only in a few exceptional cases. The extent of the improvement, however, varies considerably over the course of the day. Furthermore, when looking at the MAE results of the 1-hour forecast for the individual 15-min intervals, a deterioration of the forecast results with increasing forecast period can be observed. Generally, the prediction of the first 15-min timestep yields a lower MAE than the second step, the second step generates a lower MAE than the third step, and the least accurate results are generally achieved for the fourth timestep. Consequently, the examined

time of the hourly forecast influences the extent of the improvement, which can be achieved compared to the 1-day forecast.

Taking a look at the MAE values in the period from 10:00 a.m. to 11:00 a.m, for example, the MAE for shopping center charging rises from 3.63 kW in the first timestep to 10.74 kW in the 4th timestep. Similar behavior can be shown for the other three charging sites, where the MAE for residential charging increases from 0.81 kW to 1.6 kW, for public charging from 2.37 kW to 4.65 kW and for workplace charging from 2.29 kW to 3.81 kW. While from 10:00 a.m. to 10:15 a.m. the MAE is therefore lowered by 8.48 kW for shopping center charging, this figure drops to 3.95 kW in the period from 10:45 a.m. to 11:00 a.m.. For residential charging the MAE is reduced by 1.61 kW respectively 0.56 kW. For public and workplace charging, the reduction in MAE also falls from 11.28 kW to 8 kW and from 14.6 kW to 11 kW, corresponding to a decrease of around 3 kW in both cases.

Lastly, it can also be seen that the prediction accuracy at the various charging sites is strongly dependent on the time of day and correlates with the load level. With shopping center charging, the lowest average MAE is achieved for both forecast horizons between 12:00 a.m. and 4:00 a.m.. The lowest MAE values generally occur between 3:30 a.m. to 7:30 a.m.. The highest average MAE can be seen in the timeslot 12:00 p.m. to 16:00 p.m., but high MAE values are generally achieved throughout the entire period from 9:00 a.m. to 8.00 p.m.. For residential charging, the lowest MAE values are also obtained in the period from 12:00 a.m. to 8:00 a.m. and the highest in the time between 4:00 p.m. to 8:00 p.m., which again shows the correlation between the load level and the prediction error. In the case of public and workplace charging, the lowest MAEs are likewise observed in the period from 12:00 a.m. to 4:00 a.m., while low values are also recorded from 8:00 p.m. to 12:00 a.m.. The highest average MAE values are obtained for the 1-hour forecast in the period 4:00 a.m. to 8:00 a.m. and for the 1-day forecast in the period 8:00 a.m. to 12:00 p.m..

Load deviation, time deviation and MAPE results of the daily forecast

To conclude, special attention is given to the 1-day forecast results by reviewing the peak load deviation in absolute and relative terms as well as the average deviation of the peak load time. Table 6.8 lists the peak load deviations and MAPE results in kW and % respectively as well as the average deviation between the time of the true and predicted peak load in 15-min timeslots for all charging sites.

Table 6.8: Summary of the peak deviation, MAPE and corresponding time deviation for each charging site

Charging site	Shopping center			Residential			Public			Workplace		
Metric	Peak dev. (kW)	MAPE_ peak (%)	Time dev. (nslot)	Peak dev. (kW)	MAPE _peak (%)	Time dev. (nslot)	Peak dev. (kW)	MAPE_ peak (%)	Time dev. (nslot)	Peak dev. (kW)	MAPE_ peak (%)	Time dev. (nslot)
All days	29.16	53.19	10.17	6.37	34.38	5.86	14.17	37.16	6.92	17.85	128.97	19.00
Weekdays	30.96	53.65	12.43	6.24	34.71	5.74	13.75	23.50	3.48	20.54	35.50	2.30
Weekends	26.37	53.11	6.30	6.29	32.00	6.60	11.39	43.28	7.70	4.73	100.00	-
Public holidays	24.65	49.90	5.67	7.64	39.80	4.33	26.62	121.46	30.67	41.01	942.13	29.00

Over the period of 36 days, the average deviation between true and predicted peak load amounts to 29.16 kW for shopping center charging, 6.37 kW for residential charging, 14.17 kW for public charging and 17.85 kW for workplace charging. In terms of percentages, the MAPE amounts to about 53.19 % for shopping center charging, about 34.38 % for residential charging, about 37.16 % for public charging and 128.97 % for workplace charging. The average deviation between real and predicted time of peak load expressed in number of 15-min timeslots is about 10, 6, 7 and 19 for shopping center, residential, public and workplace charging.

Examining the scores grouped by weekdays, weekends and public holidays, a number of discrepancies between the charging sites can be highlighted. Regarding shopping center charging, it can be seen that on weekends and public holidays, the absolute and relative errors can be reduced relative to weekdays. In addition, the deviation between true and predicted occurrence of peak load can be significantly lowered as well. For residential charging, in contrast, the highest relative and absolute errors are now seen for public holidays, although the time deviation is minimized. For public and workplace charging it is clearly visible that by far the most precise results are achieved during the week. For workplace charging predictions on weekends, the absolute error amounts to 4.73 and MAPE to 100 % due to the already mentioned constant prediction of 0 kW. Therefore, no deviation between the time of the true and predicted peak load can be calculated. The highest error scores between real and predicted peak load are recorded on public holidays for both charging sites and the time deviation between true and predicted peak load amounts to a multiple of the deviation during the week.

7. Discussion of results

The analysis is conducted accordingly to the presentation of the results in the previous chapter and is structured in three sections. First the results of the initial training are evaluated, followed by the analysis of the hyperparameter tuning. Subsequently, a detailed assessment of the prediction results is provided.

7.1. Analysis of initial training results

When looking at the results of the initial training, three key findings emerge. To begin with, in the majority of cases the problem of early overfitting with an increased number of units is clearly evident. Such behavior indicates that the complexity of the model may be too high relative to the size of the dataset available. Within this work, no more than one year's data is available, of which 78 % is used for training the LSTM. The availability of charging data over a period of several years might therefore be of particular importance for the performance of the model. The availability of charging date measured over a longer period of time would also facilitate the use of a larger dataset for validation, ensuring that the validation data set is truly representative of the entire dataset. Looking at the load curve analyzed in Chapter 5.2.3, it seems questionable whether the validation data used are representative for the whole time series due to the limited size of the sample.

Next, it could be demonstrated that the different encoding approaches and input timestep variants only cause negligible differences in the validation loss but that the optimal choice is dependent on the charging site and forecast horizon. Given the minor improvements in validation loss that can be achieved, the additional time and complexity required to implement the various encoding and input timestep variants cannot be justified. Hence, the same number of input timesteps and encoding choice for all charging sites seems to be more reasonable. While sine/cosine encoding provides a viable encoding variant, a number of 16 input timesteps for the 1-hour prediction and 96 input timesteps for the 1-day prediction seem to be appropriate.

Ultimately, the initial training results reveal another valuable insight. Under the given

conditions the stateful mode does not provide benefits compared to the stateless mode. Therefore, the charging load seems to rather be correlated with short term-events than long term-events in the past. Since the strength of the LSTM is therefore not needed to map long-term relationships, the prediction task could possibly have been performed by a simpler ANN as well. However, when multi-year data becomes available, the stateful mode may reveal its strengths, as seasonal relationships and long-term dependencies can be identified by the LSTM. Consequently, with the availability of data over several years, a further comparison of the different modes is advisable.

7.2. Analysis of hyperparameter tuning results

With regard to the hyperparameter tuning results, three relevant conclusions can be drawn. First of all, the often significant variation between minimum and maximum validation loss during tuning reveals that the different hyperparameter configurations exert a major influence on the generalizability of the LSTM model and, ultimately, on its predictive power. Thus, the importance of carefully selecting the hyperparameter is demonstrated and the relevance of hyperparameter tuning for selecting the most appropriate LSTM model is outlined.

Moreover, it could be shown, that in general for hourly predictions, a high dropout rate or a high learning rate can lead to particularly high MSE values. For the 1-day forecast a small number of units does not seem to be able to accurately map the complexity of the forecast, leading to the highest validation losses. Despite these general observations, it is also evident that the varying combinations of hyperparameters exert a fundamentally different influence on the validation loss for each charging site and forecast horizon in most cases. Thus, it is essential to perform the hyperparameter tuning separately for each use case.

Finally, for all charging sites the analyzed hyperparameter configurations by random search show only minor or no improvement compared to the minimal loss during the initial training. A variety of conclusions can be made in this respect. First, the findings reveal that the originally chosen hyperparameters proved to be a fairly good choice. At the same time, only 50 variations of random search were tested within this

work due to time constraints. A higher number of test runs eventually would have led to further improvements of the validation loss. However, taking into account the lowest validation losses for each charging site and forecast horizon, the result suggests that an increase in test runs might not lead to substantial improvements in the validation loss. An extension of the search space by varying more hyperparameter might result in a more conspicuous change of the validation loss.

7.3. Analysis of forecast results

Finally, the analysis of the forecast results is carried out in three stages. First, the results of the 1-day load forecast are analyzed. Next, a discussion of the 1-hour prediction results is carried out. Ultimately, several implications for the main stakeholders are derived on how to use the charging load forecast in practice.

7.3.1. Discussion of 96-timesteps forecast results

In this section the results of the 1-day load forecast are reviewed. At first, the discrepancies between the results obtained for the individual charging sites are addressed. Subsequently, the overall findings are summarized and potential explanatory approaches are provided.

Main differences between the charging sites
To begin with, differences in the ranking of the charging sites become apparent with regard to the NMAE1 and NMAE2 results. The poor performance of residential charging using NMAE2 compared to NMAE1 can be attributed to the significantly lower peak load of around 41.1 kW compared to the other charging sites. Although shopping center charging exhibits the highest peak load of around 105.1 kW and second highest mean charging load of around 12.34 kW, it occupies the penultimate place for the NMAE1 and also ranks second based on the NMAE2. This circumstance can possibly be attributed to the high complexity of the load profile, showing the most irregular load profile compared to the other charging locations as shown in Section 5.2.3. The large gap in the ranking of workplace charging when using

NMAE1 and NMAE2 can be attributed to the relatively high peak load of around 89.87 kW and low mean charging load of 9.07 kW. Compared to the similar mean charging load amounting to 8.62 kW for residential charging, workplace charging achieves significant less accurate forecast results.

Further evident discrepancies involve the varying NMAE1 results for weekdays, weekends and holidays between the different charging sites. While the most accurate results are obtained for shopping center and residential charging on public holidays, contrary observations emerge for public and workplace charging. The forecasts on public holidays are significantly less accurate than the predictions on weekdays and weekends. However, the differences are attributable to the load profile characteristics shown in Section 5.2.3 rather than to the different ability of the LSTM to identify public holidays. While the load profile for workplace and public charging shifts radically on public holidays, this behavior can not be seen for residential charging. With shopping center charging the load profile also varies noticeably, but the entire load series is characterized by strongly fluctuating load characteristics. Therefore the impact of public holidays remains small.

Similar findings can be seen when looking at the peak deviation and MAPE results of the individual charging sites. While for shopping center and residential charging the MAPE difference between weekdays and public holidays amounts to only 3.75 respectively 5.09 percentage points, the figure for public and workplace charging rises by 97.96 respectively 906.63 percentage points. Once again, the discrepancies can be explained by the already discussed variations regarding the general characteristics of the load profile and the impact of public holidays on the load pattern. When looking at the weekday results, the significant discrepancies in MAPE and time deviation for each charging site likewise indicate that the characteristics of the load profile have a decisive influence on the accuracy of the prediction. Shopping center charging has a highly irregular load profile, which is why both the predicted peak load level and the time of the peak load display the poorest accuracy. Workplace charging, on the other hand, as shown in Section 5.2.3, exhibits the most steady pattern with regard to the time of peak load, which is why the average time deviation is lowest at 35 minutes.

