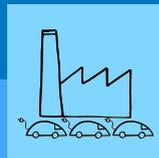
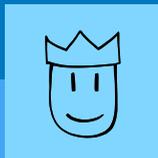


KARL SCHWENK

# A Smart Charging Assistant for Electric Vehicles Considering Battery Degradation, Power Grid and User Constraints





# **A Smart Charging Assistant for Electric Vehicles Considering Battery Degradation, Power Grid and User Constraints**

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)

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

**Dissertation**

von

M.Sc. Karl Schwenk

Tag der mündlichen Prüfung:	22.06.2022
Hauptreferent:	Prof. Dr.-Ing. Veit Hagenmeyer
Korreferenten:	Prof. Dr. rer. nat. Carsten Agert apl. Prof. Dr.-Ing. Ralf Mikut

04.07.2022, Weil der Stadt, Germany



This document—except for the cover—is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0):  
<https://creativecommons.org/licenses/by-sa/4.0/deed.en>

© Cover Design by Karl Schwenk, 2022



DOI: 10.5445/IR/1000148141





# Abstract

The rise of intermittent renewable power generation increasingly impedes an efficient and reliable utility grid operation. Simultaneously, the number of electric vehicles, which require significant amounts of electric energy to charge, is growing rapidly. The energy and mobility sectors are therefore inevitably coupled, implying that reliable electric mobility depends on robust power supply. Furthermore, vehicle users perceive a limitation of their individual mobility, as electric vehicles currently provide less driving range and require more time to recharge compared with internal combustion engine vehicles.

The present thesis therefore presents a novel concept and a software application (*charging assistant*) supporting users when charging their electric vehicles, while considering the interests of all involved stakeholders. To achieve this, design features of possible software architectures are initially compared to define a suitable structure of modules and their interconnection. Real data is then used to develop, enhance, and validate both energy consumption and battery models representing electric vehicle driving and charging characteristics. The major contributions of this thesis arise from design and validation of the charging assistant's following three core components.

First, users' individual mobility behavior is modeled and evaluated using both recorded and semi-synthetic electric vehicle trip data. Particularly, a novel two-stage clustering algorithm is developed to determine users' frequently visited locations. Ensembles of random forest models are then used to predict next locations to visit and typical parking times there.

Second, stochastic mixed-integer programming is used to most conveniently and cost-effectively schedule charging stops in a future time horizon. In this context, a graph-based algorithm helps to quantify energy demand and occurrence probability of an electric vehicle user's mobility scenarios. For validation, two alternative charging strategies are defined and compared with the proposed scheme.

Third, a non-linear optimization scheme is engineered to leverage unused time and energy flexibility in electric vehicle charging processes. Integrating a detailed elec-

tric vehicle battery model allows to precisely quantify cost savings due to reduced battery aging and dynamic electricity tariffs. Using data from real electric vehicle charging events, influences on the profitability of *vehicle-to-grid* applications can be elaborated. Implementing the proposed scheme in a realistic environment yields an architecture blueprint and communication concept for optimization-based smart charging systems. In this context, further challenges regarding standardized charging communication, interventions of power suppliers, and user acceptance are discovered.

# Kurzfassung

Der Anstieg intermittierender Stromerzeugung aus erneuerbaren Energiequellen erschwert zunehmend einen effizienten und zuverlässigen Betrieb der Versorgungsnetze. Gleichzeitig steigt die Zahl der Elektrofahrzeuge, die zum Aufladen erhebliche Mengen an elektrischer Energie benötigen, rapide an. Energie- und Mobilitätssektor sind somit unweigerlich miteinander verbunden, was zur Folge hat, dass zuverlässige Elektromobilität von einer robusten Stromversorgung abhängt. Darüber hinaus empfinden Fahrzeugnutzer ihre individuelle Mobilität als eingeschränkt, da Elektrofahrzeuge im Vergleich zu Fahrzeugen mit Verbrennungsmotor derzeit eine geringere Reichweite aufweisen und mehr Zeit zum Aufladen benötigen.

In der vorliegenden Arbeit wird daher ein neuartiges Konzept sowie eine Softwareanwendung (*Ladeassistent*) vorgestellt, die den Nutzer beim Laden seines Elektrofahrzeuges unterstützt und dabei die Interessen aller beteiligten Akteure berücksichtigt. Dafür werden zunächst Gestaltungsmerkmale möglicher Softwarearchitekturen verglichen, um eine geeignete Struktur von Modulen und deren Verknüpfung zu definieren. Anschließend werden anhand realer Daten sowohl Energieverbrauchs- als auch Batteriemodelle entwickelt, verbessert und validiert, welche die Fahr- und Ladeigenschaften von Elektrofahrzeugen abbilden. Die wichtigsten Beiträge dieser Arbeit resultieren aus der Entwicklung und Validierung der folgenden drei Kernkomponenten des Ladeassistenten.

Als Erstes wird das individuelle Mobilitätsverhalten der Nutzer modelliert und anhand von aufgezeichneten und halbsynthetischen Fahrdaten von Elektrofahrzeugen ausgewertet. Insbesondere wird ein neuartiger, zweistufiger Clustering-Algorithmus entwickelt, um häufig besuchte Orte der Nutzer zu ermitteln. Anschließend werden Ensembles von Random-Forest-Modellen verwendet, um die nächsten Aufenthaltsorte und die dort typischen Parkzeiten vorherzusagen.

Als Zweites wird gemischt-ganzzahlige stochastische Optimierung angewandt, um Ladestopps in einem zukünftigen Zeithorizont möglichst komfortabel und kostengünstig zu planen. Dabei wird ein graphenbasierter Algorithmus eingesetzt, um

den Energiebedarf und die Eintrittswahrscheinlichkeit von Mobilitätsszenarien eines Elektrofahrzeugnutzers zu quantifizieren. Zur Validierung werden zwei alternative Ladestrategien definiert und mit dem vorgeschlagenen System verglichen.

Als Drittes wird ein nichtlineares Optimierungsschema entwickelt, um vorhandene Zeit- und Energieflexibilität in Ladevorgängen von Elektrofahrzeugen zu nutzen. Die Integration eines detaillierten Batteriemodells ermöglicht eine genaue Quantifizierung der Kosteneinsparungen aufgrund einer geringeren Batteriealterung und dynamischer Stromtarife. Anhand von Daten aus realen Ladevorgängen von Elektrofahrzeugen können Einflüsse auf die Rentabilität von *Vehicle-to-Grid*-Anwendungen herausgearbeitet werden. Aus der Umsetzung des vorgestellten Ansatzes in einer realistischen Umgebung geht ein Architekturentwurf und ein Kommunikationskonzept für optimierungsbasierte intelligente Ladesysteme hervor. Dabei werden weitere Herausforderungen im Zusammenhang mit standardisierter Ladekommunikation, Eingriffen der Energieversorger und Nutzerakzeptanz aufgedeckt.

# Prologue

Since my earliest childhood memories, a strong fascination for science and technology has always shaped my perception. Whether it was huge machinery on construction sites that caught my attention during a walk with my grandmother, or chemical reactions that my grandfather showed and explained to me while I skipped kindergarten. Growing up in a rural area, I also soon realized the importance of mobility for people to achieve certain things. My fascination for cars, planes, trains, and boats was born and I could hardly wait to finally receive my own driver's license. After finishing school, all this technical enthusiasm led me—as a logical consequence—to study mechanical engineering; and it was amazing: intense math, scientific formulae, technical drawings, and theoretical background to all my childhood visions. Entering the automotive industry as an intern, I was finally able to apply all the knowledge I had acquired. At about the same time, I discovered another fascination. It was the magic of computer programs turning series of numbers and letters into valuable output, or even controlling those machines that already fascinated me. I guess at this time my career path was set: to develop technical software, preferably for cars. When I was given the opportunity to pursue a PhD in collaboration with the automotive industry, it felt like connecting all the elements: technology, mobility, software engineering, and science.

Preparing the present thesis, however, would have been impossible without having the support of many as a foundation. First of all, I would therefore like to give special thanks to my direct supervisor apl. Prof. Dr.-Ing. Ralf Mikut. Your time spent in numerous constructive discussions, and your encouragement, advice and support helped me develop both professionally and personally. I would also like to express my sincere thanks to Prof. Dr. rer. nat. Carsten Agert from the University of Oldenburg for reviewing this thesis and to Prof. Dr.-Ing. Christoph Stiller for chairing the examination committee. I would furthermore like to acknowledge the support of Prof. Dr.-Ing. Veit Hagenmeyer for the supervision and review of this thesis, for the opportunity to be part of the Institute for Automation and Applied Informatics, and for enabling real-world smart charging experiments in the *Energy Lab 2.0*. In the latter context I would also like to thank Dr.-Ing. Simon Waczowicz and his colleagues

for diligent hands-on support. Thanks also goes to all members of the Institute for Automation and Applied Informatics—especially my colleagues within the research area *Automated Image and Data Analysis*. Your supportive collaboration doubtlessly contributed to successfully completing this work. In this context, I would like to express my special thanks to Dr.-Ing. Riccardo Remo Appino. You always and tirelessly encouraged, mentored, and supported me, especially during my first months as a doctoral student.

While conducting the research and experiments of this thesis, I was employed at Mercedes-Benz AG in Sindelfingen, Germany. I would like to give special thanks to Dipl.-Ing. Frank Hepperle and Dr.-Ing. Tim Harr. You gave me the opportunity to be an appreciated member of your team, provided me honest and constructive feedback, and constantly supported my professional and personal growth. Likewise, I would like to thank all my former team colleagues and members of the *eDrive Innovations* department under the leadership of Dr.-Ing. Martin Hermsen. During the past years I also had the honor to supervise the work of students myself; a truly rewarding task that helped me grow and learn a lot about mutual collaboration. Many thanks therefore goes to M.Sc. Maximilian Hentsch, M.Sc. Stefan Meisenbacher, and M.Sc. Julian Frick.

Finally, I would like to express deep gratitude to my parents and family. Your comprehensive support and encouragement always enabled me to pursue my goals such as this thesis. Many thanks also goes to my friends and especially my girlfriend Mareike who always pushed my motivation in case I was struggling myself.

July 4, 2022, Weil der Stadt, Germany

*Karl Schwenk*

# Contents

<b>Abstract</b> . . . . .	<b>i</b>
<b>Kurzfassung</b> . . . . .	<b>iii</b>
<b>Prologue</b> . . . . .	<b>v</b>
<b>1 Introduction</b> . . . . .	<b>1</b>
1.1 State of the Art . . . . .	5
1.1.1 Grid Integration of Electric Vehicles . . . . .	5
1.1.2 Modeling of Electric Vehicles . . . . .	11
1.1.3 Electric Vehicle Users . . . . .	19
1.1.4 Design of Charging Assistants . . . . .	26
1.2 Open Questions . . . . .	28
1.3 Objectives . . . . .	29
<b>2 Concept and Architecture</b> . . . . .	<b>31</b>
2.1 Stakeholders . . . . .	34
2.2 Modules of an Automated Charging Assistant . . . . .	36
2.2.1 The Electric Vehicle as a Means of Transportation . . . . .	36
2.2.2 Predicting the Mobility of Electric Vehicle Users . . . . .	38
2.2.3 Scheduling of Charging Events . . . . .	39
2.2.4 Optimization of Charging Events . . . . .	40
2.3 Environment of Implementation . . . . .	42
2.3.1 Acquisition of Data . . . . .	43
2.3.2 Storage of Data . . . . .	44
2.3.3 Application of Data . . . . .	44
2.4 Conclusion . . . . .	46
<b>3 Modeling of Electric Vehicles</b> . . . . .	<b>47</b>
3.1 Consumption Model . . . . .	48
3.1.1 Selection and Preprocessing of Training Data . . . . .	51
3.1.2 Data-driven Models to Estimate Energy Consumption . . . . .	54

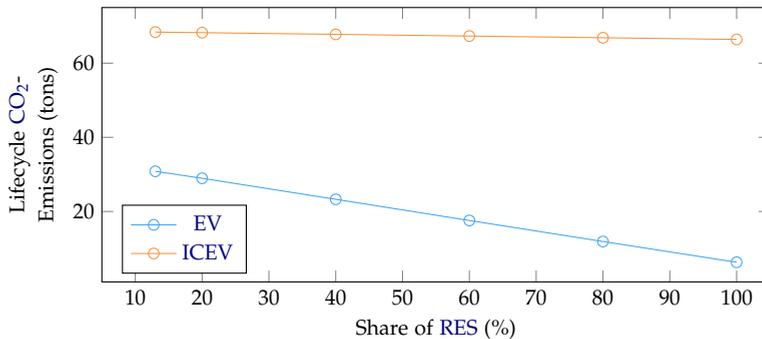
3.1.3	Validation of Consumption Model . . . . .	57
3.2	Battery Model . . . . .	61
3.2.1	Electric Battery Model . . . . .	62
3.2.2	Thermal Battery Model . . . . .	65
3.2.3	Battery Aging Model . . . . .	70
3.2.4	Validation of Battery Model . . . . .	72
3.3	Conclusion . . . . .	83
<b>4</b>	<b>Mobility Prediction . . . . .</b>	<b>85</b>
4.1	Available Data of Users' Mobility Behavior . . . . .	87
4.1.1	Data of Real Mobility Behavior . . . . .	87
4.1.2	Generation of Semi-Synthetic Data . . . . .	87
4.2	Determination of Users' Known Locations . . . . .	91
4.2.1	Data Cleaning . . . . .	91
4.2.2	Two-Stage Density-Based Spatial Clustering . . . . .	92
4.2.3	Evaluation of Clustering . . . . .	93
4.3	Prediction of Next Location . . . . .	97
4.3.1	Data Engineering for Next Location Prediction . . . . .	98
4.3.2	Models to Predict the Next Location to Visit . . . . .	100
4.3.3	Validation and Results . . . . .	102
4.4	Estimation of Parking Duration at Known Locations . . . . .	110
4.4.1	Data Engineering for Parking Duration Estimation . . . . .	110
4.4.2	Models to Estimate Parking Duration . . . . .	113
4.4.3	Validation and Results . . . . .	115
4.5	Conclusion . . . . .	122
<b>5</b>	<b>Charging Scheduling . . . . .</b>	<b>125</b>
5.1	Scenario Generation . . . . .	127
5.2	Optimization Problem . . . . .	131
5.2.1	Optimization Objective of Charging Scheduling . . . . .	132
5.2.2	Decision Variables . . . . .	133
5.2.3	State Variables . . . . .	134
5.2.4	Chance Constraints on Adapted Minimum Energy . . . . .	135
5.2.5	Reformulated Optimization Problem . . . . .	138
5.3	Simulation Setup to Validate Charging Scheduling . . . . .	140
5.3.1	Parameter Setting for Monte Carlo Simulation . . . . .	140
5.3.2	Alternative Strategies for Charging Scheduling . . . . .	144
5.4	Results of Validating Charging Scheduling . . . . .	146
5.4.1	Robustness through Adapted Minimum Energy . . . . .	146
5.4.2	Benefits over Alternative Strategies for Charging Scheduling . . . . .	149