Overall findings
Overall, two important outcomes can be outlined. The MAE results of the different

timesteps reveal that the prediction error in times of low load is smaller than in times of high load and therefore correlates with the the load patterns. The reason for this outcome can be naturally explained by the fact that during periods of high demand, the fluctuation in charging load also tend to increase, therefore making it difficult for the LSTM to reproduce the correct load level. In times of no or very low charging load, the MAE is consequently lower.

More importantly, the findings of the 1-day prediction results indicate that the LSTM is not able to provide a reliable forecast given the complex load characteristics and high resolution of 15-min. There are several explanations for the lack of forecasting ability of the 96-timesteps prediction.

It is conceivable that the sharp increase in aggregated charging load over the course of the year, illustrated in Section 5.2.3, and the resulting altered load profile prevents the LSTM from learning and thus leads to poor prognosis results. Due to the limited amount of data spanning only one year, the training and validation data possess insufficient or unrepresentative data to allow the LSTM to establish the correlation between the inputs and outputs as present in the test data, which was already indicated in Section 7.1.

A similar problem is posed by the large number of holidays in the test data occurring in a sequence not previously seen by the LSTM, resulting in the LSTM not being able to locate the public holidays. In addition, the changes in the load characteristics between Christmas and Epiphany, demonstrated in Section 5.2.3 in particular for residential, workplace and public charging, caused by the winter holidays, might explain the frequently large prediction errors in those days.

7.3.2.   Discussion of 4-timesteps forecast results

In the following the results of the 1-hour forecast are reviewed as well. As for the 1-day forecast, the different results among the charging sites are discussed first before a summary of the overall findings is provided.

Differences between charging sites

Looking at the NMAE1 and NMAE2 results of the 1-hour forecast two main findings emerge. On the one hand, the frequently poorly obtained results for shopping

center charging show that even with a shorter forecast period, a strongly fluctuating load curve without a clear pattern poses difficulties for the LSTM and leads to poor prediction results. On the other hand, weekends and public holidays seem to yield less accurate predictions particularly for public and workplace charging, although significantly less than seen with the 96-timesteps forecast. This circumstance can be explained by the fact that due to the shorter forecast period, the preceding load values assist the LSTM in mapping the height and course of the load curve more precisely.

Overall findings and explanatory approaches
Similar to the 1-day forecast, two general statements can be derived for the 1-hour load forecast, applicable to all charging sites. It can be pointed out that by reducing the forecast period from 96 to four timesteps, the forecast accuracy can be significantly increased for all charging sites. The impact of public holidays and weekends on the forecast accuracy can also be significantly reduced. However, the MAE results of the individual time steps still show a noticeable increase in the prediction error with increasing time deviation from the last input load value.

7.3.3. Implications for the practical application of the forecast

Several implications for sub-aggregators and network operators can be derived concerning the application of the charging load forecast in practice, which are discussed in the following.

Implications for sub-aggregator
To begin with, it is imperative to note that, from the point of view of a sub-aggregator, the forecasting accuracy achieved by the 1-day load prediction is far from sufficient to predict flexibility potentials for the provision of CR. In order to improve the poor forecast result obtained by the 96-timesteps forecast more data is needed. The dataset needs to cover a much longer period of time with the training and evaluation data containing representative data that exhibit comparable properties to the test data. It can be assumed that the prediction accuracy improves if charging data of the same period as the test data from previous years are also present in the training data. Furthermore, a higher forecast accuracy can be achieved by either shortening

the forecast horizon or prolonging the time resolution of the forecast.

The results of the 1-hour forecast show that considerably more accurate predictions can be achieved compared to the 1-day forecast. Furthermore, as shown in Section 5.2.3, the load profiles of residential, public and workplace charging indicate that the highest load level usually falls within a certain time period of the day. However, the identified patterns are mainly valid for weekdays. Weekends and public holidays are often accompanied by changes in the user behavior that are difficult to predict and also exhibit a much lower aggregation potential, which is particularly evident for workplace and public charging.

The above mentioned circumstances suggest two courses of action on how sub-aggregators can implement the proposed LSTM within this thesis to achieve optimized prediction results. On the one hand, based on the aforementioned reasons, it seems beneficial to limit the charging load forecast to weekdays. It is likely that by focusing the forecast on weekdays a higher prediction accuracy can be obtained. On the other hand, the forecast should only be carried out for the most relevant periods of a day, when the aggregated charging load is highest. By shortening the forecast period, the results are improved as shown in this paper. Furthermore, it can be anticipated that the accuracy of the forecast will be further enhanced by focusing the LSTM on a specific time period of a day. These two measures would assist the sub-aggregator to provide a more reliable load forecast in the important periods of a day to be able to exploit the huge potential of PEV aggregation for CR provision as outlined in Chapter 3.

Implications for network operators

Similar recommendations for action can be formulated for the network operator. Under the current circumstances, times of high charging loads and peak loads, which are particularly important for network operators to prevent possible bottlenecks in the low voltage network caused by the simultaneous charging of PEVs, cannot be reliably predicted with the 1-day load forecast. The forecast could be targeted towards the time range of the highest aggregated load at each charging site, where the simultaneous charging is likely to result in the most severe bottlenecks in the low voltage network. Thus, the forecast could still be used for congestion management. In addition, the temporal resolution could be reduced down to, for example, one hour to enhance the accuracy of the forecast.

8. Conclusions and future work

The following chapter summarizes the contribution and main findings of this work and provides an outlook on possible focal points for future research.

8.1. Summary of the work

This thesis aims at developing an ANN for predicting the aggregated charging load for different charging sites clustered according to distinct behavioral patterns. The charging at home (residential), the charging at work (workplace), the charging at public car parks (public) and the charging at shopping centers (REDI) are chosen for the analysis. Moreover, two different forecast horizons are selected in the form of a 1-hour and a 1-day charging load forecast with a 15-min resolution and implemented for every charging site. A special form of RNNs, the LSTM, is applied within this thesis due to the good performance demonstrated for time series in the past. Based on the research shortcomings pointed out by the literature review in Chapter 2.4 it can be demonstrated that the multivariate multi-step LSTM charging load forecast presented in this work introduces a new approach within the literature. Two main contributions of this work can be highlighted. The two main achievements of this thesis comprise the analysis of two forecast horizons, which by far exceed the prediction horizons reported in the literature so far, and the comparison of different charging sites based on both real and actual charging load data, providing another valuable contribution to the poorly researched field of PEV charging load forecasting.

8.2. Findings of the work

In the following the most important findings of the charging load prediction are summarized. First and foremost, the forecast results indicate that the evaluation of the results is highly dependent on the chosen metrics. Based on the MAE, for both forecast horizons the smallest error is observed for residential charging, followed by workplace and public charging. The highest MAE values are seen for shopping

center charging. To eliminate the scale dependency, the MAE results are normalized using the mean charging load (NMAE1) or difference between maximum and minimum charging load (NMAE2). Applying the NMAE1, the most precise results are achieved at both forecast horizons for public charging followed by residential charging. For hourly forecasting the least accurate results are obtained for shopping center charging and for 1-day forecasting for workplace charging. In contrast, when using the NMAE2 as measure of evaluation, the most accurate prediction results for both forecast horizons are obtained for workplace charging. While public charging ranks second for 1-hour predictions, shopping center charging comes second for the 1-day prediction. The least accurate results for hourly predictions are seen for shopping center charging and with the daily prediction for residential charging.

Furthermore, it is demonstrated that the forecast accuracy strongly correlates with the respective forecast horizon. The 1-hour load forecast achieves substantially better results than the 1-day load prediction. The results of the 1-hour forecast also indicate that the prediction error increases as the time to be forecasted lies further in the future.

Last but not least, it can also be seen that the prediction results are strongly dependent on the time and the type of day. The MAE values correlate with the changing load level during the course of the day. In times of low load, the MAE also remain low and increase in times of higher aggregated load. Additionally, the prediction results vary depending on whether it is a weekday, weekend or public holiday. This influence is mainly visible for the 1-day load prediction. Moreover, under the given circumstances, the LSTM is not able to predict the peak loads accurately at the respective charging sites based on the 1-day forecast.

The findings of this work benefit a wide range of stakeholders. Sub-aggregators have a great interest in a reliable load forecast in order to offer flexibility potentials for the CR market. Network operators, on the other hand, are keen on forecasting the charging load in order to identify possible bottlenecks in the low voltage network caused by simultaneous PEV charging at an early stage. The results of this work provide valuable insights on how LSTMs can be applied most effectively for a valid charging load prediction and thus provide useful guidelines for future research which will be discussed in the following.

8.3. Future work

The outlined weaknesses of the 1-day forecast can be addressed by improving the LSTM and training it on data collected over a longer period of time. With data over several years available in the future, a renewed load forecast would be beneficial to investigate the influence on the accuracy of the forecast. In this context, the LSTM could be extended by further indicators such as the month, the stateful mode could be investigated again and time series cross validation could be performed to increase the robustness of the model and overcome early overfitting.

In view of the fact that the accuracy of the forecast increases significantly with a shortened forecast horizon, it is advisable to focus the forecast on the suitable days and time periods in the course of the day of each charging site when a sufficient volume of PEV is available for pooling the PEV flexibility. A future research objective thus involves the development of a customized LSTM model tailored to each charging site. For residential charging, the forecast period from 4:00 p.m. to 8:00 p.m. seems to be a good choice. For public charging, however, the time period 8:00 a.m. to 12:00 p.m. would be reasonable and for workplace charging a focus on the time period from 6:00 a.m. to 10:00 a.m. is recommended. For shopping center charging no recommendation can be given due to the strongly fluctuating load. For all charging sites, the load forecast should be targeted to weekdays, since user behavior on weekends and holidays is often difficult to predict and in many cases, as with workplace charging, the aggregation potential is not sufficient enough.

Furthermore, in order to predict the flexibility potential of PEVs, a fundamentally different approach is conceivable for future research, focusing on predicting the start of charging, the end of charging and the departure time in order to identify the flexibility potential without restricting the charging process outcome of the user.

Finally, the outcome of the 1-day forecast has shown that with the chosen approach it is currently not possible to reliably predict the level and time of peak loads. The previously discussed targeting and customization of the LSTM to the period of the highest load could significantly improve the accuracy of the forecast. However, another research approach is also feasible, focusing on the sole prediction of the peak load level and the timing of the peak load.

A. Appendix

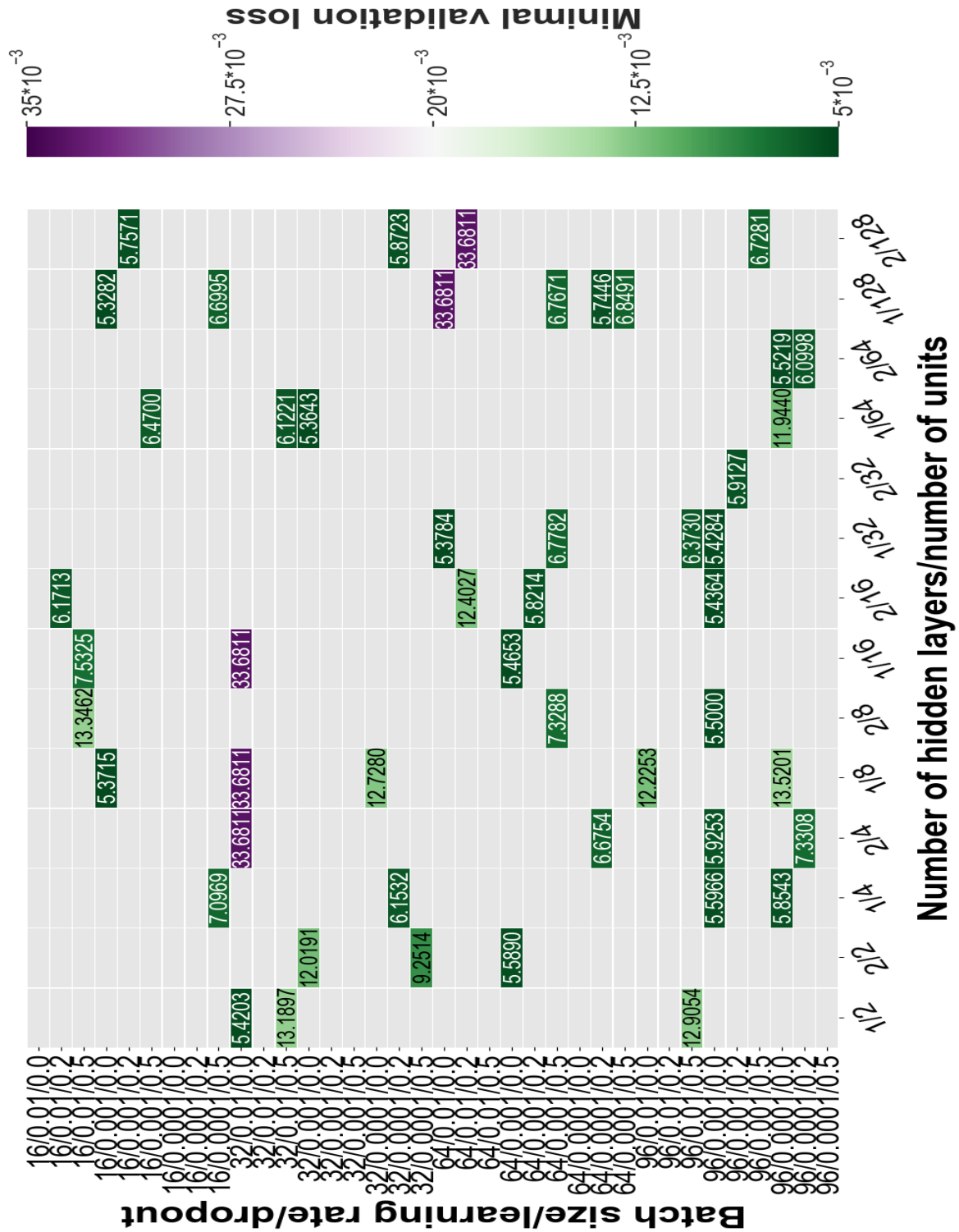## A.1. Hyperparameter tuning results

Figure A.1: Minimal validation losses of random search hyperparameter combinations for the 4-timesteps shopping center charging forecast

Figure A.2: Minimal validation losses of random search hyperparameter combinations for the 96-timesteps shopping center charging forecast
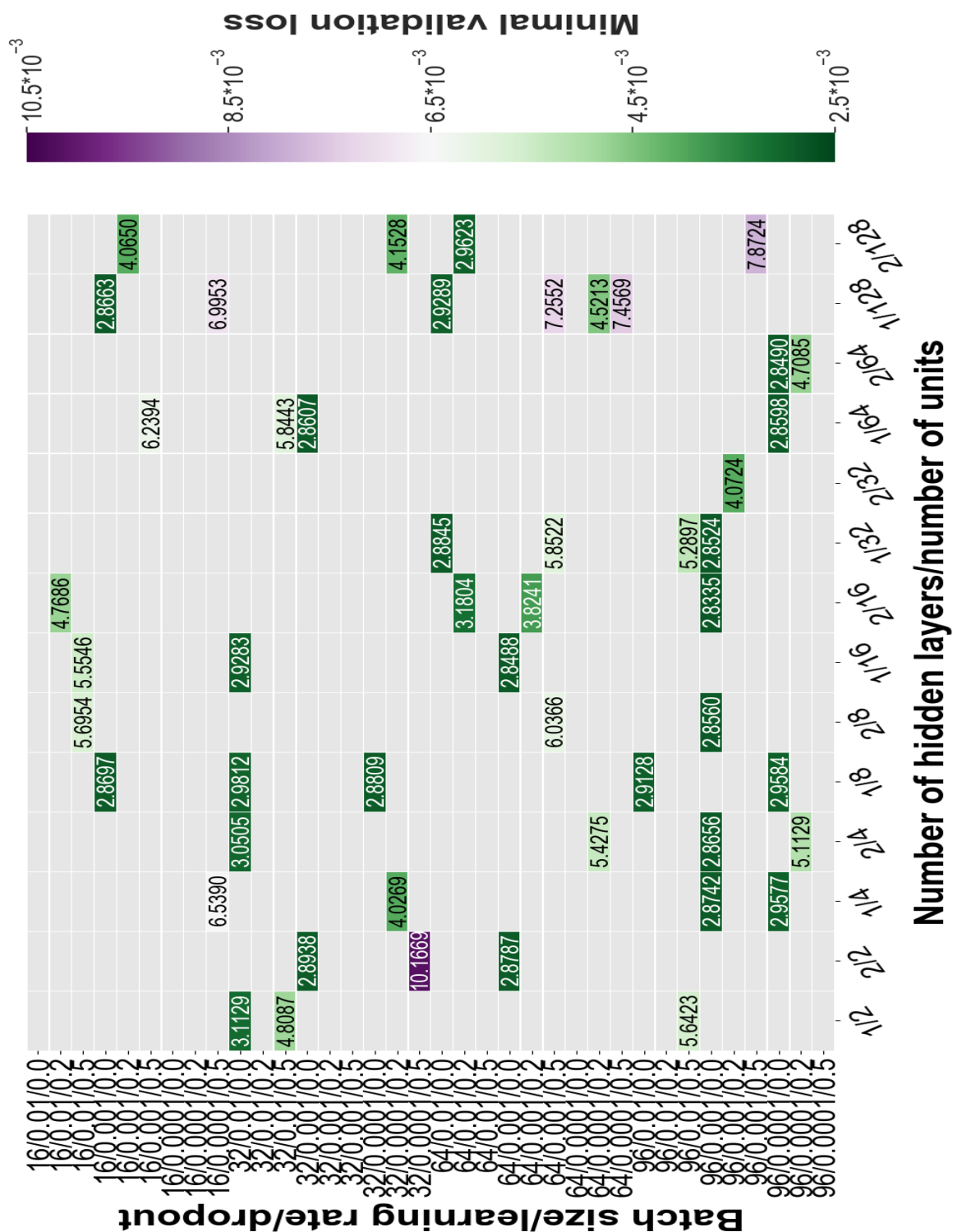
Figure A.3: Minimal validation losses of random search hyperparameter combinations for the 4-timesteps residential charging forecast
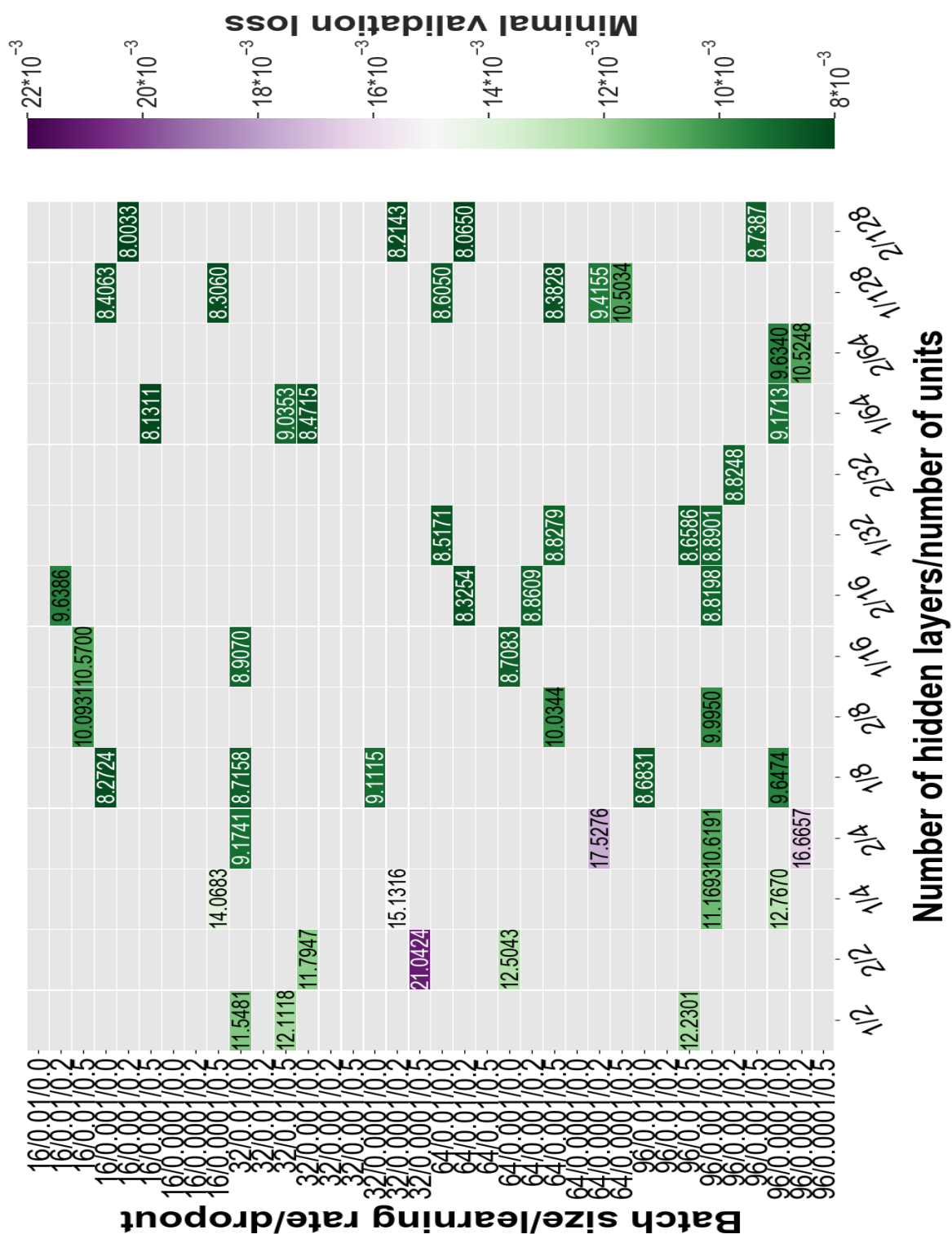
Figure A.4: Minimal validation losses of random search hyperparameter combinations for the 96-timesteps residential charging forecast