5.5	Conclusion . . . . .	153
<b>6</b>	<b>Charging Optimization . . . . .</b>	<b>155</b>
6.1	Charging Optimization Scheme . . . . .	157
6.1.1	Objective of Charging Optimization . . . . .	158
6.1.2	Decision and State Variables . . . . .	159
6.1.3	Including Dynamics of the Electric Vehicle Battery . . . . .	160
6.2	Validation and Case Study . . . . .	161
6.2.1	Simulation of Historical Charging Events . . . . .	161
6.2.2	Real-World Implementation of Charging Optimization . . . . .	164
6.3	Results of Charging Optimization Case Study . . . . .	167
6.3.1	Comparison of Electric Vehicle Operating Cost . . . . .	167
6.3.2	Influence Factors on Vehicle-to-Grid Profitability . . . . .	168
6.3.3	Effects of Thermal Modeling . . . . .	173
6.3.4	Insights from Real-World Implementation . . . . .	175
6.4	Conclusion . . . . .	178
<b>7</b>	<b>Conclusion and Outlook . . . . .</b>	<b>179</b>
<b>A</b>	<b>Important Evaluation Metrics . . . . .</b>	<b>183</b>
A.1	Metrics to Evaluate Feature Performance . . . . .	183
A.1.1	Spearman Rank Correlation Coefficient . . . . .	183
A.1.2	Gini Importance in Random Forest Models . . . . .	184
A.2	Metrics to Evaluate Model Performance . . . . .	185
A.2.1	Accuracy . . . . .	185
A.2.2	Coefficient of Determination . . . . .	185
A.2.3	Absolute Error Metrics . . . . .	186
A.2.4	Squared Error Metrics . . . . .	186
<b>B</b>	<b>Discrete Dynamic Programming Solving Algorithms . . . . .</b>	<b>187</b>
	List of Abbreviations . . . . .	xi
	List of Symbols . . . . .	xiii
	List of Figures . . . . .	xvii
	List of Tables . . . . .	xxv
	List of Own Publications . . . . .	xxix
	Bibliography . . . . .	xxxix



# 1 Introduction

In the previous century, the eventual shortage of fossil fuels, such as gasoline and diesel, has been prophesied repeatedly [1]. Additionally, Carbon Dioxide ( $\text{CO}_2$ ) emissions from combusting these fuels are assumed to negatively affect the earth's climate [2]. These aspects heavily influence the current-day mobility system consuming more than 50 % of all worldwide allocated fossil fuels [3]. Electric mobility powered from sustainable, environmental-friendly energy sources constitutes a feasible alternative [4,5]. Some researchers even suppose a potential to mitigate anthropogenic climate change when employing (Battery) Electric Vehicles (EVs) combined with low-emission energy sources [6–8]. For instance, the life-cycle  $\text{CO}_2$  emissions of EVs could be decreased by 55 % compared with Internal Combustion Engine Vehicles (ICEVs), if charged with 13 % of Renewable Energy Sources (RESs). As shown in Figure 1.1, using 100 % RESs would even decrease life-cycle  $\text{CO}_2$  emissions of an EV by 90%.<sup>1</sup>

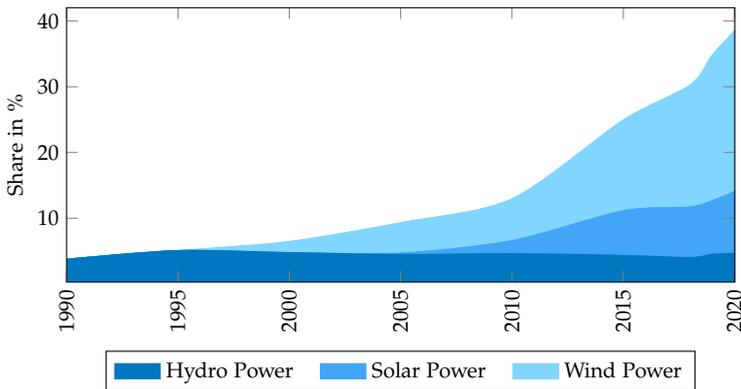


**Figure 1.1:** Comparison of life-cycle  $\text{CO}_2$  emissions of ICEV and EV for varying RES-share [6].

<sup>1</sup>Assuming a midsize EV with a life cycle mileage of 240 000 km, including production and scrapping.

However, the practical use of EVs in combination with renewable energies requires some basic conditions. First, a reliably available infrastructure providing sustainable energy to charge EVs. Second, EVs fulfilling the user requirements in terms of reliability, convenience, and value-stability. The latter, in turn, requires a sufficient understanding of behavioral motives and potential influence to define user requirements.

In terms of electric energy supply, nuclear power was assessed as putatively clean and reliable resource. The nuclear waste storage, however, poses a huge risk of environmental pollution. Furthermore, catastrophic incidents—both man-made (*Tschernobyl*, Ukraine 1986) and natural disasters (*Fukushima*, Japan 2011)—encouraged the political decision to exit nuclear power perennially in many nations, e.g., Germany [9]. Advancing environmental pollution due to fossil fuel emissions prompted governments to also abolish coal-fired power plants, e.g., Germany in 2020 [10]. These decisions call for a tighter coupling of energy sectors such as electricity, transport and heat demand and favor the rise of RESs [11–13]. Recently, an upheaval on the part of energy producers can thus be observed. Figure 1.2 shows that energy generation shifts towards wind, solar, and hydro power.



---

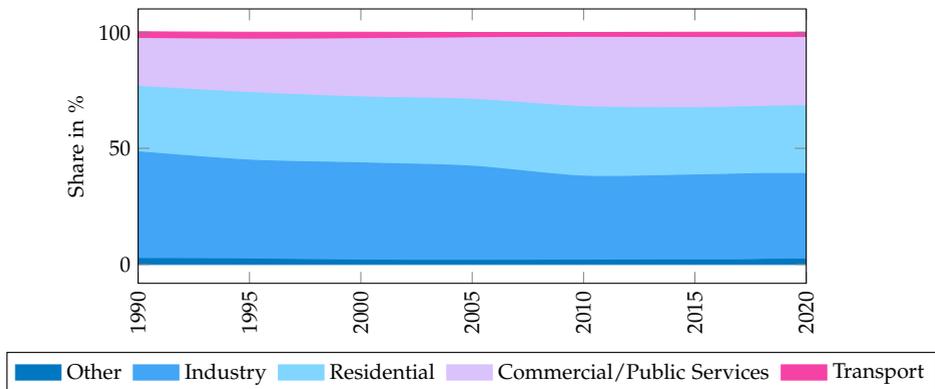
**Figure 1.2:** Relative share of hydro power, solar power, and wind power in total electricity generation in Germany, 1990 to 2020 [14].

Unlike stationary, centralized power plants, however, RESs distributedly inject energy into the grid [15]. Additionally, the power output of RESs depends on environmental conditions, e.g., solar radiation and wind. Increasing shares of RESs, hence, impede a reliable and predictable energy allocation [16]. The risk of local over- or

---

underproduction arises and needs to be handled accordingly, e.g., using adequate forecasting methods [17,18] together with (stationary) energy storages [19,20].

Increasing uncertainty also arises from energy consumers. Public services, residential areas, and industry dominated the hitherto electricity consumption. As Figure 1.3 shows, the transport sector, in contrast, only accounted for approx. 3% of all consumed electricity. With rising electric mobility, EV charging causes load profiles to become more fluctuating. Primarily, users' random behavior determines the points in time an EV will be charged. Furthermore, the power rates and energy amounts needed for EV charging exceed the residential levels by far. This may cause a congestion of the distribution grid [21–24]. Together, RESs and EVs' energy consumption complicate an efficient and reliable grid operation [25] and call for a coordination of EV charging processes [23,24].

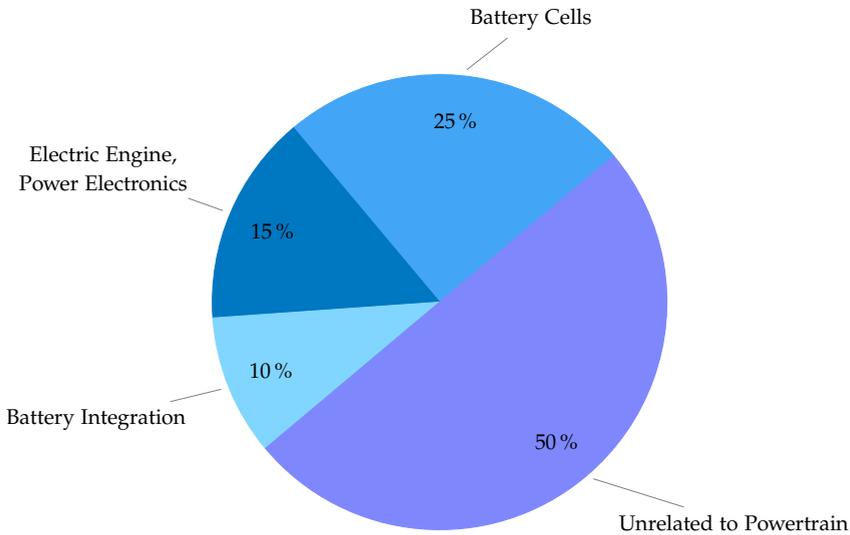


---

**Figure 1.3:** Electricity consumption in the European Union grouped by sectors; illustration of the transport sector's reluctant consumption, 1990 to 2020 [14].

With the market penetration of EVs, vehicle manufacturers also face new challenges [26]: The complexity of ICEVs was mainly based on mechanical components such as shiftable transmission or combustion engine. Even though mechanical components of EVs simplify, the complexity moves towards power electronics and energy storage. Low energy density of state-of-the-art battery systems, and thus high battery weight, requires increased construction safety. High power in- and outputs cause battery heating. According monitoring and cooling is required to reduce the risk of inflammation. Battery degradation, i.e., increased internal resistance and the loss of storage capacity, however, poses the possibly biggest challenge. Figure 1.4 shows a cost breakdown of EV manufacturing, with battery related costs account-

ing for half of it. Thus, battery degradation causes a significant value loss of the entire EV [27].



---

**Figure 1.4:** Breakdown of EV manufacturing cost focusing on powertrain-related components [28].

Insufficient user acceptance impairs the effectiveness of a strategy that combines the divergent interests of car manufacturers and power supply [29]. Using ICEVs over decades shaped the expectations of users towards EVs in terms of flexibility and reliability. Low driving range and long charging times of EVs, however, increase the manual effort of planning trips. For instance, estimating the energy consumption or searching for charging stations requires multiple information resources. Acquiring and combining this information is neither trivial, nor comfortable for users. A software tool considering individual user needs is inevitable to improve flexibility and reliability of EVs, and thus ensure convenient electric mobility. Therein, the secure storage and proper treatment of users' personal information need to be considered [30].

## 1.1 State of the Art

This section points out the context of the thesis by reviewing the state of the art and discussing related work. First, [Section 1.1.1](#) presents the setup of the electric utility grid, the effects of [RESs](#), and the integration of [EVs](#) into the grid. [Section 1.1.2](#) reviews essential models of [EV](#) batteries and compares different modeling approaches. Then, typical driving and charging habits of users, methods to predict mobility behavior, and user acceptance issues are described in [Section 1.1.3](#). At last, [Section 1.1.4](#) presents the combination of the primarily mentioned aspects in form of software applications (*charging assistants*) and architectural consideration therefore.<sup>2</sup>

### 1.1.1 Grid Integration of Electric Vehicles

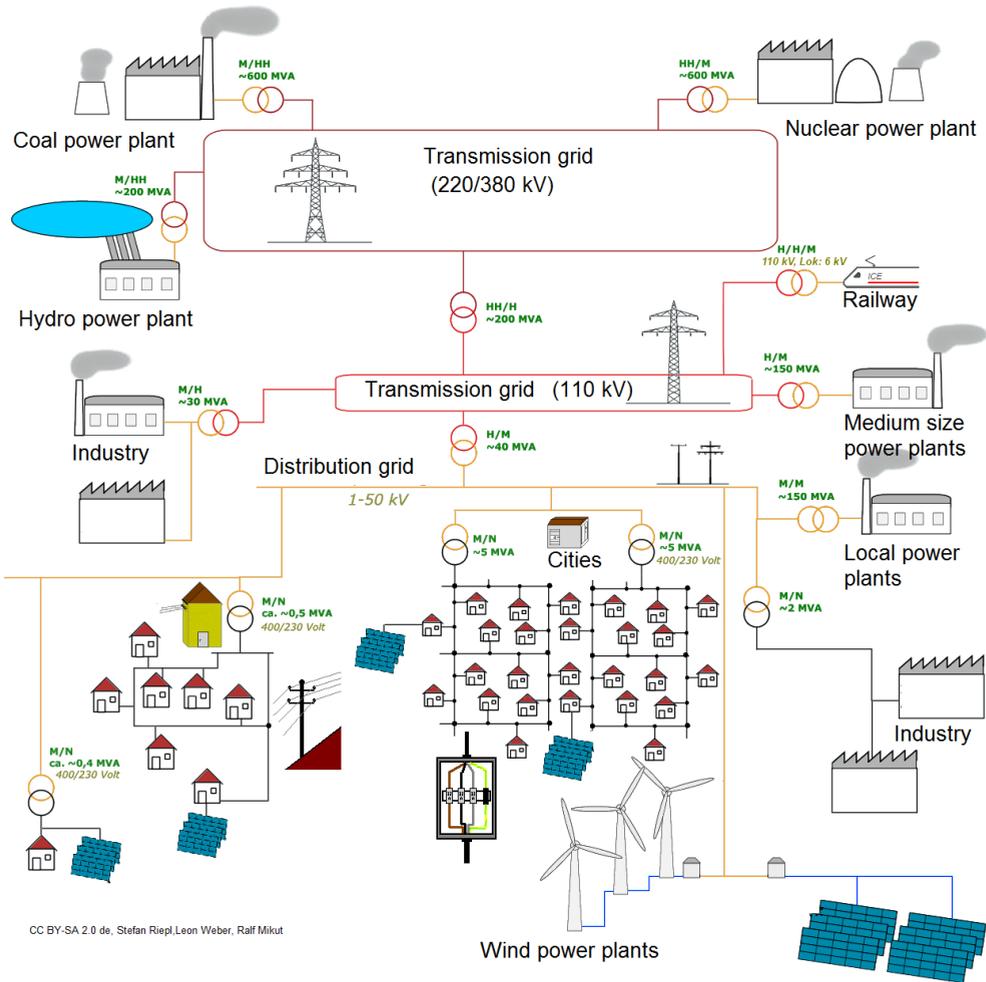
Today, electric energy is available as ubiquitous resource. For this, reliable and efficient transport of electricity constitutes a prerequisite [31]. Most utility grids all over the world, but especially in Europe have a hierarchical topology of different voltage levels, as seen in [Figure 1.5](#). Power transformer substations connect grid areas of different voltage among each other. Independent grid areas with the same voltage can be connected with switch substations. For areas with different or asynchronous frequencies, electric energy converters, e.g., motor-generator combinations or power electronics are used [31].