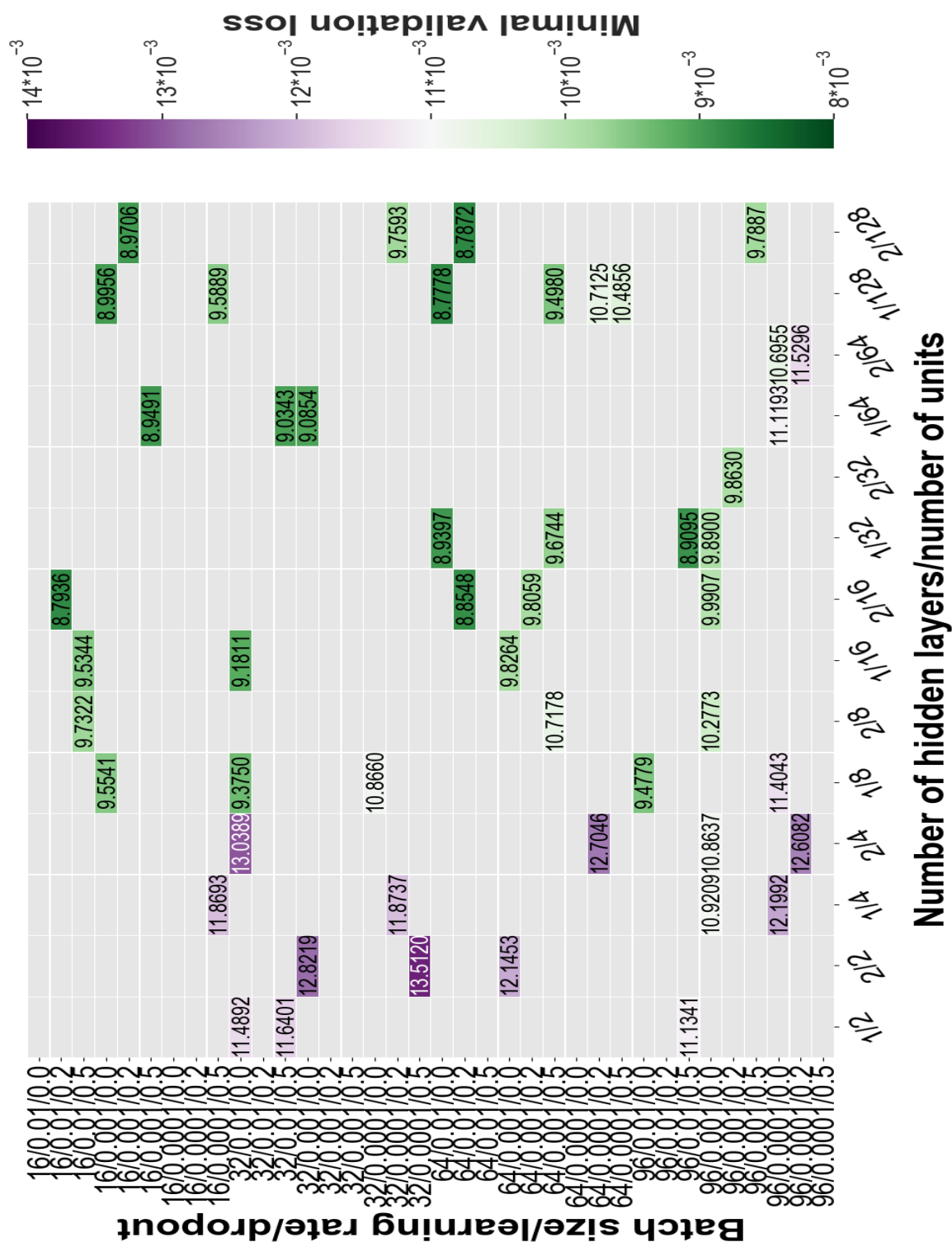Figure A.5: Minimal validation losses of random search hyperparameter combinations for the 4-timesteps public charging forecast

Figure A.6: Minimal validation losses of random search hyperparameter combinations for the 96-timesteps public charging forecast

Figure A.7: Minimal validation losses of random search hyperparameter combinations for the 4-timesteps workplace charging forecast

Figure A.8: Minimal validation losses of random search hyperparameter combinations for the 96-timesteps workplace charging forecast

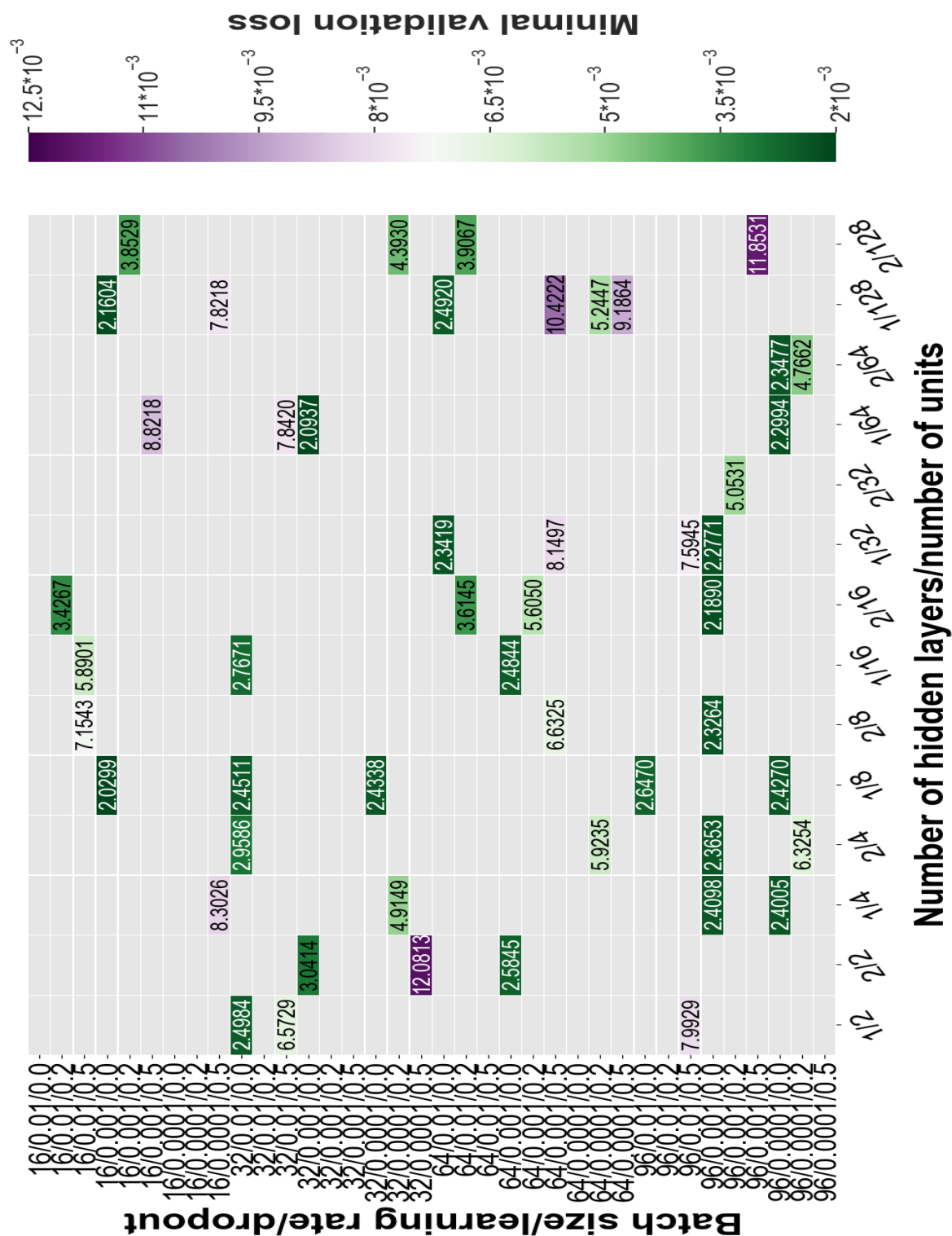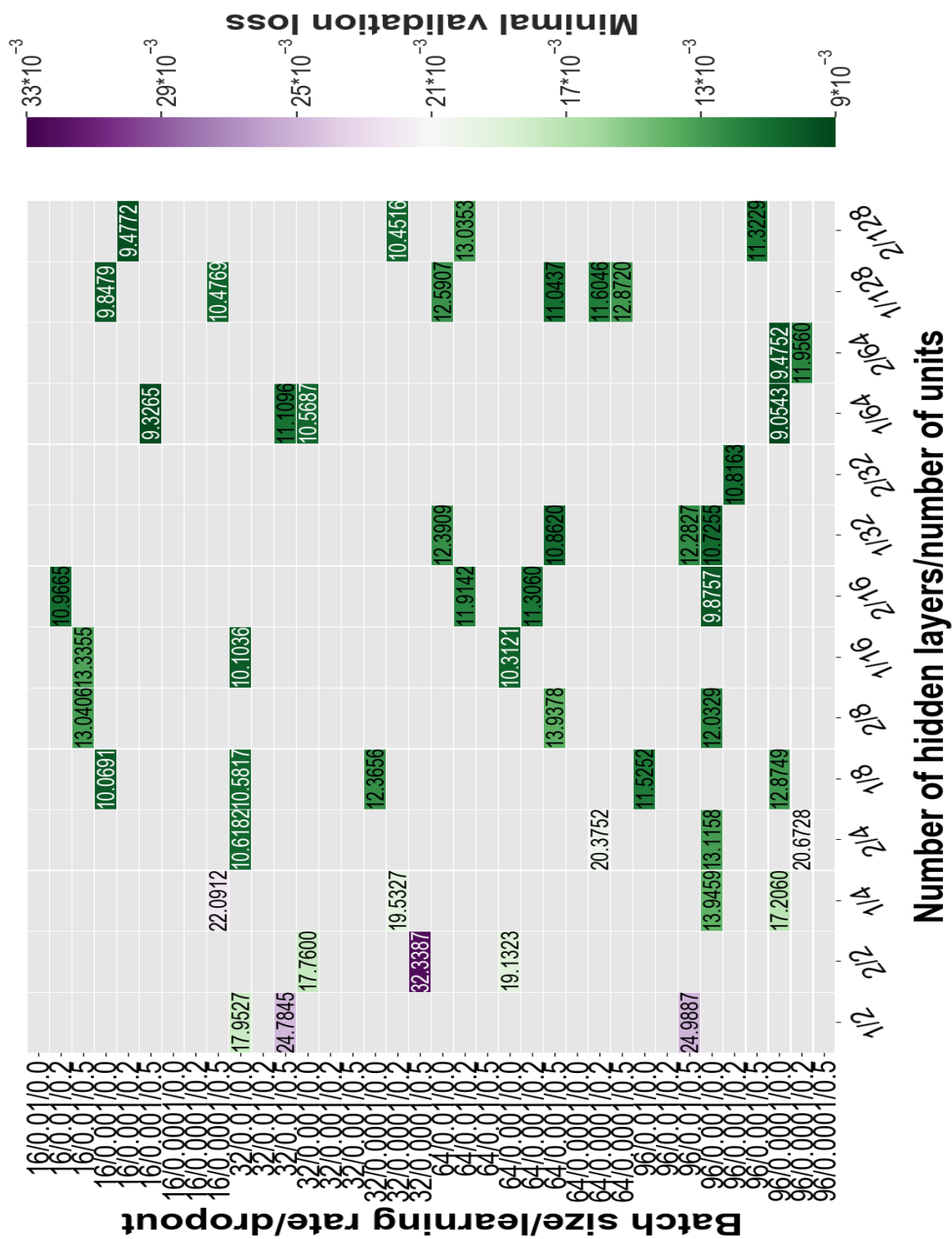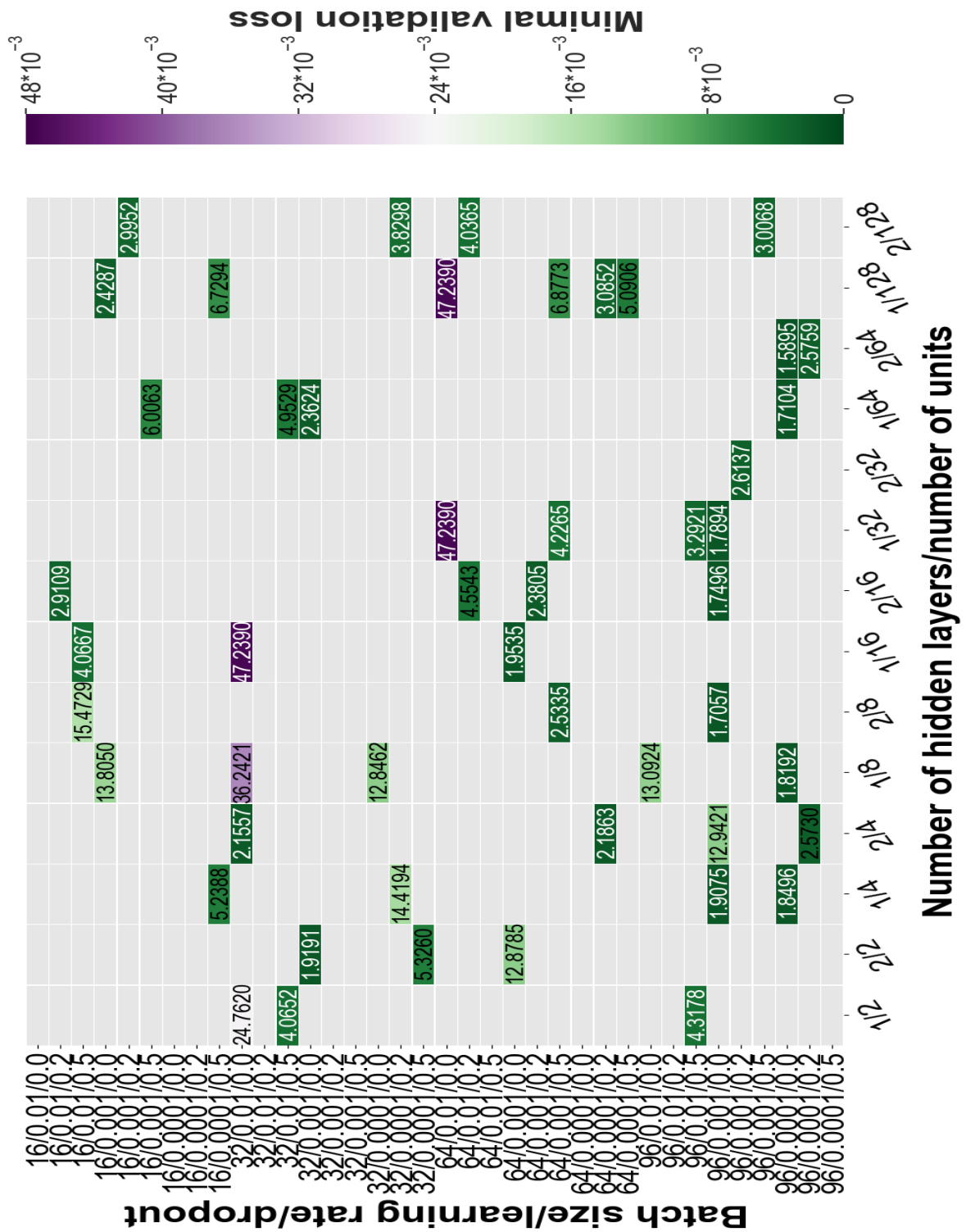## A.2. Forecast results

Timeslot	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th	16th	Average
Shopping center	0.55/1.24	1.15/1.36	0.57/0.46	0.38/0.41	0.29/0.42	0.85/0.93	0.85/0.98	0.66/0.82	0.55/0.26	0.26/0.35	0.24/0.15	0.15/0.15	0.15/0.15	0.29/0.29	0.29/0.29	0.29/0.29	0.47/0.53
Residential	1.07/1.61	1.37/1.73	1.39/1.67	1.38/1.66	0.68/1.54	0.89/1.49	0.96/1.51	1.00/1.46	0.37/1.40	0.59/1.61	0.71/1.67	0.75/1.57	0.43/1.47	0.53/1.42	0.53/1.39	0.59/1.31	0.83/1.53
Public	0.95/1.22	0.73/1.45	0.70/1.53	0.92/1.69	0.90/1.41	0.90/1.50	1.17/1.40	1.38/1.40	0.76/1.17	0.80/1.26	1.37/1.35	1.59/1.82	1.00/1.68	1.08/2.00	1.19/2.02	1.30/2.03	1.05/1.56
Workplace	0.28/0.58	0.24/0.58	0.23/0.55	0.25/0.40	0.13/0.34	0.13/0.40	0.15/0.59	0.12/0.84	0.11/0.84	0.08/0.75	0.18/0.67	0.17/0.74	0.13/0.69	0.28/0.73	0.77/0.88	1.15/1.18	0.27/0.67

Timeslot	17st	18th	19th	20th	21st	22nd	23rd	24th	25th	26th	27th	28th	29th	30th	31st	32nd	Average
Shopping center	0.29/0.29	0.29/0.29	0.29/0.29	0.29/0.29	0.29/0.29	0.29/0.29	0.29/0.29	0.14/0.14	0.22/0.22	0.22/0.22	0.38/0.38	1.14/1.20	1.37/2.28	2.69/3.12	2.99/3.27	4.10/4.76	0.96/1.10
Residential	0.26/1.33	0.50/1.38	0.59/1.55	0.70/1.46	0.24/1.48	0.27/1.42	0.59/1.47	0.82/1.52	0.32/1.56	0.58/1.67	0.69/1.76	0.84/1.87	0.40/1.78	0.53/1.62	0.74/1.83	1.03/1.94	0.57/1.60
Public	1.00/2.07	1.67/2.62	2.78/3.36	4.64/4.41	2.24/5.99	4.08/8.80	5.26/9.41	6.37/10.94	3.20/11.50	3.30/10.44	5.01/11.20	6.19/10.83	2.24/10.85	3.06/10.79	5.67/12.09	7.66/12.73	4.02/8.63
Workplace	0.81/1.58	1.38/1.90	2.66/2.53	4.08/4.03	1.16/5.51	1.95/7.86	3.59/10.92	5.43/13.59	1.56/14.70	2.96/14.00	6.06/13.86	10.57/13.13	3.59/13.04	6.82/14.4	9.46/15.80	11.38/16.54	4.59/10.21

Timeslot	33rd	34th	35th	36th	37th	38th	39th	40th	41st	42nd	43rd	44th	45th	46th	47th	48th	Average
Shopping center	1.56/4.99	2.92/6.29	5.42/8.49	8.87/10.69	3.63/12.11	7.65/14.12	9.59/15.87	10.74/14.69	6.04/13.90	7.64/15.64	8.11/13.24	8.23/12.37	4.64/12.43	8.31/11.92	8.29/12.10	9.81/11.5	6.97/11.90
Residential	0.34/1.92	0.69/1.95	0.99/2.16	1.33/2.42	0.81/2.42	1.33/2.50	1.40/2.36	1.60/2.16	0.76/2.49	1.01/2.50	1.36/2.26	1.36/2.17	0.68/2.27	1.11/2.16	1.42/2.15	1.83/2.32	1.13/2.26
Public	1.96/12.85	2.35/12.79	3.63/12.71	4.83/12.58	2.37/13.65	3.50/13.15	4.54/13.65	4.65/12.65	2.03/12.53	3.18/11.95	4.07/11.95	4.01/10.87	3.56/8.17	4.43/8.03	4.65/7.72	4.46/7.07	3.64/11.40
Workplace	2.37/16.62	3.48/17.03	4.78/17.35	4.97/17.29	2.29/16.89	3.08/16.67	3.31/15.48	3.81/14.84	2.98/14.84	4.57/14.50	4.94/13.39	3.94/10.96	2.15/9.13	2.22/8.40	2.48/8.34	3.09/7.96	3.40/13.73

Timeslot	49st	50th	51st	52nd	53rd	54th	55th	56th	57th	58th	59th	60th	61st	62nd	63rd	64th	Average
Shopping center	7.29/13.55	7.94/12.13	8.33/11.04	9.78/12.39	6.22/13.41	9.18/15.40	12.10/16.78	13.01/17.61	6.10/16.74	6.27/14.31	6.58/14.29	7.75/13.72	6.86/16.52	8.51/12.89	8.88/10.88	10.16/10.68	8.44/13.9
Residential	1.26/2.65	1.66/2.91	1.85/3.35	2.35/3.27	1.26/3.43	1.66/3.16	2.40/3.28	2.93/3.53	1.81/3.57	2.27/3.71	2.65/3.59	3.15/3.77	2.02/4.06	2.30/3.74	3.06/4.23	3.53/4.75	2.26/3.56
Public	2.60/6.94	3.60/6.24	3.86/6.08	4.91/6.64	2.91/6.96	3.75/6.89	3.80/7.02	3.98/6.91	2.46/7.15	2.51/5.92	3.11/5.61	3.52/5.85	2.41/5.96	2.99/5.74	3.13/5.68	3.08/5.48	3.29/6.32
Workplace	1.53/7.23	1.87/6.50	2.10/5.72	2.47/5.79	1.27/5.16	1.34/4.65	1.56/4.13	1.74/3.80	0.88/3.06	1.25/3.10	1.39/3.03	1.70/2.62	1.13/2.71	1.41/2.58	1.44/2.42	1.56/2.48	1.54/4.06

Timeslot	65th	66th	67th	68th	69th	70th	71st	72nd	73rd	74th	75th	76th	77th	78th	79th	80th	Average
Shopping center	4.16/10.90	6.38/11.00	7.87/10.49	8.73/10.21	5.45/11.46	7.29/11.42	9.16/11.91	9.59/11.60	5.37/10.00	8.02/10.58	7.57/9.49	8.39/10.01	4.68/8.78	6.54/8.63	7.23/8.69	7.35/8.23	7.11/10.21
Residential	2.23/5.98	2.24/5.60	2.60/5.96	3.31/6.34	2.08/6.23	3.04/6.11	3.67/6.16	4.29/6.44	1.59/6.34	2.71/6.84	2.78/6.68	2.83/6.72	1.39/6.83	1.96/6.73	2.34/6.59	3.00/6.43	2.63/6.37
Public	2.32/5.89	2.85/5.24	3.50/5.98	3.51/5.75	1.63/5.86	1.86/5.57	1.97/5.36	2.41/5.05	2.33/4.82	2.61/4.17	3.05/3.93	2.94/4.05	1.94/3.56	2.56/3.22	2.75/3.02	3.21/3.08	2.59/4.66
Workplace	0.85/2.37	1.02/2.38	1.46/2.59	1.61/2.39	0.52/2.27	0.65/2.30	0.75/2.21	1.00/2.03	0.64/1.82	0.76/1.87	0.91/1.93	1.06/1.88	0.47/1.88	0.58/1.96	0.81/1.97	0.84/2.07	0.87/2.12