The **transmission grid** transfers power from centralized generation to areas of high consumption. In this context, *transmission grid operators* monitor power flow and maintain hardware. The *highest voltage* level (220-380 kV) takes in the energy production from hydro power generation, coal-fired and nuclear power plants. Prospectively, off-shore wind parks also contribute. Mostly, no consumers are directly connected to this voltage level. The *high voltage* level (110 kV) is the transmission grid's second voltage level. Mainly industrial power plants feed in energy. The consumers comprise large industry and railway transportation.

Once the energy is transmitted to high consumption areas, the **distribution grid** spreads energy within metropolitan areas. Here, *distribution grid operators* are in charge of monitoring and maintenance. Terrestrial wind parks, solar power plants,

---

<sup>2</sup>Regional differences, e.g., regarding electric utility grid, [EV](#) market penetration, or user behavior may exist. For the scope of this thesis, Europe, or particularly Germany is used as reference.



CC BY-SA 2.0 de, Stefan Riepl, Leon Weber, Raff Mikut

**Figure 1.5:** Schematic grid topology with transmission and distribution grids in different voltage levels, as well as energy producers and consumers, see also [31].

and municipal power plants inject energy to the *mid voltage* level (1-50 kV). This voltage level supplies most of the industry consumers and some remote private homes. The *low voltage* level (230-400 V) is the most subordinated level. It takes in energy generation from distributed RESs such as private solar panels and wind turbines. As the power output of these resources depends on volatile environmental conditions, e.g., solar radiation and wind, they pose the risk of local over- or underproduction. Private homes both in rural and urban areas account for the major consumption in this voltage level.

Nowadays, mainly residential and industry consumers dominate the energy demand. Their load profiles repeat in multiple time levels (time of day, day of week, month). With a growing market penetration of EVs, EV charging increasingly contributes to electricity consumption [32,33]. EV charging loads, however, are more volatile and less predictable. Primarily, users' behavior, determines the points in time an EV will be plugged in to charge. Due to *range anxiety*, users tend to charge their vehicles as often as possible and in an uncoordinated way [21,34].

Furthermore, the power rates and energy amounts needed for EV charging usually exceed the household levels by far. This might especially endanger the distribution grid's performance; in particular, the mid voltage level for fast charging at highways, and the low voltage level for residential areas. In already critical peak load situations, e.g., people coming home from work in the evening, while using several electric devices, the power demand concentrates. As few monitoring or controlling exists for these voltage levels, arising overload might not be detected instantaneously [35–37]. Power quality issues, e.g., voltage fluctuation or frequency instability could be the consequence. In the worst case, local distribution grid areas may break down due to overload. Combined with distributed RESs, an efficient and reliable grid operation becomes thus more complicated [25].

A growing number of EVs connected to the grid, may also enable grid-beneficial applications, often concluded as Vehicle-to-Grid (V2G) services [22,38–40]. For this, *smart charging*, i.e., to charge a defined amount of energy at a certain place and time, however, is a prerequisite [41]. Many V2G services base on bidirectional power flow and can provide additional revenue streams for charging station providers and/or EV users [5,36,42–46]. According to customer surveys, the concerns of users about V2G comprise advanced battery degradation and limited flexibility [47]. For receiving monetary rewards in return, however, users could accept providing their EVs for V2G services [48].

### 1.1.1.1 Vehicle-to-Grid Development

The evolution of V2G services can be classified in five phases as shown in Figure 1.6.<sup>3</sup> The already established **grid-compliant charging** bases upon the com-

---

<sup>3</sup>The specified time line only gives a rough estimation. Various factors may propel or delay this development.

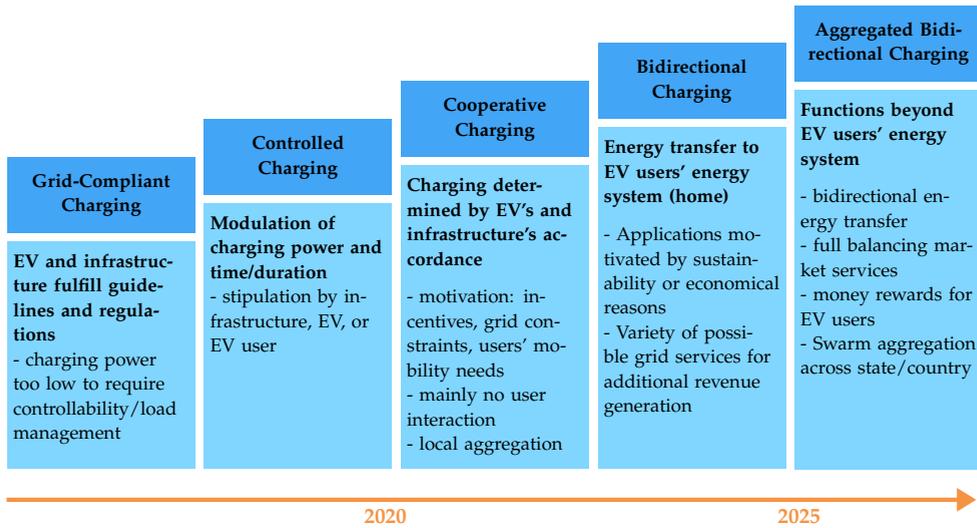


Figure 1.6: Road map of EV charging modes also considering a connection to the electric utility grid, approx. 2015 to 2025, according to [49].

pliance with guidelines and regulations. As charging power is either limited to un-critical levels or monitored decentralized, active controllability or centralized load management is not required.

The on-the-rise **controlled charging** is based on modulations of charging power or a temporal shift of the charging process, if applicable. Either infrastructure (charging station provider, grid operator), the EV, or the EV user provide necessary signals for the power modulation. Electricity only flows from the grid to the EV, not vice versa.

The principle of **cooperative charging** currently develops. EVs and charging infrastructure automatically negotiate and agree upon charging power, time and tariff. Active user interaction is mainly not required. Grid constraints, incentives for users and mobility needs of users motivate the negotiated charging strategy. A local aggregation of several EVs can be included, e.g., in parking lots.

In upcoming years, **bidirectional charging** can be expected. Therein, energy can also be transferred from EVs to the infrastructure. First, applications could support home energy systems of EV users motivated by sustainability or economic reasons, see e.g., [50]. Subsequently, a variety of services to the grid can offer additional revenue streams.

In the long term, **aggregated bidirectional charging** could be implemented. Bidirectional energy transfer mostly exceeds the own energy system of users and can be aggregated across states or countries. Thus, full balancing market services could be generated also enabling monetary rewards for EV users. The balancing of contradictory objectives regarding EV charging, however, requires advanced charging strategies. For instance, block-chain technology could help balance charging and discharging of multiple EVs. The *internet of energy* could provide a suitable communication platform, respecting security and privacy issues. For instance, secure data transmission, centralized architectures with mutual authentication, and individual remote access for payments [30,51].

### 1.1.1.2 Vehicle-to-Grid Services

The applications based on grid-connected EVs—V2G services—can be divided in five major categories as shown in Table 1.1 [52,53].

**Table 1.1:** Literature review on V2G services grouped by categorical characteristics of the services [53]

Category	Reference	Comments
Frequency Regulation	Wu et al. [54]	two-stage stochastic programming
	DONADEE et al. [55]	stochastic dynamic programming
Spinning Reserves	RAHMANI-ANDEBILI et al. [56]	optimal fleet management
	ALIASGHARI et al. [57]	micro-grid demand response
Power Quality	BEAUDE et al. [58]	voltage regulation
	KNEZOVIC et al. [59]	voltage support by EVs and solar power
	RAJAEI et al. [60]	ramp provision cost reduction
Energy Arbitrage	DENHOLM et al. [61]	system impacts of energy arbitrage
	MELO et al. [62]	energy arbitrage with battery degradation
	GAN et al. [63]	demand-side management
	ERDEN et al. [64]	peak shaving
	KHEMAKHEM et al. [65]	flattening power load curves
Renewable Energy Integration	ASHIQUE et al. [66]	solar-powered fast chargers
	LUND et al. [67]	minimize wind generation curtailment
	GHOFRANI et al. [68]	resource management strategies
	APPINO et al. [16]	reliable dispatch of RES
	NGUYEN et al. [69]	managed charging with user information
	DALLINGER et al. [70]	price-responsive EVs
	MOZAFAR et al. [71]	optimal RES allocation
SALAH et al. [72]	price-based RES integration via EV charging	

Some of them, e.g., energy arbitrage or power quality services, require mandatory bidirectional power flow, i.e., energy feedback from the EV to the grid. Within this context, dynamic electricity tariffs and market structures serve to monitor and control V2G services [72–75]. Energy price models help to calculate appropriate electricity prices, e.g., using stochastic programming [76,77]. The quantification of the impacts of V2G services on the power supply requires appropriate metrics. The literature recommends for instance i) the change in the minimum net load, ii) the change in cumulative negative net load, or iii) the residual load change [70].

Mismatching electric energy generation and loads cause the system frequency to deviate from the nominal value [53]. Controlled energy injecting or withdrawal—**frequency regulation**—can eliminate these deviations. The associated market requires energy storage systems to react within a few minutes. These technical qualifications are well within the specifications of EV batteries [78]. Despite high market clearing prices, increased battery degradation due to high energy throughput, i.e., cyclic battery aging, limits profitability. Hence, battery degradation costs often exceed the remuneration of frequency regulation.<sup>4</sup>

**Spinning reserves** provision extra generation to provide power in case of a sudden loss of generation [53]. To attain permanent service, a combination with stationary energy storages, e.g., large capacitors, seems conceivable [79]. Dispatch signals are issued occasionally and not on a regular schedule (20-50 times per year [80, 81]). The infrequent dispatch calls cause low additional energy throughput of the EV battery, i.e., low cyclic aging. As remuneration is paid for having capacity on hand, regardless of it being called, spinning reserves yields high revenue per energy throughput [82].

**Power quality** services correct voltage and current deviations from an ideal sinusoidal profile with precise power injections or withdrawals on sub-cycle timescales. Such deviations can negatively affect the operation of electrical devices. The high-dynamic and easy controllable power characteristics of EV batteries grant EVs a promising role therein [58,59].

**Energy arbitrage** services leverage energy price spreads.<sup>5</sup> Presupposed a sufficient price difference to exceed round-trip energy losses, trading fees, and operational

---

<sup>4</sup>The EV battery is assumed to be property of the remuneration recipient, e.g., the EV user.

<sup>5</sup>Energy arbitrage describes economically motivated V2G services solely based upon energy acquisition and disposal according to temporal or spatial price discrepancies. Price signals can be motivated by applications such as load balancing or peak shaving.

cost of the energy storage, revenue can be generated. Despite its simplicity, grid-beneficial energy arbitrage requires a suitable control of EV charging [70]. Energy arbitrage can be proceeded on a retail level, e.g., vehicle-to-building applications, or on a wholesale level. In the latter case, the charging demand of several EVs can be aggregated and traded on the wholesale energy market.

**Renewable energy integration**, e.g., powering EVs entirely from RESs, could enable an environmental-friendly and sustainable energy system [6]. By means of suitable optimization algorithms [83], grid-responsive EVs can help to integrate RESs into the electricity system [84, 85]. For instance, balancing the variable power output of RESs by controlled charging and discharging of EVs can yield a power transfer efficiency of up to 92% [86]. The authors of [68] develop resource management strategies to enable V2G in support of RES integration. Another approach for reliable dispatch of RES generation by means of EVs is presented in [16]. More specific approaches, e.g., solar-powered fast chargers [66] or minimize wind generation curtailment [67], could remediate distribution system impacts.

Note that all five aforementioned V2G service categories mainly possess a Technology Readiness Level (TRL) according to MANKINS [87] between 2 (“Technology concept and/or application formulated”) and 5 (“Component and/or breadboard validation in relevant environment”).

### 1.1.2 Modeling of Electric Vehicles

Unlike ICEVs, the complexity and major cost of EVs arise from energy storage and power electronics. For the energy storage, two partly competing technologies exist. Fuel Cell Electric Vehicles (FCEVs) consuming hydrogen can be refueled fast and operated locally emission-free. Producing hydrogen from natural gas reformation, however, depends on fossil resources, e.g., methane. Water electrolysis would be an alternative, but has low efficiency [88]. Furthermore, the low energy density of gaseous hydrogen requires energy and cost-intensive compression and storage.<sup>6</sup> Thus, FCEVs only yield about half the efficiency of EVs using electro-chemical storages, i.e., battery-electric systems. When directly using electricity to drive, no inefficient conversion of chemical to electrical energy is required. Furthermore, the losses

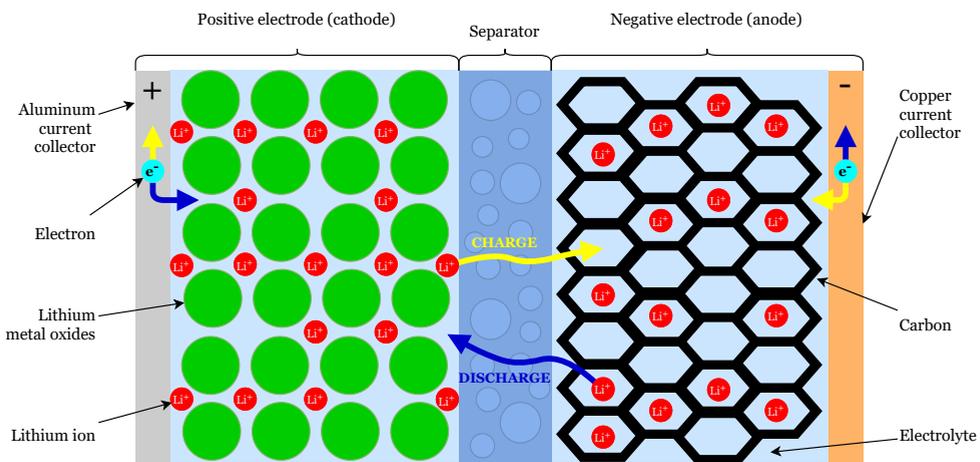
---

<sup>6</sup>Given the physical limitations of that issue, future research is not expected to bring significant advantages.

during transport and refueling are significantly lower [88]. Especially mobile applications, e.g., EVs that require frequent recharging, qualify for batteries as electric energy storages. This thesis therefore takes a closer look at the latter concept.

### 1.1.2.1 Lithium Ion Batteries

Most state-of-the-art batteries use lithium as active material, which theoretically allows higher energy density compared with liquid fossil fuels. Figure 1.7 illustrates a schematic view of a lithium ion battery [89].



**Figure 1.7:** Schematic section view of lithium ion battery. Positive electrode ( $\text{LiCoO}_2$ ) releases electrons when charging; negative electrode (graphite) releases electrons when discharging; lithium ions diffuse through the separator accordingly [89].