Timeslot	81st	82nd	83rd	84th	85th	86th	87th	88th	89th	90th	91st	92nd	93rd	94st	95th	96th	Average
Shopping center	4.08/7.64	6.14/7.87	4.81/6.37	4.31/5.94	1.83/5.09	2.68/4.30	2.70/3.56	2.15/1.98	0.80/1.15	0.61/0.41	0.53/0.41	0.53/0.43	0.24/0.28	0.76/0.76	0.76/0.76	1.41/1.41	2.15/3.02
Residential	1.36/6.19	1.62/5.99	2.26/6.05	2.54/5.99	1.13/5.71	1.92/5.31	2.48/4.84	2.73/5.34	1.06/5.04	1.59/4.62	1.79/4.21	2.10/4.24	1.03/4.11	1.45/3.96	1.68/3.37	1.61/3.41	1.77/4.90
Public	2.01/3.14	2.29/2.90	2.19/2.74	2.10/2.40	1.82/2.23	1.49/2.72	1.54/2.29	1.68/2.85	1.66/2.43	1.34/2.47	1.26/2.68	1.18/2.40	1.27/2.36	1.21/2.03	1.15/2.20	1.36/2.28	1.60/2.51
Workplace	0.43/1.97	0.59/1.72	0.67/1.59	0.58/1.51	0.29/1.33	0.38/1.24	0.34/1.17	0.38/1.04	0.36/1.06	0.42/1.16	0.45/1.23	0.49/1.18	0.29/1.28	0.26/1.21	0.23/1.22	0.32/1.02	0.40/1.31

Table A.1: Summary of the MAE results in kW for each 15-min timeslot (4-timesteps prediction MAE in kW/96-timesteps prediction MAE in kW)

References

[1] Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit, "Internationale Klimapolitik," Sep. 2017, accessed: 23.03.2020. [Online]. Available: https://www.bmu.de/themen/klima-energie/klimaschutz/internationale-klimapolitik/

[2] European Environment Agency, "Greenhouse gas emissions from transport in Europe," Dec. 2019, accessed: 23.03.2020. [Online]. Available: https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12

[3] ——, "Greenhouse gas emissions by aggregated sector," Dec. 2019, accessed: 23.03.2020. [Online]. Available: https://www.eea.europa.eu/data-and-maps/daviz/ghg-emissions-by-aggregated-sector-5#tab-dashboard-02

[4] European Comission, "2050 long-term strategy," accessed: 23.03.2020. [Online]. Available: https://ec.europa.eu/clima/policies/strategies/2050_en#tab-0-0

[5] Z. Yang, K. Li, and A. Foley, "Computational scheduling methods for integrating plug-in electric vehicles with power systems: A review," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 396–416, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032115005778

[6] T. Gnann, T. S. Stephens, Z. Lin, P. Plötz, C. Liu, and J. Brokate, "What drives the market for plug-in electric vehicles? - A review of international PEV market diffusion models," *Renewable and Sustainable Energy Reviews*, vol. 93, pp. 158–164, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032118301497

[7] N. Daina, A. Sivakumar, and J. W. Polak, "Modelling electric vehicles use:

a survey on the methods," *Renewable and Sustainable Energy Reviews*, vol. 68, pp. 447–460, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032116306566

[8] T. P. Narins, "The battery business: Lithium availability and the growth of the global electric car industry," *The Extractive Industries and Society*, vol. 4, no. 2, pp. 321–328, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2214790X17300175

[9] J. Martìnez-Lao, F. G. Montoya, M. G. Montoya, and F. Manzano-Agugliaro, "Electric vehicles in Spain: An overview of charging systems," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 970–983, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032116310152

[10] O. Egbue and S. Long, "Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions," *Energy Policy*, vol. 48, pp. 717–729, 2012, special Section: Frontiers of Sustainability. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0301421512005162

[11] Z. Ji and X. Huang, "Plug-in electric vehicle charging infrastructure deployment of China towards 2020: Policies, methodologies, and challenges," *Renewable and Sustainable Energy Reviews*, vol. 90, pp. 710–727, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S136403211830220X

[12] Y. J. Jang, "Survey of the operation and system study on wireless charging electric vehicle systems," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 844–866, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X18304649

[13] BDEW, DKE, ZVEH, ZVEI and VDE/FNN, "Der Technische Leitfaden – Ladeinfrastruktur Elektromobilität Version 3," Tech. Rep., 2020.

[14] Nationale Plattform Elektromobilität, "Ladeinfrastruktur für Elektrofahrzeuge in Deutschland - Statusbericht und Handlungsempfehlungen 2015," Tech. Rep., 2015.

[15] P. Komarnicki, J. Haubrock, and Z. A. Styczynski, *Elektromobilität und Sektorenkopplung: Infrastruktur-und Systemkomponenten*. Springer-Verlag, 2018.

[16] Bundesnetzagentur, "Ladesäulenkarte," accessed: 13.02.2020. [Online]. Available: https://www.bundesnetzagentur.de/DE/Sachgebiete/ ElektrizitaetundGas/Unternehmen_Institutionen/HandelundVertrieb/ Ladesaeulenkarte/Ladesaeulenkarte_node.html

[17] Nationale Plattform Elektromobilität, "Fortschrittsbericht 2018 – Markthochlaufphase," Tech. Rep., 2018.

[18] T. Gnann, P. Plötz, J. Globisch, U. Schneider, E. Dütschke, S. Funke, M. Wietschel, P. Jochem, M. Heilig, M. Kagerbauer, and M. Reuter-Oppermann, "Öffentliche Ladeinfrastruktur für Elektrofahrzeuge: Ergebnisse der Profilregion Mobilitätssysteme Karlsruhe," Fraunhofer-Institut für System- und Innovationsforschung ISI, Karlsruhe, Tech. Rep., 2017.

[19] Deutsches Zentrum für Luft- und Raumfahrt, IMU Institut and bridgingIT, "Strukturstudie BWe mobil 2019 - Transformation durch Elektromobilität und Perspektiven der Digitalisierung," Tech. Rep., 2019.

[20] Tilastokeskuks, "Cars by driving power, 1990 - 2019," accessed: 01.05.2020. [Online]. Available: http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin_ _lii__mkan/statfin_mkan_pxt_11ie.px/

[21] Kraftfahrt-Bundesamt, "Bestand," accessed: 01.05.2020. [Online]. Available: https://www.kba.de/DE/Statistik/Fahrzeuge/Bestand/bestand_node.html

[22] Bundesregierung, "Klimaschutz – Verkehr," accessed: 01.05.2020. [Online]. Available: https://www.bundesregierung.de/breg-de/themen/klimaschutz/ verkehr-1672896

[23] Die Bundesregierung, "So funktioniert der neue Umweltbonus," Nov. 2019, accessed: 27.06.2020. [Online]. Available: https://www.bundesregierung.de/ breg-de/aktuelles/umweltbonus-1692646

[24] The Association of Automobile Importer in Finland, "Purchase incentives for battery electric vehicles," accessed: 01.05.2020. [Online]. Available: https://www.autotuojat.fi/en/current_affairs/purchase_incentive_of_battery_electric_vehicles

[25] A. Zakariazadeh, S. Jadid, and P. Siano, "Multi-objective scheduling of electric vehicles in smart distribution system," *Energy Conversion and Management*, vol. 79, pp. 43–53, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196890413007632

[26] M. Aro, K. Mäki, P. Järventausta, and H. Vesa, "Sub-aggregator as a Key Enabler in Harnessing Demand Response Potential of Electric Vehicles," in *2019 16th International Conference on the European Energy Market (EEM)*, Sep. 2019, pp. 1–5. [Online]. Available: https://doi.org/10.1109/EEM.2019.8916429

[27] A. Ruuska, O. Parkkonen, P. Ruokolainen, and M. Aro, "Smart Otaniemi webinar – Aggregator business models," Sep. 2019, Accessed: 15.03.2020. [Online]. Available: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwj6uajzub_mAhVcxcQBHZWTBPAQFjAAegQIBBAC&url=https%3A%2F%2Fsmartotaniemi.fi%2Fwp-content%2Fuploads%2F2019%2F09%2FSmartOtaniemi_Aggregator_business_webinar_05092019.pdf&usg=AOvVaw0OZE-URXvWOoCNhhJCxTsk

[28] I. Koprinska, M. Rana, and V. G. Agelidis, "Correlation and instance based feature selection for electricity load forecasting," *Knowledge-Based Systems*, vol. 82, pp. 29–40, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0950705115000714

[29] F. Martínez-Álvarez, A. Troncoso, G. Asencio-Cortés, and J. Riquelme, "A Survey on Data Mining Techniques Applied to Electricity-Related Time Series Forecasting," *Energies*, vol. 8, no. 11, pp. 13 162–13 193, Nov. 2015. [Online]. Available: http://dx.doi.org/10.3390/en81112361

[30] M. Rana and A. Rahman, "Multiple steps ahead solar photovoltaic

power forecasting based on univariate machine learning models and data re-sampling," *Sustainable Energy, Grids and Networks*, vol. 21, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352467719304023

[31] S. B. Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7067–7083, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417412000528

[32] H. M. Louie, "Time-Series Modeling of Aggregated Electric Vehicle Charging Station Load," *Electric Power Components and Systems*, vol. 45, no. 14, pp. 1498–1511, 2017. [Online]. Available: https://doi.org/10.1080/15325008.2017.1336583

[33] M. H. Amini, A. Kargarian, and O. Karabasoglu, "ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation," *Electric Power Systems Research*, vol. 140, pp. 378–390, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378779616302115

[34] M. Pertl, F. Carducci, M. Tabone, M. Marinelli, S. Kiliccote, and E. C. Kara, "An Equivalent Time-Variant Storage Model to Harness EV Flexibility: Forecast and Aggregation," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 1899–1910, Apr. 2019. [Online]. Available: https://doi.org/10.1109/TII.2018.2865433

[35] H. Jahangir, H. Tayarani, A. Ahmadian, M. A. Golkar, J. Miret, M. Tayarani, and H. O. Gao, "Charging demand of Plug-in Electric Vehicles: Forecasting travel behavior based on a novel Rough Artificial Neural Network approach," *Journal of Cleaner Production*, vol. 229, pp. 1029–1044, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0959652619314428

[36] J. Zhu, Z. Yang, Y. Chang, Y. Guo, K. Zhu, and J. Zhang, "A novel LSTM based deep learning approach for multi-time scale electric vehicles

charging load prediction," in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, May 2019, pp. 3531–3536. [Online]. Available: https://doi.org/10.1109/ISGT-Asia.2019.8881655

[37] J. Zhu, Z. Yang, M. Mourshed, Y. Guo, Y. Zhou, Y. Chang, Y. Wei, and S. Feng, "Electric Vehicle Charging Load Forecasting: A Comparative Study of Deep Learning Approaches," *Energies*, vol. 12, no. 14, p. 2692, July 2019. [Online]. Available: http://dx.doi.org/10.3390/en12142692

[38] J. Zhu, Z. Yang, Y. Guo, J. Zhang, and H. Yang, "Short-Term Load Forecasting for Electric Vehicle Charging Stations Based on Deep Learning Approaches," *Applied Sciences*, vol. 9, p. 1723, Apr. 2019. [Online]. Available: https://doi.org/10.3390/app9091723

[39] A. Gerossier, R. Girard, and G. Kariniotakis, "Modeling and Forecasting Electric Vehicle Consumption Profiles," *Energies*, vol. 12, no. 7, p. 1341, Apr. 2019. [Online]. Available: http://dx.doi.org/10.3390/en12071341