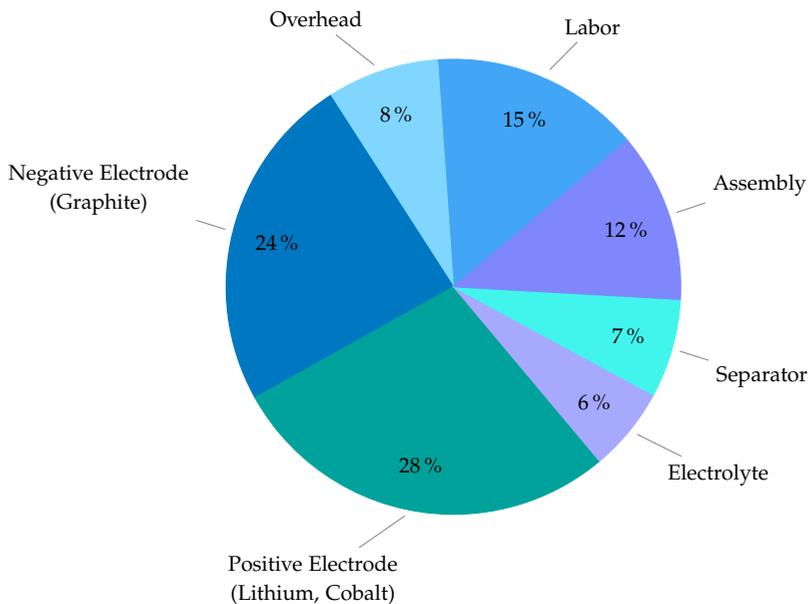
Battery cells as used in EVs generally consist of two compartments divided by a **separator**. **Current collectors** line the outsides of the two compartments (positive: aluminum, negative: copper).

The inner of the **positive electrode** consists of lithium active material, e.g., Lithium-Cobalt-Oxide ( $\text{LiCoO}_2$ ). Therefore, most batteries contain cobalt, which appears critical regarding environmental pollution and raw material scarcity. Emerging battery recycling, however, will improve resource availability, while current research could allow to entirely replace cobalt in future batteries [90]. The **negative electrode** contains a carbon-based material that can store positive lithium ions. The inner of the battery cell is filled with a liquid-state **electrolyte** that appears critical

in terms of flammability and toxicity. Using solid-state electrolytes in the future, however, could eliminate these issues [27,91].

When charging, electrons from the positive electrode's active material drain off through the aluminum current collector, see the yellow arrows in Figure 1.7. The lithium ions in the positive electrode are pushed through the separator into the carbon material. Simultaneously, the copper current collector releases electrons into the carbon material. In discharging mode, the charging process is inverted as indicated by the blue arrows in Figure 1.7. The positive electrode absorbs electrons that flow into the active material. Positive lithium ions are pushed through the separator, while the negative electrode releases electrons towards the copper current collector.

In terms of cost, the battery accounts for about half of the vehicle's value. About two thirds of the battery cost arise from active components such as electrodes, electrolyte, and separator, as shown in Figure 1.8. However, a maturing EV market and battery technology improvements might cause battery production prices to decrease within the next years [27,92,93].



**Figure 1.8:** Breakdown of battery manufacturing cost comprising raw material values, labor and overhead costs [27].

Aboard the EV, the Battery Management System (BMS) supervises the operation of the EV battery. This includes functions that monitor and control the battery while charging or discharging. If a safety-critical situation is likely, a warning system alerts the user, or the safety control shuts down the battery system entirely. Furthermore, the battery state, i.e., the State of Charge (SOC), the State of Health (SOH), and the internal battery temperature are estimated [94]. Therefore, appropriate models, i.e., an **electrical model**, a **thermal model**, and a **degradation model**, are required [95].

### 1.1.2.2 Electrical Battery Models

To describe the charging and discharging characteristics of batteries, an electrical model is required. In this context, the momentary battery voltage  $U_{\text{bat}}$  and current  $I_{\text{bat}}$  can be measured. To quantify the battery level, the battery energy  $e$  is used, usually represented by the dimensionless SOC

$$S_C = \frac{e}{e_{\text{nom}}}, \quad (1.1)$$

with the nominal storage capacity  $e_{\text{nom}}$ .<sup>7</sup> For real-time SOC estimation, a variety of methods exist as shown in Table 1.2 [96].

For  $I_{\text{bat}} = 0$ , the SOC of a battery correlates with the open circuit voltage  $U_{\text{OCV}}$  and the battery temperature  $\theta$ , see Figure 1.9. These relationships can be mapped to the SOC and stored into **look-up tables**. Consequently, if the external parameters are measurable, the battery SOC can be inferred by consulting this table [97,98]. However, a precise  $U_{\text{OCV}}$  measurement requires cutting off the power and having the battery rest for an extended period to attain electro-chemical equilibrium [110]. Thus, the method's applicability in EVs is limited [110,111].

**Coulomb counting** bases on integrating all currents  $I_{\text{bat}}$  flowing into or out of the battery over time [110]. Presuming the initial SOC is known, and  $I_{\text{bat}}$  is precisely measurable, the remaining SOC can be calculated. Inevitable current measurement errors, however, compromise the precision of the SOC estimation. Furthermore, the method requires a capacity recalibration according to the battery's operating and aging condition [111,112]. Combining coulomb counting with look-up tables [113] or physical models [100,101] helps to overcome these drawbacks.

---

<sup>7</sup>In the literature, formulation (1.1) is sometimes also referred to as the battery's *state of energy*.

**Table 1.2:** Literature review on electric battery models grouped by model categories.

Category	Reference	Comments
Look-up Tables	HE et al. [97]	look-up table for open-circuit voltage
	XING et al. [98]	open-circuit voltage with temperature influence
Coulomb Counting	KUSTIMAN et al. [99]	recursive least square estimation
	JUANG et al. [100]	with physical thermal model
	HE et al. [101]	real-time error correction
Electro-chemical Model	RAHMAN et al. [102]	parameter identification, particle swarm optimization
	ZHENG et al. [103]	electro-chemical-enhanced equivalent circuit
Equivalent Circuit Model	ZHANG et al. [104]	resistor-capacitor-network
	LI et al. [105]	distributed parameter model
Data-driven Estimation	DENG et al. [106]	support vector machine
	GONG et al. [107]	neural network (battery pack modeling)
	WANG et al. [108]	neural network (thermo-coupled SOC estimation)
	LEE et al. [109]	growing hierarchical self-organizing maps (with aging)

**Electro-chemical models** describe the mass, energy, moment, and momentum transport in each phase and component of a battery cell [102]. Thus, macroscopic quantities, e.g., cell current and voltage can be derived from kinetic and charge transfer processes on a microscopic scale. The model complexity, however, restricts a direct use in **BMSs** [102]. For this reason, electro-chemical models are mainly used to infer design rules in battery research [111].

**Equivalent circuit models** aim to describe batteries' electrical behavior with components such as resistors, capacitors, or voltage sources in circuit networks [101]. Solving the state equations of the circuit network allows to determine the **SOC** evolution for given voltage and current profiles. Due to the simple model structure and few model parameters, real-time applications often utilize equivalent circuit models [104,111]. Figure 1.9 shows an equivalent circuit that consists of a voltage source  $U_{OCV}$  serially connected to the internal resistance  $R_i$  [114]. This circuit suitably represents low-dynamic or stationary battery operation, e.g., **EV** charging.

Complex chemical processes and varying operating conditions often reduce the accuracy of above-mentioned methods. A broader applicability can be achieved with **data-driven estimation**. Despite the huge variety of model types, in the literature mostly neural networks [107,108] and support vector machines [106] are used. Independent of the model type, the model reconstructs the inner system behavior by

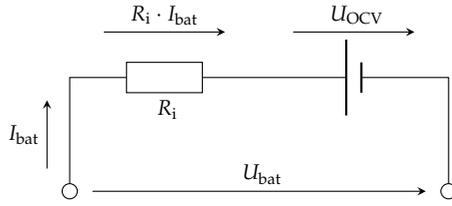


Figure 1.9: Equivalent circuit network of EV battery for low-dynamic or stationary operation [104].

mapping observed outputs for various inputs. Achieving a high accuracy at varying conditions, however, requires highly diverse training data [109–111, 115].

### 1.1.2.3 Thermal Battery Models

The inner battery temperature  $\theta$  influences battery performance, degradation, and the risk of inflammation. To monitor  $\theta$ , BMSs utilize thermal battery models [116]. The literature differentiates three model types, as Table 1.3 shows.

Table 1.3: Literature review on thermal battery models grouped by model categories.

Category	Reference	Comments
Heat Generation Model	DAI et al. [117]	adaptive KALMAN filter
	YOO et al. [118]	internal resistance derived from electro-chemical kinetics
Heat Transfer Model	GUO et al. [119]	finite element method
	PANCHAL et al. [120]	infrared analysis for thermal safety
Reduced-order Model	MURATORI et al. [121]	internal temperature estimation of battery pack
	GAMBHIRE et al. [122]	electro-chemical thermal response

**Heat generation models** describe the sources of heat within the battery during operation. The Joule heating mechanism causes the largest source of heat due to OHMIC losses at the internal resistance of the battery. Furthermore, exothermic or endothermic electro-chemical reactions generate heat sources or sinks [123]. Exemplary approaches assess the batteries’ real-time heat generation [117], or incorporate electro-chemical kinetics for higher accuracy [118].

Inside the battery, generated heat mostly distributes non-uniformly. The ability to measure the internal battery temperature, however, is limited. **Heat transfer models** help to avoid overheating of certain areas (*hot spots*) that might cause increased

battery degradation or risk of inflammation. To this end, heat convection, conduction, and radiation within and outside the battery need to be considered. These mechanisms can then be concluded in a three dimensional, distributed-parameter model as in [119]. To assess the risks of battery overheating, infrared analysis can be helpful [120].

**Reduced-order models** describe the simplified thermal battery behavior, i.e., heat generation and transfer, on a macroscopic level. Simplifications could be a homogeneous temperature distribution within the battery, or the conclusion of internal losses into one internal resistance [124]. Therefore, the electrical power loss during charging and discharging is assumed to be completely dissipated into heat. A second heat flow comprises all heat exchange between the battery and the environment. To maintain the battery temperature within a range that guarantees efficient and safe operation, the BMS may actively inject or withdraw additional heat. Using the battery's heat capacity, the differential temperature evolution of the battery can be derived from the energy balance. Due to their lower computational complexity these simplified models are often used within BMSs, presupposed their accuracy suffices [121, 122].

#### 1.1.2.4 Battery Degradation

About half of EVs' total cost is accounted for by electrical components, i.e., the battery and power electronics. Especially battery degradation (also *battery aging*) dominates the depreciation of the entire EV [125]. The mechanisms causing battery degradation comprise i) loss of lithium active material, i.e., lithium irreversibly reacting with other components, ii) lithium crystallization on the negative electrode (*lithium plating*), and iii) film formation by reaction products depositing on active material. While the first mechanism decreases the capacity available for energy storage, the latter two cause increased internal resistance and hinder diffusion [125, 126]. One differentiates *cyclic aging* and *calendar aging*. Cyclic aging is mainly caused by charging or discharging of the battery. Calendar aging, i.e., a temporal decay of the battery material, is promoted by high battery temperatures, high SOCs, over-voltage, and mechanical vibrations. To quantify battery degradation, one refers to the SOH

$$S_H = \frac{e_{\max}}{e_{\text{nom}}}, \quad (1.2)$$

as the relative loss of storage capacity.<sup>8</sup> For measuring or estimating  $S_H$ , in the literature different methods have been developed, as shown in Table 1.4 [127].

**Table 1.4:** Literature review on battery degradation grouped by model categories.

Category	Reference	Comments
Model-free	ROSCHER et al. [128]	standard capacity test
	COLEMAN et al. [129]	pulse current test
	JIANG et al. [130]	electro-chemical impedance spectroscopy
Model-based	BI et al. [131]	particle filter
	GHOLIZADEH et al. [132]	KALMAN filter
	ZHENG et al. [133]	analyses of charging curves, particle swarm optimization
	PETIT et al. [134]	thermo-coupled empirical model
Data-driven	NUHIC et al. [135]	lithium ion cell load collectives, support vector machine
	HU et al. [136]	particle swarm optimization, k-nearest-neighbor regression
	YOU et al. [137]	real-time SOH estimation, neural network
	NG et al. [138]	remaining useful life prediction, naive BAYES
	HU et al. [139]	sample entropy of battery voltage, BAYES learning
	HUANG et al. [90]	discharging voltage, unit time voltage drop, linear regression
	KLASS et al. [140]	on-board SOH estimation method, support vector machine
	EDDAHECH et al. [141]	recurrent neural networks with impedance spectroscopy
	SUSILO et al. [142]	GAUSSIAN kernels, support vector machine
LI et al. [143]	ensemble learning	

Those methods help to estimate battery degradation costs, or determine a reasonable point in time, the battery should be replaced.<sup>9</sup>

**Model-free** methods base upon direct or indirect measurements of characteristic battery quantities. Examples are the standard capacity test [128], the pulse current test [129], or the electro-chemical impedance spectroscopy [130]. Due to extensive measurement equipment and difficult execution, those methods are usually used in an laboratory environment and disqualify for the usage aboard an EV.

<sup>8</sup>Alternatively, the relative increase of the internal resistance could be used.

<sup>9</sup>Model characteristics and parameters originate from comprehensive cells tests and therefore are mostly confidential.

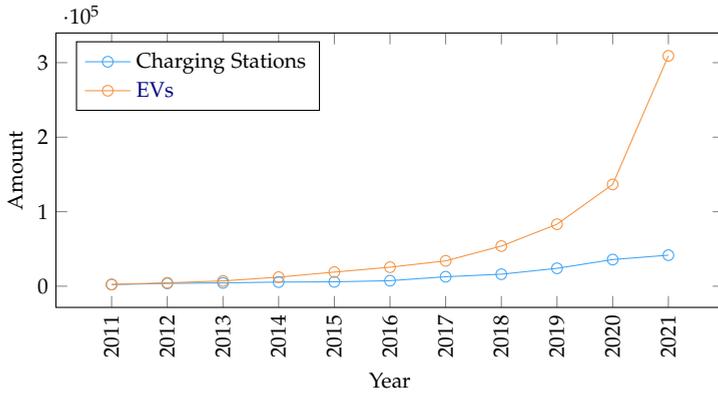
**Model-based** methods generally base upon empirical laboratory tests of single battery cells. The battery's physical characteristics, manifested in electrical or electrochemical models, change throughout battery degradation. Observing the temporal change thereof, allows to infer the actual SOH. The time-varying parameters of these models can be estimated e.g., by means of particle filters [131], or KALMAN filters [132]. As the models usually are based on observations of single battery cells, transferring the results on entire battery packs, as used in EVs is limited.

The complexity of model-free measurement methods, and the limited generalization of model-based methods can be overcome by using **data-driven** approaches. According models correlate accumulated aging influences with the momentary SOH. Existing approaches utilize i) support vector machines [135,140], ii) k-nearest-neighbor regression [136], iii) Linear Regression (LR) [90], iv) Artificial Neural Networks (ANNs) [137], v) and BAYES learning [138,139]. Depending on the data used for training, environmental influences, or real degradation behavior, e.g., of an entire battery pack, can also be represented. As no additional measurement equipment is required, real-time SOH estimations are also possible.

### 1.1.3 Electric Vehicle Users

From a user's perspective, using an EV differs from the usage of an ICEV. While an ICEV can be easily refueled in short time, charging an EV consumes more time. The current-day availability of charging stations is also more limited compared with the number of gas stations. While the number of registered EVs constantly increases, the public charging infrastructure limps behind, as shown in Figure 1.10 [144,145].

In addition, current-day EVs generally have a lower driving range than ICEVs. Charging is thus required more often. Altogether, those circumstances cause EV users to develop *range anxiety*, the fear of being stuck on the road with an empty battery and no charging station in reach [36,146]. Overcoming this fear requires users to adapt their behavior in a certain manner. Future trips need to be planned in advance. This requires the combination of several sources of information and is not trivial for users. For instance, the search for charging stations, estimating energy consumption, and considering charging stops in travel time. The results of



**Figure 1.10:** Number of public charging stations and disproportionately growing amount of registered EVs in Germany, 2011-2021 [144, 145].

recent customer surveys [48, 147, 148] support this effect.<sup>10</sup> Table 1.5 presents users’ relevant concerns and according expectations towards future EVs.

**Table 1.5:** Concerns and future expectations towards EV usage according to recent customer surveys [48, 147, 148].

Concern	Details	Future expectations
Reliability	mistrust in driving range	improved/individualized remaining range estimation
	low driving range	increased battery capacity (> 400 km)
Convenience	long charging duration	advanced fast charging ability/infrastructure
	effort for trip preparation	smart navigation, i.e., charging assistant
	complicated pay systems	standardized/centralized payment platform
	poor charging infrastructure	accessible information on available charging stations, reservation functionality
Costs	missing knowledge on battery care	detailed guidance on appropriate battery usage
	unknown cost of ownership	transparency of depreciation/amortization

<sup>10</sup>Survey [48]: 157 participants from Germany, USA, China, and Great Britain. Survey [147]: 782 participants from Germany, USA, China, Great Britain, Japan, and South Korea. KÜHL et al. [148]: evaluation of customer feedback on EVs via Twitter.

Among the surveyed participants the majority specified **reliability**, **convenience**, and **costs** as momentous concerns when using an **EV**. For instance, users expect a driving range of more than 400 km and a battery that lasts more than 8 years before total degradation.<sup>11</sup> In return, users are also willing to accept a surcharge of 13% to 15% for an **EV** over an **ICEV**. Although advancing technology and increasing market penetration of **EVs** are expected to mitigate most concerns in [Table 1.5](#), other challenges may arise that **EV** users are not yet aware of. For instance, occupied fast charging infrastructure on highly frequented travel days.

To encounter these issues and support users with **EV** usage on a daily basis, one needs to understand users' **driving and charging habits**, aspects of **user acceptance**, and methods for **mobility prediction**. [Table 1.6](#) shows related work elaborating these topics.

### 1.1.3.1 Driving and Charging Habits

The habits characterizing driving can be differentiated between a macroscopic purpose and a subordinated driving style. According to recent customer surveys [48], the major purposes of **EV** comprise private trips, commuting, and business trips. While exclusively private purposes account for 41% of all driven trips, only 17% of the total distance is caused by daily commutes to work. Trip with exclusive occupational purpose account for 12%. Evaluating the daily distance driven by an average vehicle user, see [Figure 1.11](#), reveals most of the trips to be shorter than 50 km. This distance is within the range of most current **EV** batteries [171]. Hence, the phenomena of range anxiety appears to be mainly psychologically enforced [36].

Driving style can e.g., be quantified via the vehicles' speed, steering wheel angle, and longitudinal and lateral acceleration [150]. Existing evaluations show that driving styles vary broadly, while e.g., a correlation to the drivers' age exists. Younger people drive more aggressively, e.g., accelerate harshly [149], while older people drive more efficiently, e.g., with moderate speed [152]. Aggressive driving style negatively affects the vehicles' efficiency and wear. As users accustomed to **ICEVs** transfer their driving style on **EV** usage, both maximum driving range and battery lifetime decrease [151, 172]. For instance, users that brake excessively instead

---

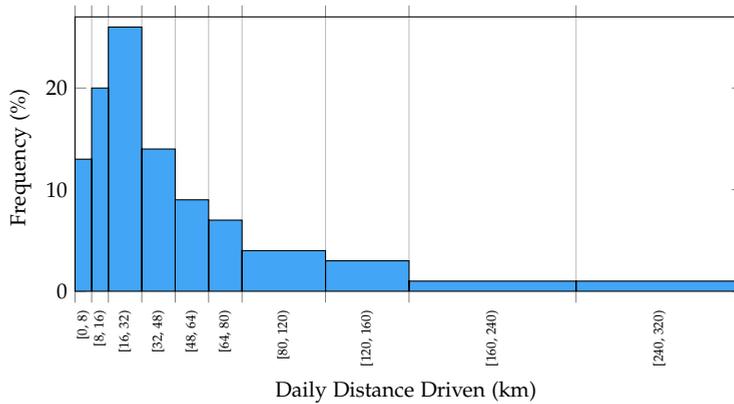
<sup>11</sup>**EV** batteries are assumed to fulfill automotive requirements as long as the nominal energy capacity has faded less than 20%, i.e., SOH  $\geq$  80 %.

**Table 1.6:** Literature review on EV usage, i.e., driving and charging habits, mobility prediction, and user acceptance.

Category	Reference	Comments
Driving/Charging Habits	ABU-BAKAR et al. [149]	driver behavior classification, deep learning
	DONG et al. [150]	driver classification, deep learning
	LI et al. [151]	two-level driving pattern clustering
	KNOWLES et al. [152]	driving style effects in EV efficiency
	FARIA et al. [153]	comparison charging scenarios
	AZADFAR et al. [154]	classification charging behavior
	NETZE BW AG [155]	small scale field experiment, charging evaluation
	DESAI et al. [156]	charging pattern analysis
	SCHMIDT et al. [157]	charging behavior at travel destinations
GOEBEL et al. [158]	driving behavior forecasting for V2G	
Mobility Prediction	GAMBS et al. [159]	next place prediction, MARKOV chain
	ZHANG et al. [160]	time-related route prediction, ensemble learning
	RATHORE et al. [161]	trajectory prediction framework
	KOOLWAL et al. [162]	trajectory prediction, possible use cases
	LIU et al. [163]	next location prediction, spatio-temporal recurrent ANN
	ZHANG et al. [164]	next location recommendation, hidden content, meta-data
	FRENDO et al. [165]	prediction of EV departure time
	ZONG et al. [166]	prediction of trip destinations
	ZHAO et al. [167]	prediction of individual trip properties
ROSSI et al. [168]	prediction of taxi trip destinations	
User Acceptance	EIDER et al. [36]	individualized incentivization methods
	BROADBENT et al. [169]	EV penetration influences, incentive recommendations
	HUBER et al. [170]	user acceptance smart charging, possible incentives
	WILL et al. [29]	influence factors on user acceptance of smart charging

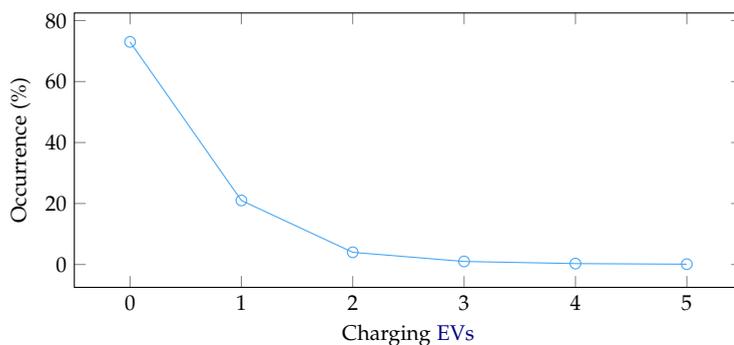
of coasting disregard EVs' possibility of recuperation.<sup>12</sup> In contrast, anticipatory driving yields a much higher efficiency.

<sup>12</sup>Coasting is the disengagement of wheels and powertrain during deceleration, in order to decrease drag torque. Recuperation is regenerating energy while decelerating, using the electric engine as generator.



**Figure 1.11:** Histogram illustrating the frequency of daily driven distance [171].

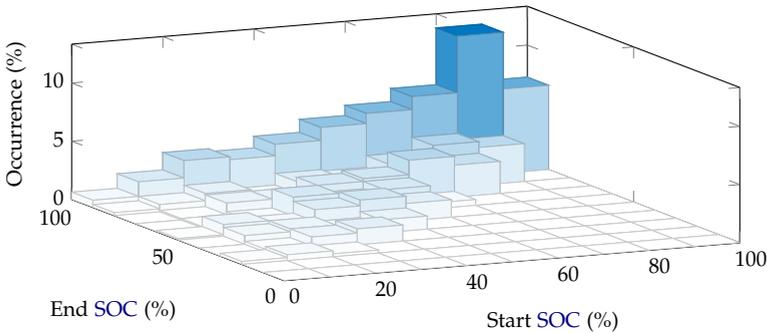
Charging habits can generally be characterized by i) the frequency an EV is charged, ii) the appliances used for charging, iii) the charging duration, and iv) the energy amount charged. Even though range anxiety does not constitute a problem from a technical perspective, it influences users' charging behavior. EV users charge more often than needed and in an uncoordinated way [154]. However, recent small-scale experiments, such as *E-Mobility Allee* [155] show that this is not generally the case. In this field experiment including ten EVs used in a residential area, the authors evaluated charging habits and respective impacts on grid load. They found that most of the time (73%) no EVs were charging at all, as shown in Figure 1.12. In



**Figure 1.12:** Characterization of charging habits according to frequency of simultaneously charging EVs in residential areas [155].

less than 2% of the time, two or more EVs were charging simultaneously. Negative effects of the charging behavior of multiple EVs, e.g., grid overload, are thus very unlikely. Recent customer surveys support these results. Only half of all EV users specified to charge their car on a daily basis [48,173]. However, a growing number of EVs might significantly affect power supply, which in turn calls for adequate power demand forecasting e.g., as in [18,174–183]

To the end of commonly used charging appliances, most EV users prefer either a conventional household socket (38%) or a wallbox (28%) at home.<sup>13</sup> The usage of public charging stations (12%) and charging at wallboxes at work (14%) each only accounts for less than one fifth of all charging events. Only fast charging stations (8%) are used more infrequently [173]. The amounts of energy charged are usually significantly below the EVs' total battery capacity, as the histogram in Figure 1.13 illustrates. Therein, all charging events of an EV fleet are classified according to their SOC at the start and the end in bins of 10%. The majority of charging events ends at an SOC of 100%, while most of them are starting at SOC of 60% or more. Note



**Figure 1.13:** Characterization of charging habits according to SOC at the beginning and the end of charging events [184].

that the aforementioned evaluations may change with advancing technology and increasing market penetration of EVs in the future; e.g., the duration of common trips taken or the preferred charging appliances. Appropriate modeling of EV users' driving and charging behavior, however, is still expected to be relevant [23].

<sup>13</sup>Wallboxes are charging devices usually connected to higher voltage supplies than the conventional household level. Common power levels comprise 7 kW to 22 kW.

### 1.1.3.2 Mobility Prediction

Appropriately integrating EV users' behavior into automated charging requires a prediction of the future actions of a user. In particular, the next locations and trips in between need to be predicted based upon historical movements and whereabouts. Then, suitable decisions—in this case regarding charging, e.g., where and how much to charge—can be made, according to the prediction. Alike works in the literature aim to predict users' next locations or travel trajectories. For this purpose, ensemble learning [160], Markov chains [159], or neural networks [163,164] are used. Other approaches semantically describe the problem of predicting travel trajectories [162] or define generic frameworks for this [161]. These works, however, pursue different purposes and are unrelated to charging strategies.

### 1.1.3.3 User Acceptance

A strategy combining various objectives, e.g., improve efficiency, prevent battery degradation, or provide grid services, requires sufficient user acceptance. Users must be granted a benefit to ensure an adoption of such strategies. Some researchers even doubt that without adequate user integration technological advancement alone will not increase the market share of EVs [154]. Possible incentives comprise cost advantages, convenience, environmental protection, or health benefits [36,170,185]. As driving and charging habits, work hours, and personality characteristics vary broadly, there is no universal approach to influence all users equally effective. Appropriate methodology needs to be adapted individually per user [170,186]. To this end, usage profiles and behavioral patterns can be deduced from case studies [154] and customer surveys [36]. For instance, most EV users specified environmental protection (34%) or cost savings (31%) as compelling reasons for using EVs [48,187]. Accordingly, suitable incentives must be selected and individualized for every user.

### 1.1.4 Design of Charging Assistants

Considering the combination of all, partly divergent objectives described in the previous sections, a multi-layered and multivariate decision problem originates. It requires information from diverse sources, e.g., historical user data, energy prices, physical characteristics of EVs and the power grid [41,188,189]. For users both the allocation of such information and solving the decision problem itself is non-trivial [190,191]. Therefore, the support of automation in terms of a software application is required. Table 1.7 presents related work proposed in the literature. The approaches are categorized according to three perspectives relevant to this thesis; considerations of EV fleets, e.g., as in [192–196], are not examined further here.

The first perspective deals with the efficient utilization of charging **infrastructure**. In particular, the objectives comprise to size or locate charging station in an efficient way—according to EV charging demand—and thus maximize profits [197–203]. Approaches assigned to the **EV user** perspective aim to simplify electric mobility from a customers’ point of view. Therefore, user-individual habits, or monetary and social incentives are integrated in EV charging assistants [204–216]. Holistic **power supply** concepts aim to increase the reliability and efficiency of the electric energy system. They, for instance, leverage market mechanisms to develop novel business cases [16,57,217–231].

For a safe and efficient operation, the architecture of charging assistants needs to be capable of scaling computation and memory resources to increased demand. The broad variety of possible architectures is reviewed e.g., in [41,232,233]. The use of cloud-based *microservices* allows to handle scalability accordingly [234]. For this, each microservice is deployed in its own, stateless and stand-alone virtual environment (*container*). Usually, a superordinate software application (*orchestration*) coordinates several dependent containers and adapts their available resources to momentary demand. Several concepts describing alike architectures can be found in the literature. Examples comprise use cases such as i) scalability of meta-heuristic optimization [235], ii) speed up of meta-heuristics via virtualization/parallelization [236], iii) semantic data handling with web services [237], iv) application-level cloud orchestration [238], and v) scalable machine learning/data analysis tasks [239].