[40] X. Zhang, "Short-Term Load Forecasting for Electric Bus Charging Stations Based on Fuzzy Clustering and Least Squares Support Vector Machine Optimized by Wolf Pack Algorithm," *Energies*, vol. 11, p. 1449, June 2018. [Online]. Available: https://doi.org/10.3390/en11061449

[41] Y. Li, Y. Huang, and M. Zhang, "Short-Term Load Forecasting for Electric Vehicle Charging Station Based on Niche Immunity Lion Algorithm and Convolutional Neural Network," *Energies*, vol. 11, no. 5, p. 1253, May 2018. [Online]. Available: http://dx.doi.org/10.3390/en11051253

[42] Y. Lu, Y. Li, D. Xie, E. Wei, X. Bao, H. Chen, and X. Zhong, "The Application of Improved Random Forest Algorithm on the Prediction of Electric Vehicle Charging Load," *Energies*, vol. 11, p. 3207, Nov. 2018. [Online]. Available: https://doi.org/10.3390/en11113207

[43] L. Buzna, P. De Falco, S. Khormali, D. Proto, and M. Straka, "Electric vehicle load forecasting: A comparison between time series and machine learning approaches," in *2019 1st International Conference on Energy Transition in*

the Mediterranean Area (SyNERGY MED), May 2019, pp. 1–5. [Online]. Available: https://doi.org/10.1109/SyNERGY-MED.2019.8764110

[44] M. Majidpour, C. Qiu, P. Chu, H. R. Pota, and R. Gadh, "Forecasting the EV charging load based on customer profile or station measurement?" *Applied Energy*, vol. 163, pp. 134–141, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261915014348

[45] H. Lin, K. Fu, Y. Liu, Q. Sun, and R. Wennersten, "Modeling charging demand of electric vehicles in multi-locations using agent-based method," *Energy Procedia*, vol. 152, pp. 599–605, 2018, Cleaner Energy for Cleaner Cities. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1876610218307628

[46] H. Lin, K. Fu, Y. Wang, Q. Sun, H. Li, Y. Hu, B. Sun, and R. Wennersten, "Characteristics of electric vehicle charging demand at multiple types of location - Application of an agent-based trip chain model," *Energy*, vol. 188, p. 116122, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544219318171

[47] New 4.0 GbR, "Entwicklung eines Sub-Aggregators," accessed: 04.03.2020. [Online]. Available: https://new4-0.erneuerbare-energien-hamburg.de/de/new-40-projekte/details/entwicklung-eines-sub-aggregators.html

[48] 50Hertz Transmission GmbH, Amprion GmbH, TenneT TSO GmbH and TransnetBW GmbH, "General information on control reserve - technical aspects," accessed: 01.05.2020. [Online]. Available: https://www.regelleistung.net/ext/static/technical?lang=en

[49] Bundesnetzagentur, "Regelleistung," accessed: 01.05.2020. [Online]. Available: https://smard.de/home/wiki-article/446/396

[50] M. Kriener and K. Simons, "Driving the Energy Transition: wie Elektromobilität die Energiewende unterstützen kann," WWF Deutschland and LichtBlick SE, Tech. Rep., 2017.

[51] Bundesnetzagentur, "So funktioniert der Strommarkt," May 2020, accessed: 01.05.2020. [Online]. Available: https://www.smard.de/home/wiki-article/446/ 384

[52] G. Göhler, C. Schmaus, and A.-L. Klingler, "Netzbelastungen und Netzdienstleistungen durch Elektrofahrzeuge: Metastudie," 2019.

[53] Next Kraftwerke GmbH, "Was ist Sekundärregelleistung (SRL)?" accessed on: 01.05.2020. [Online]. Available: https://www.next-kraftwerke.de/wissen/ sekundaerreserve

[54] ——, "Was ist Minutenreserveleistung (MRL)?" accessed on: 01.05.2020. [Online]. Available: https://www.next-kraftwerke.de/wissen/ minutenreserve-tertiaerregelung

[55] P. Fenner, "Efficient use of the existing real estate infrastructure for electric vehicle charging," Master's thesis, 2019.

[56] R. Henger, H. Scheunemann, M. Barthauer, C. Giesemann, M. Hude, B. Seipelt, and A. Toschka, "Büroimmobilien. Energetischer Zustand und Anreize zur Steigerung der Energieeffizienz," Deutsche Energie-Agentur GmbH (dena), Tech. Rep., May 2017.

[57] Statistisches Bundesamt, "Bauen und Wohnen. Baugenehmigungen / Baufertigstellungen. Lange Reihen z. T. ab 1949," Statistisches Bundesamt (Destatis), Tech. Rep. 5311101187004, June 2019.

[58] ——, "Bauen und Wohnen. Baugenehmigungen/ Baufertigstellungen von Nichtwohngebäuden (Neubau). Lange Reihen z. T. ab 1980," Statistisches Bundesamt (DESTATIS), Tech. Rep. 5311105187004, July 2019.

[59] E. Heinrichs, M. Schreiber, S. Rath, I. Kosarev, and L. Weinke, "Untersuchung von Stellplatzsatzungen und Empfehlungen für Kostensenkungen unter Beachtung moderner Mobilitätskonzepte," LK Argus GmbH, Tech. Rep., July 2015.

[60] Architektenkammer Baden-Württemberg, "Verwaltungsvorschrift des Ministeriums für Verkehr und Infrastruktur über die Herstellung notwendiger Stellplätze (VwV Stellplätze)," May 2015, accessed: 05.04.2020. [Online]. Available: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web& cd=1&ved=2ahUKEwj8j_Sy38HmAhXM5KQKHT7cCVQQFjAAegQIARAC& url=https%3A%2F%2Fwww.akbw.de%2Ffileadmin%2Fdownload% 2Fdokumenten_datenbank%2FAKBW_Merkblaetter%2FBaurecht_ Planungsrecht%2FMerkblatt593-VWV-Stellplaetze2015.pdf&usg= AOvVaw1zp6dGX5TqCZiogLXzltBn

[61] Geschäftsstelle Zukunftsnetz Mobilität NRW, "Kommunale Stellplatzsatzungen. Leitfaden zur Musterstellplatzsatzung NRW," accessed: 05.04.2020. [Online]. Available: https://www.zukunftsnetz-mobilitaet.nrw.de/sites/default/ files/downloads/znm_nrw_stellplatzsatzung_handbuch_2019_final.pdf

[62] Ministerium für Umwelt, Energie, Bauen und Klimaschutz Niedersachsen, "Bauaufsicht; Ausführungsbestimmungen zu § 47 NBauO," July 2016, accessed: 05.04.2020. [Online]. Available: http://www.nds-voris.de/jportal/?quelle=jlink&query= VVND-210720-MS-20160706-SF&psml=bsvorisprod.psml&max=true

[63] Behörde für Stadtentwicklung und Umwelt Amt für Bauordnung und Hochbau, "Fachanweisung zugleich Dienstanweisung für die Bauprüfabteilung der Hamburg Port Authority und das Genehmigungsreferat HafenCity," Jan. 2013, accessed: 05.04.2020. [Online]. Available: https://www.hamburg.de/contentblob/ 3810612/3c5b1093e482f0ee77f1bb88fcf88f93/data/fa-1-2013.pdf

[64] Freie Hansestadt Bremen, "Ortsgesetz über Kraftfahrzeugstellplätze und Fahrradabstellplätze in der Stadtgemeinde Bremen (Stellplatzortsgesetz Bremen - StellplOG)," accessed: 05.04.2020. [Online]. Available: https://www.transparenz.bremen.de/sixcms/detail.php?gsid=bremen2014_tp. c.70002.de&template=20_gp_ifg_meta_detail_d#_XY_d238385e12738

[65] 50Hertz Transmission GmbH, Amprion GmbH, TenneT TSO GmbH and TransnetBW GmbH, "Regelleistung.net

- Datencenter," accessed: 01.05.2020. [Online]. Available: https://www.regelleistung.net/apps/datacenter/tenders/?productTypes= SRL,MRL&from=2020-04-06&to=2020-04-10&tid=SRL_20200406_D1

[66] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032117303155

[67] I. N. da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, *Artificial neural networks: A Practical Course*. Springer, 2017.

[68] T. Hong, "Short Term Electric Load Forecasting," Ph.D. dissertation, Jan. 2010.

[69] W. Schellong, *Analyse und Optimierung von Energieverbundsystemen*. Springer, 2016.

[70] S. Ryu, J. Noh, and H. Kim, "Deep Neural Network Based Demand Side Short Term Load Forecasting," *Energies*, vol. 10, no. 1, 2017. [Online]. Available: https://www.mdpi.com/1996-1073/10/1/3

[71] K. Backhaus, B. Erichson, and R. Weiber, *Fortgeschrittene multivariate Analysemethoden: eine anwendungsorientierte Einführung*. Springer-Verlag, 2015.

[72] T. Hossen, S. J. Plathottam, R. K. Angamuthu, P. Ranganathan, and H. Salehfar, "Short-term load forecasting using deep neural networks (DNN)," in *2017 North American Power Symposium (NAPS)*, Sep. 2017, pp. 1–6. [Online]. Available: https://doi.org/10.1109/NAPS.2017.8107271

[73] F. Mohammad, K. B. Lee, and Y. Kim, "Short Term Load Forecasting Using Deep Neural Networks," *CoRR*, vol. abs/1811.03242, 2018. [Online]. Available: http://arxiv.org/abs/1811.03242

[74] M. Q. Raza and A. Khosravi, "A review on artificial intelligence

based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032115003354

[75] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, *Recurrent neural networks for short-term load forecasting: an overview and comparative analysis*. Springer, 2017.

[76] W. De Mulder, S. Bethard, and M.-F. Moens, "A survey on the application of recurrent neural networks to statistical language modeling," *Computer Speech & Language*, vol. 30, no. 1, pp. 61–98, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S088523081400093X

[77] S. Shanmuganathan, "Artificial neural network modelling: An introduction," in *Artificial Neural Network Modelling*. Springer, 2016, pp. 1–14.