**Table 1.7:** Literature review on EV charging assistants and strategies grouped by stakeholder perspectives.

Perspective	Reference	Comments
Infrastructure	CHEN et al. [197]	multiple-charger, multiple-port charging system
	CASSANDRAS et al. [198]	EV allocation at charging stations
	GUSRIALDI et al. [199]	charging scheduling/control at highway stations
	KIM et al. [200]	pricing, scheduling, energy management
	HOU et al. [201]	auction design for EV charging station scheduling
	GONZALEZ et al. [202]	charging station locating based on user habits
	LI et al. [203]	charging station planning based on user behavior
EV Users	ZHOU et al. [204]	incentive-based charging scheduling with uncertainty
	LIU et al. [205]	day-ahead charging scheduling, game model
	SUBRAMANIAN et al. [206]	optimal EV charging, electricity price peaks
	YÜCEL et al. [207]	privacy-aware supplier matching for EV charging
	VADIUM et al. [208]	charging strategies for CO <sub>2</sub> footprint reduction
	ALSABBAGH et al. [209]	distributed EV charging, social contribution concept
	ALINIA et al. [210]	EV charging scheduling, energy delivery commitment
	ALFACE et al. [211]	smart phone application for charging guidance
	HOCH et al. [212]	EV travel optimization, charging constraints
	STEFFEN et al. [213]	optimal EV charging, battery aging reimbursement
	EBRAHIMI et al. [214]	stochastic V2G, battery aging and user constraints
	ZHOU et al. [215]	charging scheduling, convenience, dynamic prices
MAIA et al. [216]	EV charging optimization to prolong battery life	
Power Supply	YOON et al. [217]	EV charging with demand response
	APPINO et al. [16]	dispatchability of EVs and RESs
	HOOG et al. [218]	market mechanism for grid constrained charging
	ALIASGHARI et al. [57]	charging scheduling with RESs, reserve markets
	ZHANG et al. [219]	load reduction for non-residential EV charging
	ALIPOUR et al. [220]	EV aggregators scheduling, energy/ancillary services
	HERTRAMPF et al. [221]	local power production, emission reduction
	GIORGIO et al. [222]	model predictive control for EV charging
	VALOGIANNI et al. [223]	sustainable EV charging via dynamic pricing
	SUNDSTRÖM et al. [224]	EV charge planning, power grid constraints
	LI et al. [225]	V2G approach with active battery anti-aging
	YANG et al. [226]	battery aging-aware V2G approach
	HESSE et al. [227]	optimization-based energy arbitrage concept
	DAS et al. [228]	multi-objective V2G approach
	WAN et al. [229]	EV charging scheduling with reinforcement learning
LUNZ et al. [230]	EV charging optimization with genetic algorithms	
VERMEER et al. [231]	vehicle-to-building, renewable energy, battery aging	

## 1.2 Open Questions

In reflection of the previously described state of the art, the following problems and research questions arise:

- Which model components serve to describe the context of a holistic charging assistant? Can model deficiencies be eliminated using advanced modeling approaches? What architectural considerations are needed for reliable communication, scalability, flexibility, and data security?
- Based on geodetic locations of EVs, how accurate can future trips and parking times be predicted? How can user requirements for a charging assistant be represented?
- To what extent do (electrical, thermal, degradation) characteristics of EV batteries limit the feasibility and profitability of V2G services?
- How does the uncertainty of user actions propagate inside a charging assistant context? Can a robust operation of the system be guaranteed?

## 1.3 Objectives

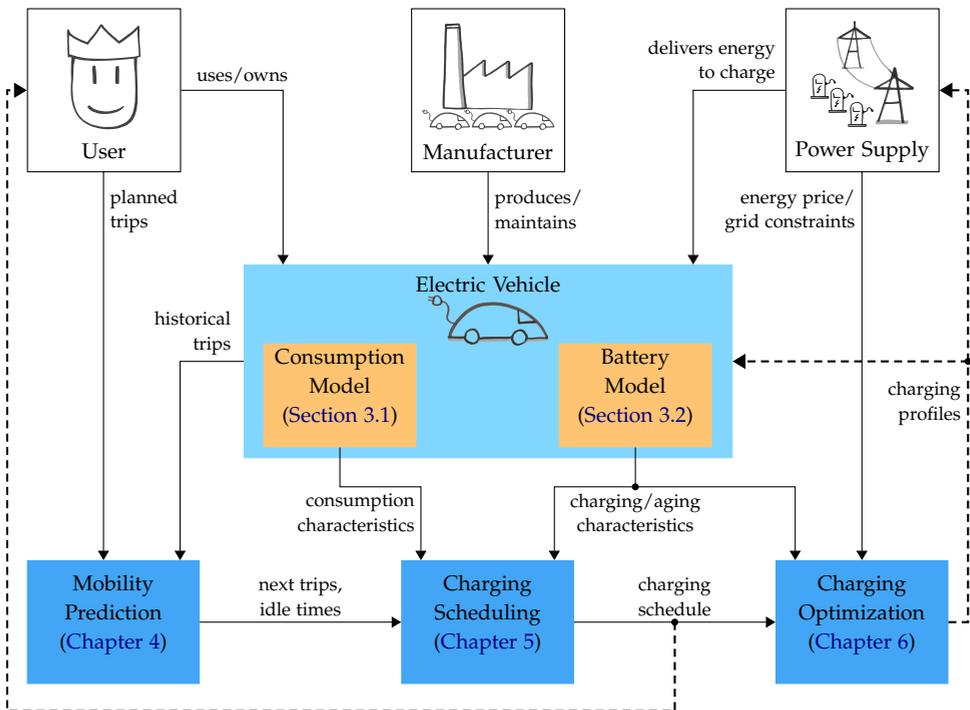
Based on the previous elaborated research questions, the objectives of this thesis comprise the following:

- **Chapter 2** outlines the requirements to tackle the open questions. All necessary components are identified and a modeling concept is presented. Then, a suitable way to connect components with each other is elaborated.
- **Chapter 3** presents models to describe the physical characteristics of EVs. Where necessary and applicable, components are modeled based on historical data to overcome model deficiencies. Then, the models are validated based on realistic data.
- **Chapter 4** evaluates and adapts concepts to predict users' future movements, typical parking locations and duration. To this end, historical data is used such as parking locations, energy consumption, or calendar appointments. Furthermore, the uncertainty of user decisions is estimated. For validation of the concept, real data is used.
- **Chapter 5** deals with scheduling charging stops based on predicted user actions. The focus here is on user requirements in terms of cost and convenience. To achieve robustness, the uncertainty of user actions is handled with multi-day scenarios. Simulating alternative strategies provides benchmark data and a proof of robustness.
- **Chapter 6** aims at adapting single charging events to user requirements, dynamic energy prices, and battery degradation. Additionally, the allocation of V2G services based on bidirectional power flow is integrated. Together with historical charging events, a real-world implementation of the concept is used for validation.



## 2 Concept and Architecture

To systematically approach the objectives outlined in [Section 1.3](#), [Figure 2.1](#) presents an overview of relevant stakeholders and subsystems. In the following, the combination of these subsystems is referred to as the *charging assistant*.



**Figure 2.1:** Layout of all stakeholders (white boxes), the EV (light blue box) represented by the vehicle-specific models (orange boxes), all developed modules (dark blue boxes), and their respective connection among each other.

Together with the content of this chapter, [Figure 2.1](#) supports the reader to understand the following three prerequisites. First, the context of an automated charging

assistant needs to be described. Therefore, all relevant stakeholders are introduced, which are represented by the white boxes in Figure 2.1. The individual perspective of each stakeholder is briefly outlined in Section 2.1. Second, a semantic segmentation of the charging assistant into technical subsystems, i.e., *modules*, is required. This modular concept enables flexibility for alternative solutions that are adaptable, e.g., to different technologies or regulatory decisions. In Figure 2.1, all modules of the charging assistant are represented by blue boxes; those are described in Section 2.2. The arrows and their respective labels represent relationships and information flows between single modules or stakeholders. Third, it is necessary to understand distinct characteristics about the implementation environment in which the system was developed and tested. Figure 2.1 shows a basic concept that is independent from manufacturer and technology. However, a systematization via the scope of each module allows to decide on suitable implementation environments. Hence, Section 2.3 provides further details on employed communication protocols, data storage, and deployment technology.

**Table 2.1:** Possible alternatives and their respective properties for design features of the charging assistant shown in Figure 2.1.

Design Feature	Alternatives	Properties
Reusability of Modules/Architecture	Proprietary Scope	simpler interface design, less development costs
	Generic Scope	high flexibility, broader application scope of components
Architecture Principle	Cloud-based Microservices	highly extensible and scalable system, availability of external information
	User End Device	advanced privacy and security, reduced communication overhead
User Representation	User-individual Models	precise adaption of charging strategy to user needs
	Generic User Models	less modeling effort, less privacy issues
	User-less Models	no data-intensive modeling
Modeling Approach	Data-driven	individual, adaptable and precise models
	Physical/Empirical	less data-intensive modeling, advanced extrapolation properties
Decision Problem Modeling	Heuristic/Rule-based	simple and inexpensive design
	Unified Optimization	precise solutions for decision problems, high computational complexity
	Two-stage Optimization	semantic separation of hierarchical decisions, manageable complexity

---

To furthermore justify the design of the charging assistant, [Table 2.1](#) presents a qualitative overview of alternatives for each design feature; the eventually chosen alternatives are briefly described in the following. To attain a reusability of the developed modules, a generic scope would hold the most benefits in terms of flexibility and extensibility. Due to limited time and resources, however, a hybrid approach of proprietary and generic modules is chosen. Scalability and flexibility also determines the cloud-based microservice architecture that is chosen. To counteract privacy and security issues, modules processing highly user-related data (e.g., **mobility prediction**) are run on user end devices. To ideally represent the **EV** user needs, user-individual models, e.g., for mobility habits, are inevitable. If user data is not available or users deny their consent to use it, generic user models are used as base alternative. In a similar way, data-driven modeling approaches are preferred over physical ones whenever data is available in sufficient amount and quality. For cases with sparse training data, physical or empirical models are used exclusively or reasonably combined with data-driven models. Decision problems—especially to schedule and optimize charging events—may be easiest modeled with rule-based heuristics; their results, however, lack precision and robustness compared with optimization models. A unified optimization approach would yield the most precise solutions but may be difficult to solve due to high computational complexity. Consequently, two hierarchical optimization models (i.e., **charging scheduling** and **charging optimization**) are chosen that combine sufficiently precise results with manageable complexity.

## 2.1 Stakeholders

To develop a holistic charging assistant, first all relevant stakeholders are identified; those are represented by the white boxes in [Figure 2.1](#). Then, each stakeholder needs to be characterized by describing their individual perspective and the relation to other stakeholders and modules. The following three stakeholders are defined:

- **User:** The person who owns and uses the [EV](#).<sup>1</sup> The interest of the user generally comprises individual and convenient mobility, cost saving, and sustainability (see [Table 1.5](#)). Especially when switching from [ICEVs](#) to [EVs](#), users may desire guidance on battery usage and charging procedure. Furthermore, typical mobility habits, as described in [Section 1.1.3](#), are determined by users themselves. The according information—assuming a previous consent of the user—can then be used to analyze and predict the user’s future mobility. To support mobility prediction, users may also provide information about planned mobility events. For instance explicitly, via set routing destinations, or implicitly using calendar appointments.
- **Manufacturer:** The car manufacturer produces and maintains the [EV](#). It further delivers experimental data on energy consumption and battery degradation used for modeling purposes, see also [Section 1.1.2](#) and [Chapter 3](#). The manufacturer’s interest comprises designing reliable and value-stable products to avoid future warranty expenses or damages to reputation. In addition, the manufacturer may have sovereignty over data on [EV](#) usage. This designates the manufacturer to be the prospective operator of the developed charging assistant.
- **Power Supply:** Within the power supply stakeholder all entities that support the allocation of electric energy are concluded, those comprise:
  - i) Producers of electric energy, either based on centralized energy sources, e.g., in coal-fired plants, or based on distributed energy sources, e.g., solar and wind energy generation.
  - ii) Operators of transmission and distribution grids that maintain and—if required—enhance physical grid infrastructure to ensure a reliable constitution of the electric utility grid.

---

<sup>1</sup>Car leasing or sharing, which are concepts that usually involve different users than the [EV](#) owner, are neglected. The developed charging assistant of this thesis is also applicable to these concepts of ownership; monetary value streams, however, may differ.

- iii) Providers of charging infrastructure that maintain and market infrastructure particularly relevant to charge EVs, e.g., public charging stations, or web-based search engines for charging stations.

All these entities support charging of EVs, see therefore also [Section 1.1.1](#). Furthermore, the power supply pursues the interest in a reliable and efficient operation of the electricity grid. Hence, it dynamically manages grid load and demand (or according forecasts) to match closely. This can be accomplished via dynamic electricity prices, which comprise costs for energy generation, and for the operation of grid and charging infrastructure. These dynamic tariffs are usually published as price tables. Additionally, grid operators may issue *hard* grid constraints, e.g., power limitations on an irregular basis and for a limited time.

These stakeholders are not elaborated in further detail. Rather, they are set in according context in the following described modules.

## 2.2 Modules of an Automated Charging Assistant

To describe the functionality of an automated charging assistant, separate modules are defined, as indicated by the blue boxes in [Figure 2.1](#). These modules represent subsystems that have a specific purpose and can be modeled and implemented individually. The **EV** is a central component illustrated by a light blue box in [Figure 2.1](#), which again is subdivided into vehicle-specific models (orange boxes). All stand-alone modules of the charging assistant, which are not vehicle-specific, are represented by dark blue boxes. In the following Sections [2.2.1](#) to [2.2.4](#), each module is characterized according to:

- Input and output values with their respective structure and interfaces
- Scope, i.e., whether the module is user-, vehicle-, or battery-specific
- Time frame, i.e., the time each module is created, trained or retrained, and used for predictions or calculations

Furthermore, links to the chapters that describe the detailed modeling and validation of each module are provided.

### 2.2.1 The Electric Vehicle as a Means of Transportation

As the **EV** is assumed to be the user's prioritized means of transportation, it is a central part of the charging assistant (light blue box in [Figure 2.1](#)). It provides historical information about:

- Movements, i.e., trips described by a start and end time, start and end location, and provided as single data points in JavaScript Object Notation (**JSON**) format. This information is vehicle-independent as it actually represents movements of the **EV** user. For this study, however, the **EV** functions as a sort of tracking device.
- Battery behavior, i.e., time series data on charging power, battery **SOC**, temperature, current, and voltage. This information is vehicle-independent, but battery-specific, i.e., it depends on the installed battery cells and power electronics.
- Energy consumption, i.e., time series data on vehicle speed and acceleration, propulsion and brake torque, and auxiliary consumption, e.g., for heating.

This information is both battery- and vehicle-specific as physical characteristics of the entire EV account for energy consumption, e.g., the air resistance or efficiency of the electric engine.

The battery- and vehicle-specific characteristics that can be derived from the provided data is most important to the algorithms of the proposed charging assistant. Based on the fundamentals in Section 1.1.2, models therefore are created that suit the application purposes in terms of accuracy and complexity. These models can be initialized with a generic baseline approach, e.g., an analytical model. Once a sufficient amount of measurement data is collected, data-driven approaches with higher accuracy can also be implemented.<sup>2</sup> Thus, either vehicle-specific models—e.g., to represent manufacturing tolerances—, or user-specific models—i.e., to incorporate the EV user’s individual behavior—can be created. The model characteristics can then explicitly be included into an optimization scheme (see Chapter 5 and Chapter 6). Alternatively, the consumption model can be deployed as a stand-alone microservice accessible via a standardized Application Programming Interface (API).

The **consumption model** represents characteristics of the EV when driving. Especially time series data on vehicle speed and acceleration, battery temperature and torque is used as input values. This time series data is assumed to be available from a *learning map*, e.g., as described in [241]. Thus, the consumption model can be used to predict the EV’s user-individual energy consumption of future trips. For the modeling process both historical and experimental consumption data can be used. In this way, not only vehicle-specific power electronics and mechanical drive-train components can be represented, but also user-individual behaviors. The details of this subsystem are described in Section 3.1.

The **battery model** describes the electrical, thermal, and aging behavior of the EV battery. For the context of this thesis, the focus lies on battery models for charging the battery, i.e., a low-dynamic operation.<sup>3</sup> For this, mostly experimental charging data is used, as the charging process generally does not include user-individual influences. All details on modeling and the model’s performance are outlined in Section 3.2.

---

<sup>2</sup>In the case of sparse training data, an ensemble approach as in [240], using both analytical and data-driven models, might be conceivable.

<sup>3</sup>To represent EVs’ driving characteristics, similar battery models can be used. However, a higher time resolution and according model types to capture further non-linearity would be required.

## 2.2.2 Predicting the Mobility of Electric Vehicle Users

Scheduling and optimizing EV charging in an automated manner requires to forecast the EV user's mobility.

Using an ensemble of different machine learning techniques the module then outputs user-specific estimations on:

- Common whereabouts, i.e., the frequently visited locations of the EV user, characterized by a geodetic center location surrounded by a radius of stay, and a differentiation into known and unknown locations (*measurement noise*). The information may also be augmented with data on charging infrastructure and user preferences.
- Typical parking times at frequently visited locations, i.e., the time the EV will be parked at a location given a specific context, e.g., day of week, time of day, or previously visited locations. The estimation may be advanced by also estimating a measure of uncertainty, e.g., the parking times' standard deviation.
- Possible next trips, i.e., the locations the EV user might visit next, with their respective probability of being visited next, again given a specific context, e.g., day of week, time of day, or previously visited locations. Furthermore, a residual probability is deduced to quantify the chance that the user will visit an unknown location next.

Furthermore, the user is able to input (deterministic) information on planned trips, e.g., via set routing destinations or via a calendar interface. Together with one-time user inputs such as *home* and *work* location, the forecast's accuracy can thus be enhanced.

The **mobility prediction** module could either be implemented as a cloud-based microservice, or directly on the user end device; besides the user's smart phone, this may also be the infotainment system of the EV. The latter approach, i.e., the EV both collects and analyzes data, only requires device-internal interfaces for raw mobility data. Only the anonymously rendered result will be made available on demand for the use in subsequent modules outside the EV infotainment system. This poses a clear benefit in terms of data security and privacy issues, see also [Section 1.1.4](#).

When initializing the module, user data first needs to be accumulated before accurate predictions can be made. Thereafter, the learning models inside the module

are retrained in frequent intervals (e.g., 30 days) considering a rolling time window (e.g., 90 days) of historical data. To continuously process the incoming stream of mobility data, accordingly adapted methods, e.g., *stream clustering* as described in [242], are required. All details on the used methods and the belonging results of the **mobility prediction** module are described in [Chapter 4](#).

### 2.2.3 Scheduling of Charging Events

Keeping track of all possible upcoming mobility needs is not easy for human individuals. Based on that, maintaining a sufficiently charged **EV** battery to stay mobile in the majority of cases is even more difficult for **EV** users.<sup>4</sup> To support the **EV** user to stay mobile, the **charging scheduling** module (see [Figure 2.1](#)) makes use of the **mobility prediction** outputs. In particular, the set of next locations, the probability to go there, and corresponding parking duration at known places are used to automatically schedule charging events. The resulting *charging schedule* considers several trips and parking periods in a given future horizon, e.g., five days, and contains information about:

- The need to plug in the **EV** at a charging station when approaching a targeted location at a specific time. This information is directly fed back to the **EV** user, e.g., via the **EV** infotainment system, as a user action is required in this case, i.e., plug in the **EV**.
- Once the **EV** was plugged in, how much energy to charge at a specific location. Together with the expected parking duration at this location, the required energy amount will subsequently be processed in the **charging optimization** module.

These two decisions are modeled in a mixed-integer optimization problem, whose solution yields the required decision policy. In this context, (linearized) characteristics of the **EV** (on consumption, charging, and battery aging, see [Section 2.2.1](#)) are also necessary. To represent the uncertainty inherent in the user's actions, a set of possible scenarios is created. The user may define a security level, i.e., the portion of cases (e.g., 5%) in which the battery is allowed to undercut a lower **SOC** limit. Then, both security level and the scenarios' probability of occurrence are considered in the optimization as chance constraints.

---

<sup>4</sup>The term "staying mobile" refers to **EV** trips that do not require additional charging stops along the way. Rather, the **EV** is only charged when it is parked anyway due to another, unrelated purpose, e.g., an appointment.

**Charging scheduling** is a generic (neither vehicle-, nor user-specific) module, which is only modeled once and not retrained during operation. Thus, a deployment as microservice in a cloud-based environment is conceivable.<sup>5</sup> In this way, the module can also virtually be replicated for several EV users. A recalculation of the charging schedule may be requested via a standard API. This could happen either manually, by the user, or automatically each time the user attempts to start a trip, i.e., gets inside the EV and starts driving. The same standard API is also used to feed back the calculated charging schedule to the EV user (see dashed line in Figure 2.1) and subsequent modules. Chapter 5 outlines the approach and according evaluations of the **charging scheduling** module in detail.

## 2.2.4 Optimization of Charging Events

The majority of charging events follows a standard procedure, i.e., plug in the EV and charge at maximum available power to 100 % SOC. However, *default charging* is sub-optimal in terms of efficiency (e.g., power conversion losses), sustainability (e.g., battery aging), and cost (e.g., electricity prices). In general, charging events offer flexibility that can be utilized:

- Time flexibility, i.e., additional time to charge is available, than actually required. In this context, the **mobility prediction** module estimates the parking duration for a given parking location and given arrival time (see Section 2.2.2).
- Energy flexibility, i.e., a lower target SOC than 100% may suffice in many cases. To this end, a superordinate scheme, i.e., the **charging scheduling** module (see Section 2.2.3) calculates the minimum required target SOC to stay mobile in subsequent trips.

The **charging optimization** module exploits the identified energy and time flexibility in a single charging process. Particularly, a time-discrete charging power profile is calculated to reduce the EV's operating cost due to electricity prices and battery aging. In addition to (power-dependent) price tables, the resulting charging power profile may also be subject to strict power limitations issued by the power supply. While the target SOC is assumed to be deterministic, both the estimated parking time and the irregular occurring power limitations introduce uncertainty. To tackle the introduced uncertainty, a stochastic optimization scheme may be used. Therein,

---

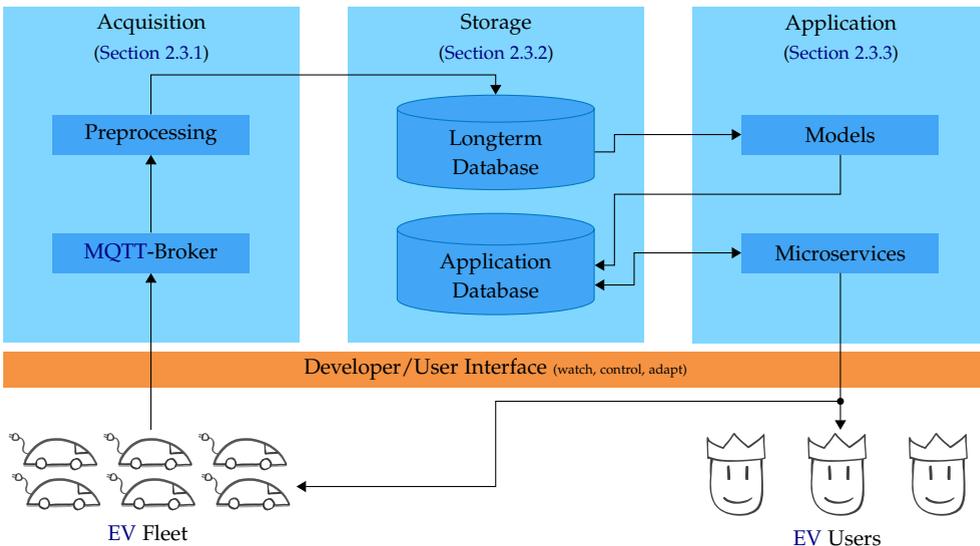
<sup>5</sup>User-specific quantities, e.g., the individual security level, or vehicle-specific quantities, e.g., consumption characteristics, are replaceable parameters to the optimization problem.

the detailed model of the EV battery is also utilized (see Section 3.2). To obtain a deterministic charging power profile despite the uncertainty, chance constraints with a user-defined security level (e.g., 1 %) can be introduced. This is, the portion of cases the user accepts the actual SOC to be below the target SOC when departing from a parking location.

Similar to **charging scheduling**, **charging optimization** is a generic (neither vehicle- nor user-specific) module, which is only modeled once and not retrained during operation. Hence, a cloud-based microservice deployment may also be used to ensure availability and scalability. Via a standardized API, a recalculation may be triggered automatically each time, the EV is parked and plugged into a charging station. Using the same API, the resulting charging profile is sent to the EV-BMS that starts and controls the charging process (see dashed line in Figure 2.1). Simultaneously, the user may monitor and adapt this process via a smart phone app in real time. Furthermore, the power supply receives the calculated charging profiles to incorporate them in load forecasts, electricity pricing, or similar. Chapter 6 presents all details and evaluations on the **charging optimization** module.

## 2.3 Environment of Implementation

Developing, evaluating, and deploying single modules of the charging assistant requires a suitable environment. Especially when processing user-related data, appropriate data security measures must be taken as required by law. Furthermore, a centralized, server-based infrastructure is inevitable to i) coordinate and monitor data efficiently, ii) make data persistently available, and iii) ensure sufficient data quality. Therefore, a cloud-based environment is used, also referred to as *backend*, as it is shown in Figure 2.2 [243,244]. For further reading on the fundamentals of cloud infrastructures and computing, see also [245].



**Figure 2.2:** Scheme of cloud-based environment for data acquisition via a MQTT-broker, storage in several databases, and application of data-driven models or microservices [243,244].

The cloud-based environment consists of three semantic subsystems for **acquisition**, **storage**, and **application** of data; those are described in the following Sections 2.3.1, 2.3.2, and 2.3.3. Via an **interface** (a set of various APIs) all data processes can be monitored and analyzed. The Interface also provides feedback to the **EV fleet** and **EV users**.

Note that Figure 2.2 shows a schematic setup for the use case of this thesis. This is, the environment mainly collects, stores, and processes sensor data of **EVs**. Furthermore, it connects **EVs** and **EV Users** (or developers during development) with each other. The underlying concept of the framework may, however, also be a template

for other development setups; for instance, a setup with different sources of data and stakeholders. Furthermore, the concept may be a blueprint for later productive systems that also comprise larger fleets of EVs.

### 2.3.1 Acquisition of Data

To develop and maintain high-quality applications based on data-driven models, a sufficient supply of both live and historical data is required. For the scope of this thesis, the data **acquisition** especially comprises on-board sensor data of EVs (see [Figure 2.2](#), left). To collect this data, a fleet of EVs is equipped with Single-Board Computers (SBCs) that run on a *Linux*-based operating system. Each SBC possesses a mobile internet connection and Global Positioning System (GPS) antenna. As the SBC enables bidirectional communication (via the Message Queuing Telemetry Transport (MQTT) protocol, a *publish and subscribe* mechanism), it functions as a gateway between EV and backend. While the SBC is installed inside the EV, its functionality can be updated remotely using packetized software.

During data collection, i.e., during driving and charging, the SBC reads physical raw data from the EV's Controller Area Network (CAN). As the amount of live data generated aboard each EV is very circumstantial and contains irrelevant components, the SBC initially filters the data stream. The selected time series signals (370 signals, 10 Hz recording frequency) are then decoded, and compressed. Before transmitting the data to a cloud-hosted MQTT broker, all signals are packetized into MQTT messages of approx. ten to thirty seconds of measurement data. The resulting data stream amounts to approx. 10 MB/h, while the SBC is also able to buffer messages in case the backend connection is lost.

Once received, the MQTT messages are decompressed in a **preprocessing** module. To maintain a high data quality, the Preprocessing module proceeds further actions:

- Filtering of corrupted data to avoid the storage of unstructured data to become *oversaturated*.
- Extracting meta-information to create a structured index that simplifies and speeds up querying of raw data.
- Standardization, i.e., assigning a unique reference name, data type, and physical unit to each signal in order to align data of different sources, e.g., data from EVs of various make and model.

After Preprocessing, the filtered and unified data is passed to the next module to be stored accordingly.

### 2.3.2 Storage of Data

Building applications based on data-driven models generally requires large amounts of data. Therefore, appropriate data **storage** is inevitable (see [Figure 2.2](#), middle). For the scope of this thesis, **EV** data is initially stored in a **longterm database**, after it was collected and pre-processed in the Acquisition module (see [Section 2.3.1](#)). According to its structure, different types of storage technologies are employed. A document-oriented *NoSQL*-database stores time series data from recorded **EV** signals. Due to vehicle-dependent features and signal availability, this data is *semi-structured*. A unified document schema simplifies the processing of large amounts of data. In particular, one stored document (in **JSON** format) represents one received **MQTT** message. Meta-information of the **EV** fleet is stored in a relational database. Both structure and content of this meta-information is almost static, i.e., it changes rarely. Hence, this data is considered as *structured*.

A separate **application database** stores data particularly relevant for applications. This comprises e.g., i) specifically processed data for model training, ii) data-driven models and their training status, iii) user-specific information, or iv) configurations of microservice applications. Most of this data is again unstructured and thus stored in a *NoSQL*-database.