[78] N. Somu, G. R. MR, and K. Ramamritham, "A hybrid model for building energy consumption forecasting using long short term memory networks," *Applied Energy*, vol. 261, p. 114131, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261919318185

[79] C. Liu, Z. Jin, J. Gu, and C. Qiu, "Short-term load forecasting using a long short-term memory network," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. IEEE, 2017, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ISGTEurope.2017.8260110

[80] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019. [Online]. Available: https://doi.org/10.1109/TSG.2017.2753802

[81] S. Muzaffar and A. Afshari, "Short-Term Load Forecasts Using LSTM Networks," *Energy Procedia*, vol. 158, pp. 2922–2927, 2019, Innovative Solutions for Energy Transitions. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1876610219310008

[82] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, "Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches," *Energies*, vol. 11, no. 7, p. 1636, June 2018. [Online]. Available: http://dx.doi.org/10.3390/en11071636

[83] K. Cho, B. van Merriënboer, Ç. Gülçehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *CoRR*, vol. abs/1406.1078, 2014. [Online]. Available: http://arxiv.org/abs/1406.1078

[84] L. Wen, K. Zhou, and S. Yang, "Load demand forecasting of residential buildings using a deep learning model," *Electric Power Systems Research*, vol. 179, p. 106073, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S037877961930392X

[85] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Short-Term Load Forecasting with Multi-Source Data Using Gated Recurrent Unit Neural Networks," *Energies*, vol. 11, no. 5, p. 1138, May 2018. [Online]. Available: http://dx.doi.org/10.3390/en11051138

[86] Y. Wang, W. Liao, and Y. Chang, "Gated Recurrent Unit Network-Based Short-Term Photovoltaic Forecasting," *Energies*, vol. 11, no. 8, p. 2163, Aug. 2018. [Online]. Available: http://dx.doi.org/10.3390/en11082163

[87] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Single and Multi-Sequence Deep Learning Models for Short and Medium Term Electric Load Forecasting," *Energies*, vol. 12, no. 1, p. 149, Jan. 2019. [Online]. Available: http://dx.doi.org/10.3390/en12010149

[88] K. Ke, S. Hongbin, Z. Chengkang, and C. Brown, "Short-term electrical load forecasting method based on stacked auto-encoding and GRU neural network," *Evolutionary Intelligence*, vol. 12, no. 3, pp. 385–394, Jan. 2019. [Online]. Available: https://doi.org/10.1007/s12065-018-00196-0

[89] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural*

computation, vol. 9, no. 8, pp. 1735–1780, 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[90] Y. Wang, D. Gan, M. Sun, N. Zhang, Z. Lu, and C. Kang, "Probabilistic individual load forecasting using pinball loss guided LSTM," *Applied Energy*, vol. 235, pp. 10–20, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261918316465

[91] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544218302056

[92] J. Kämäräinen, "TUT Narvi Cluster," accessed: 01.03.2020. [Online]. Available: https://wiki.eduuni.fi/display/tutsgn/TUT+Narvi+Cluster

[93] Q. Sun, J. Liu, X. Rong, M. Zhang, X. Song, Z. Bie, and Z. Ni, "Charging load forecasting of electric vehicle charging station based on support vector regression," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*.   IEEE, 2016, pp. 1777–1781. [Online]. Available: https://doi.org/10.1109/APPEEC.2016.7779794

[94] Next Kraftwerke GmbH, "Was ist der Intraday-Handel?"   accessed: 03.02.2020. [Online]. Available: https://www.next-kraftwerke.de/wissen/intraday-handel

[95] A. Rautiainen, S. Repo, P. Järventausta, A. Mutanen, K. Vuorilehto, and K. Jalkanen, "Statistical Charging Load Modeling of PHEVs in Electricity Distribution Networks Using National Travel Survey Data," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1650–1659, Dec. 2012. [Online]. Available: https://doi.org/10.1109/TSG.2012.2206411

[96] A. Azadeh, S. Ghaderi, and S. Sohrabkhani, "Forecasting electrical consumption by integration of neural network, time series and ANOVA," *Applied Mathematics and Computation*, vol. 186, no. 2, pp. 1753–1761, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/

S0096300306011106

[97] L. S. Chan and W. C. Tang, *Engineering-Medicine: Principles and Applications of Engineering in Medicine*.   CRC Press, 2019.

[98] F. Chollet, *Deep Learning with Python*.   Manning Publications Co., 2018.

[99] H. Chen, S. Wang, S. Wang, and Y. Li, "Day-ahead aggregated load forecasting based on two-terminal sparse coding and deep neural network fusion," *Electric Power Systems Research*, vol. 177, p. 105987, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/ S0378779619303062

[100] K. Potdar, T. S. Pardawala, and C. D. Pai, "A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers," *International Journal of Computer Applications*, vol. 175, no. 4, pp. 7–9, 2017. [Online]. Available: https://doi.org/10.5120/ijca2017915495

[101] S. Hadri, Y. Naitmalek, M. Najib, M. Bakhouya, Y. Fakhri, and M. Elaroussi, "A Comparative Study of Predictive Approaches for Load Forecasting in Smart Buildings," *Procedia Computer Science*, vol. 160, pp. 173–180, 2019. [Online]. Available: https://doi.org/10.1016/j.procs.2019.09.458

[102] S. Smyl and N. G. Hua, "Machine learning methods for GEFCom2017 probabilistic load forecasting," *International Journal of Forecasting*, vol. 35, no. 4, pp. 1424–1431, 2019. [Online]. Available: http://www.sciencedirect. com/science/article/pii/S0169207019300172

[103] L. Hernandez, C. Baladrón, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, and J. Lloret, "Short-term load forecasting for microgrids based on artificial neural networks," *Energies*, vol. 6, no. 3, pp. 1385–1408, 2013. [Online]. Available: https://doi.org/10.3390/en6031385

[104] M. Ramezani, H. Falaghi, M.-R. Haghifam, and G. Shahryari, "Short-term electric load forecasting using neural networks," in *EUROCON 2005-The International Conference on " Computer as a Tool"*, vol. 2.   IEEE, 2005,

pp. 1525–1528. [Online]. Available: https://doi.org/10.1109/EURCON.2005.1630255

[105] I. P. Panapakidis, "Clustering based day-ahead and hour-ahead bus load forecasting models," *International Journal of Electrical Power & Energy Systems*, vol. 80, pp. 171–178, 2016. [Online]. Available: https://doi.org/10.1016/j.ijepes.2016.01.035

[106] S. Li, P. Wang, and L. Goel, "Short-term load forecasting by wavelet transform and evolutionary extreme learning machine," *Electric Power Systems Research*, vol. 122, pp. 96–103, 2015. [Online]. Available: https://doi.org/10.1016/j.epsr.2015.01.002

[107] A. Gulli and S. Pal, *Deep learning with Keras: implement neural networks with Keras on Theano and TensorFlow*, 1st ed. Birmingham, UK : Packt Publishing, 2017.

[108] N. Reimers and I. Gurevych, "Optimal Hyperparameters for Deep LSTM-Networks for Sequence Labeling Tasks," *CoRR*, vol. abs/1707.06799, 2017. [Online]. Available: http://arxiv.org/abs/1707.06799

[109] J. Moolayil, Moolayil, and S. John, *Learn Keras for Deep Neural Networks*. Springer, 2019.

[110] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for Activation Functions," *CoRR*, vol. abs/1710.05941, 2017. [Online]. Available: http://arxiv.org/abs/1710.05941

[111] Z. Chang, Y. Zhang, and W. Chen, "Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform," *Energy*, vol. 187, p. 115804, 2019. [Online]. Available: https://doi.org/10.1016/j.energy.2019.07.134

[112] H. HaddadPajouh, A. Dehghantanha, R. Khayami, and K.-K. R. Choo, "A deep Recurrent Neural Network based approach for Internet of Things malware threat hunting," *Future Generation Computer Systems*, vol. 85,

pp. 88–96, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167739X1732486X

[113] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016, pp. 7046–7051. [Online]. Available: https://doi.org/10.1109/IECON.2016.7793413

[114] P. A. Joshi and J. J. Patel, "Computational Analysis and Intelligent Control of Load Forecasting Using Time Series Method," in *Proceedings of the International Conference on Intelligent Systems and Signal Processing*. Springer, 2018, pp. 297–306. [Online]. Available: https://doi.org/10.1007/978-981-10-6977-2_27

[115] A. Narayan and K. W. Hipel, "Long short term memory networks for short-term electric load forecasting," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2017, pp. 2573–2578. [Online]. Available: https://doi.org/10.1109/SMC.2017.8123012

[116] L. Yang and H. Yang, "Analysis of different neural networks and a new architecture for short-term load forecasting," *Energies*, vol. 12, no. 8, p. 1433, 2019. [Online]. Available: https://doi.org/10.3390/en12081433

[117] H. Moutarde, P. Sznajder, and J. Wagner, "Unbiased determination of DVCS Compton form factors," *The European Physical Journal C*, vol. 79, no. 7, p. 614, 2019. [Online]. Available: https://doi.org/10.1140/epjc/s10052-019-7117-5

[118] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[119] G. Casella, S. Fienberg, and I. Olkin, "An Introduction to statistical learning with Applications in R," in *Springer Texts in Statistics*. Springer New York, 2013.

[120] F. Hutter, J. Lücke, and L. Schmidt-Thieme, "Beyond manual tuning of

hyperparameters," *KI-Künstliche Intelligenz*, vol. 29, no. 4, pp. 329–337, 2015. [Online]. Available: https://doi.org/10.1007/s13218-015-0381-0

[121] L. Iliadis, I. Maglogiannis, and V. Plagianakos, *Artificial Intelligence Applications and Innovations: AIAI 2018 IFIP WG 12.5 International Workshops, SEDSEAL, 5G-PINE, MHDW, and HEALTHIOT, Rhodes, Greece, May 25-27, 2018, Proceedings*.   Springer, 2018, vol. 520.

[122] Y. Gal and Z. Ghahramani, "A theoretically grounded application of dropout in recurrent neural networks," in *Advances in neural information processing systems*, 2016, pp. 1019–1027. [Online]. Available: http://papers.nips.cc/paper/ 6241-a-theoretically-grounded-application-of-dropout-in-recurrent-neural-networks. pdf

[123] J. Bergstra, D. Yamins, and D. D. Cox, "Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms," in *Proceedings of the 12th Python in science conference*.   Citeseer, 2013, pp. 13–20. [Online]. Available: https://doi.org/10.1088/1749-4699/8/1/014008

[124] M. Zahid, F. Ahmed, N. Javaid, R. Abbasi, H. Zainab Kazmi, A. Javaid, M. Bilal, M. Akbar, and M. Ilahi, "Electricity Price and Load Forecasting using Enhanced Convolutional Neural Network and Enhanced Support Vector Regression in Smart Grids," *Electronics*, vol. 8, no. 2, p. 122, Jan. 2019. [Online]. Available: http://dx.doi.org/10.3390/electronics8020122

[125] S. Li, L. Goel, and P. Wang, "An ensemble approach for short-term load forecasting by extreme learning machine," *Applied Energy*, vol. 170, pp. 22–29, 2016. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S0306261916302707

[126] S. R. Rallapalli and S. Ghosh, "Forecasting monthly peak demand of electricity in India – A critique," *Energy Policy*, vol. 45, pp. 516–520, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/ S0301421512001814

[127] M. J. O. Panão and M. C. Brito, "Modelling aggregate hourly electricity

consumption based on bottom-up building stock," *Energy and Buildings*, vol. 170, pp. 170–182, 2018. [Online]. Available: http://www.sciencedirect.com/ science/article/pii/S0378778817330797

[128] F. Amara, K. Agbossou, Y. Dubé, S. Kelouwani, A. Cardenas, and S. S. Hosseini, "A residual load modeling approach for household short-term load forecasting application," *Energy and Buildings*, vol. 187, pp. 132–143, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/ S0378778818309228

[129] J. G. Jetcheva, M. Majidpour, and W.-P. Chen, "Neural network model ensembles for building-level electricity load forecasts," *Energy and Buildings*, vol. 84, pp. 214–223, 2014. [Online]. Available: http: //www.sciencedirect.com/science/article/pii/S0378778814006458

Eidesstattliche Versicherung
(Affidavit)

Unterluggauer, Tim

Name, Vorname
(Last name, first name)

176702

Matrikelnr.
(Enrollment number)

Ich versichere hiermit an Eides statt, dass ich die vorliegende Bachelorarbeit/Masterarbeit* mit dem folgenden Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.	I declare in lieu of oath that I have completed the present Bachelor's/Master's* thesis with the following title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution.

Titel der ~~Bachelor~~/Masterarbeit*:
(Title of the ~~Bachelor's~~/ Master's* thesis):

Forecasting the aggregated charging load of electric vehicles at different charging sites

*Nichtzutreffendes bitte streichen
(Please choose the appropriate)

Tampere, 29.06.2020

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Tampere, 29.06.2020

Ort, Datum
(Place, date)

Unterschrift
(Signature)