### 2.3.3 Application of Data

To gain value from the data collected, it is processed into data-driven **models** designed for a specific purpose, such as predicting **EVs'** energy consumption. Therefore, data-driven models are created in the **application** module (see [Figure 2.2](#), right) according to the methods described in [\[246\]](#). To this end, the full potential of the cloud-based environment can be leveraged, see also [\[239\]](#). A continuous availability and potential automation allows to efficiently develop and deploy models with an updated data stream.<sup>6</sup> Furthermore, the modular structure allows to create and train several models in parallel. For this purpose, all persistent information is stored in the Application Database, see [Section 2.3.2](#). To monitor and adapt

---

<sup>6</sup>Risks inherent in a continuous availability, e.g., information security issues, need to be handled accordingly [\[247\]](#).

the training progress of these models, the **interface** provides detailed insights for developers.

Once data-driven models are mature, i.e., their accuracy is sufficient for a dedicated use case, they are integrated into standalone **microservices**. To obtain a flexible and scalable architecture—especially when deploying multiple microservices—the following aspects are inevitable:

- Each microservice is run in its individual, lightweight virtual environment (*container*) using *Docker* [248]. This allows to automatically start, restart, and replicate microservices according to increased demand.
- All microservices are hosted on a cluster of virtual machines. The related containers are orchestrated by a parent agent called *Kubernetes* [249]. The orchestration task comprises scaling and replicating microservices to demand, as well as supervising the required dependencies with other microservices.
- Each microservice possesses standardized **APIs**. This enables to communicate with other microservices and process outside information from other services, e.g., weather data. In addition, microservices can communicate with the user's end devices, e.g., smartphones or **EVs**, using the same **APIs**. A Developer Interface can be used to monitor all microservices (and their replications) and handle errors accordingly.

With this setup, microservices based on data-driven models can be developed and evaluated in a production-like environment. Furthermore, a wide range of applications—such as the proposed charging assistant—can be realized.

## 2.4 Conclusion

Chapter 2 outlines a systematic approach to separate semantic entities of the charging assistant proposed in this study. The contributions can be briefly summarized in the following three parts of the approach.

First, individual stakeholders and their perspectives are described to identify requirements of the charging assistant. In this way, each stakeholder is set in according context within the charging assistant. Second, separate modules are defined representing subsystems that have a specific purpose and can be modeled and implemented individually. Each module is systematically characterized according to i) its input and output values, ii) its scope, i.e., if the module is user-, vehicle-, or battery-specific, and iii) its time frame, i.e., the time each module is created, trained or retrained, and used for predictions or calculations. Third, an exemplary implementation framework for data collection, storage, and application is presented. In this way, privacy and security issues are discussed, as the charging assistant processes sensitive, user-related data. This framework may function as a blueprint for other development setups or later productive systems with different sources of data and stakeholders.

# 3 Modeling of Electric Vehicles

For the scope of this thesis, the EV is assumed to be the user's prioritized means of transportation. Hence, a sufficient understanding of its characteristics while driving and charging is essential. To achieve this, according models to estimate energy consumption and battery state are designed based on the fundamentals in Section 1.1.2. These models can be initialized with a generic baseline approach, e.g., an analytical model. If necessary, data-driven approaches with higher accuracy may also be implemented. For the latter, however, historical data on the EV user's movements, the battery behavior, and the EV's energy consumption is required in sufficient amount and quality.

Section 3.1 outlines the design of a consumption model to estimate the EV's energy consumption for future trips. The consumption model possesses a hybrid structure as it consists of a vehicle-specific, analytical baseline model and a user-specific data-driven model. For modeling of the latter model, especially time series data from on-board measurements is selected and processed, see Section 3.1.1; the modeling approach is partly based on the ideas presented by SCHWENK et al. in [243]. Then, different machine learning approaches are chosen (Section 3.1.2) and subsequently validated on an independent test set (Section 3.1.3).

Section 3.2 describes the structure of a three-part battery model to simulate EV charging events; parts of this section are based on the battery model described by SCHWENK et al. in [250]. First, an electrical model to estimate the energy evolution of the battery is outlined in Section 3.2.1. Second, Section 3.2.2 presents a thermal model to estimate the battery temperature progression. Third, a semi-empirical battery aging model to estimate the decay of the battery's storage capacity is described in Section 3.2.3. All three components of the battery model can be operated with typical EV on-board sensor data. Accordingly, the battery model validation is also proceeded with such data in Section 3.2.4.

### 3.1 Consumption Model

For a thoroughly adapted and reliable charging strategy, i.e., both scheduling and optimization of charging, an accurate estimation of the future energy demand is inevitable. The consumption model presented in this chapter therefore estimates the energy consumption of all future trips a user may take. Figure 3.1 shows an overview of the consumption model that is built by use of the framework presented in Section 2.3. All individual components are described in the following.

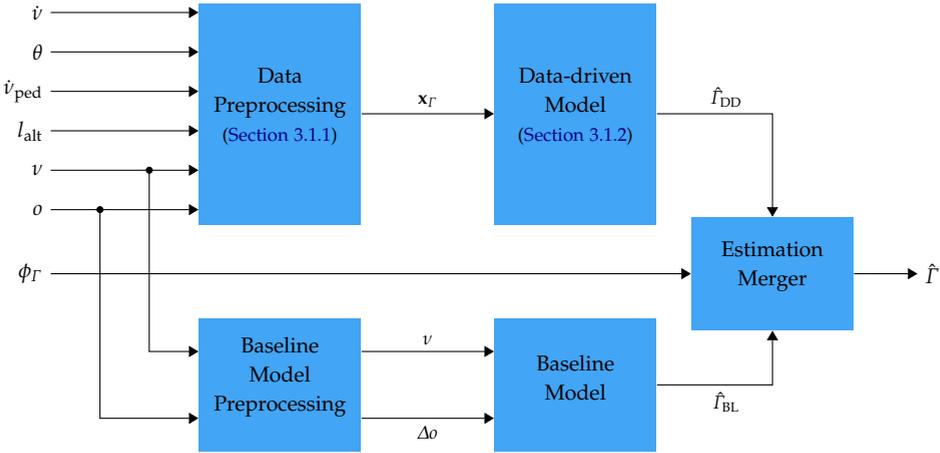
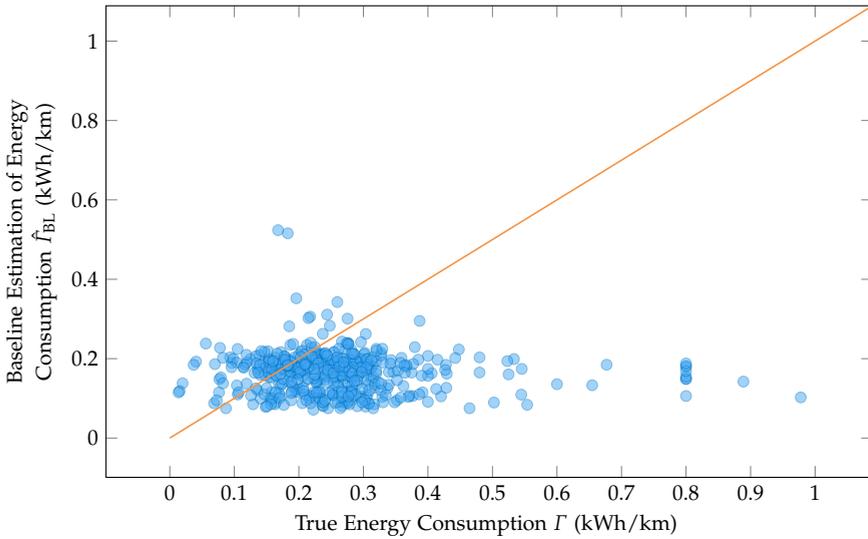


Figure 3.1: Schematic block diagram of the consumption model and its components.

A **baseline model** estimates the EV's specific energy consumption  $\hat{\Gamma}_{BL}$  based on the trip distance  $\Delta o$  and speed  $v$  along a trip. This model is built upon experimental data and the expertise of the EV manufacturer and represents the vehicle's physical characteristics in terms of energy efficiency.<sup>1</sup> Independent of user-related data, thus a robust estimation for  $\Gamma$  can be achieved. However, both user's driving style and environmental influences diminish the prediction quality of the Baseline Model. To illustrate this, Figure 3.2 compares the true energy consumption  $\Gamma$  of recorded trips with the estimations  $\hat{\Gamma}_{BL}$  of the Baseline Model. It can be seen that the majority of estimations  $\hat{\Gamma}_{BL}$  lays within a range of approx. 0.15 to 0.25 kWh/km. The

<sup>1</sup>The details of the Baseline Model are subject to the confidentiality of the manufacturer and are therefore not described further here. Similar approaches may be found, e.g., in [251,252]; for the fundamentals of EV consumption models, see [253].



**Figure 3.2:** Comparison of energy consumption estimation  $\hat{\Gamma}_{BL}$  of the Baseline Model (see Figure 3.1) with the true energy consumption  $\Gamma$  of recorded EV trips; the orange line indicates ideal model behavior.

true consumption  $\Gamma$ , however, reaches up to approx. 0.8 kWh/km for some cases. This is likely to be caused by user-individual driving maneuvers such as excessive acceleration and breaking, which reduce the EV's overall energy efficiency.

If historical trip data of the EV user is available, the **data preprocessing** module extracts additional features characterizing the user's driving style and environmental conditions, see Section 3.1.1.<sup>2</sup> Then, an additional **data-driven model** can be trained and used to estimate the user-specific energy consumption  $\hat{\Gamma}_{DD}$ , see Section 3.1.2. In this way, a higher estimation accuracy may be achieved compared with the Baseline Model, see Section 3.1.3.

An **estimation merger** module subsequently weights and adds up both estimations. Thus, the final energy consumption estimation is given by

$$\hat{\Gamma} = \phi_{\Gamma} \cdot \hat{\Gamma}_{BL} + (1 - \phi_{\Gamma}) \cdot \hat{\Gamma}_{DD}. \quad (3.1)$$

<sup>2</sup>The user's consent to the recording and use of this data for modeling purposes constitutes a prerequisite.

The coefficient  $\phi_\Gamma \in [0, 1]$  represents the weight assigned to the Baseline Model. If the amount or quality of user-related data is low, i.e., the Data-driven Model yields poor results, the weight for the baseline consumption  $\hat{I}_{BL}$  is  $\phi_\Gamma \cong 1$ . Accordingly, the data-driven consumption  $\hat{I}_{DD}$  is weighted with  $(1 - \phi_\Gamma) \cong 0$ . This also enables conformity in terms of data security and privacy. For instance, if user-related data is available in sufficient quality, but the user has revoked their consent to the use of this data. If, on the other hand, user-related data is available and can be used, the user-specific estimation  $\hat{I}_{DD}$  of the Data-driven Model may outperform the Baseline Model. The baseline estimation  $\hat{I}_{BL}$  may then be enhanced by increasing the weight for  $\hat{I}_{DD}$ , i.e.,  $\phi_\Gamma$  is decreased.

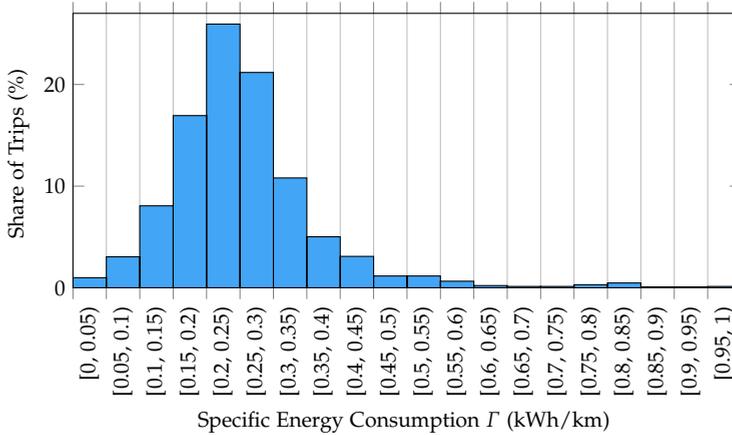
A suitable value for  $\phi_\Gamma$  could be adapted continuously depending on the available data and its quality. To this end,  $\phi_\Gamma$  could be dynamically chosen based on the performance of both models in different scenarios; e.g., for trips that are well or not well represented in the training data. In this thesis, however, the determination of  $\phi_\Gamma$  will not be elaborated in further detail; however, an exemplary approach for this may be found in [240].

### 3.1.1 Selection and Preprocessing of Training Data

To capture both the EV user's driving style and environmental conditions, the Data-driven Model uses additional information. For this purpose, a fleet of ten EVs provides 370 measured signals, which are recorded with a sampling rate of 10 Hz, see also Section 2.3.1.<sup>3</sup> The Data Preprocessing module then extracts relevant features, see Figure 3.1. For future trips, this data is assumed to be available from a *learning map*, e.g., as described in [241]. Therein, characteristic profiles for each signal are inferred based on historical trips a user has taken. For model training, historically recorded trips  $\mathcal{J}$  are used. Each trip  $j \in \mathcal{J}$  starting at  $t_{j,0}$  and ending at  $t_{j,\tau}$  represents one training sample, whereas  $\tau_j = t_{j,\tau} - t_{j,0}$ . As the consumption model targets to estimate the EV's energy consumption, the training label

$$\Gamma_j = \frac{e_{j,\tau} - e_{j,0}}{o_{j,\tau} - o_{j,0}}, \quad \forall j \in \mathcal{J}, \quad (3.2)$$

represents the specific energy consumption of a trip  $j$ . Therein,  $e_{j,0}$  is the battery energy at the beginning of the trip and  $e_{j,\tau}$  the battery energy at the end. Similarly,  $o_{j,0}$  is the mileage (in kilometers) at the beginning of the trip and  $o_{j,\tau}$  at the end, respectively. Figure 3.3 shows a histogram of  $\Gamma$  for the available training data  $\mathcal{J}$ . In the



**Figure 3.3:** Histogram of training data classified by the specific energy consumption  $\Gamma$  in bins of 0.05 kWh/km.

<sup>3</sup>This equals a data stream of approx. 12 MB/h per vehicle.

available data set  $\mathcal{J}$ , the average energy consumption amounts to 0.293 kWh/km.<sup>4</sup> In addition, a few outliers can manually be detected on the upper end of the distribution. They occur due to very short trips taken in cold ambient temperatures. In this case, a high share of consumed energy is used for heating both the EV battery and interior. For model training, these outliers are removed.

To the end of relevant features, all available signals are manually screened based on fundamental knowledge of EV modeling, see Section 1.1.2. The signals presented in the leftmost column of Table 3.1 appear to be most reasonable.

**Table 3.1:** Selection of consumption model features according to univariate feature relevance based on the SPEARMAN-correlation coefficient (see Section A.1.1) in relation to the specific energy consumption  $\Gamma$  and manual screening.

Signal		Aggregation	SPEARMAN-Correlation $q_{\text{Sp}}(\bullet, \Gamma)$	Selected
Mileage	$o$	difference	0.081	yes
Vehicle	$v$	mean	0.243	yes
Speed		minimum	0.082	no
		maximum	0.280	yes
Longitudinal Acceleration	$\dot{v}$	mean	-0.157	yes
		minimum	-0.071	yes
		maximum	-0.007	no
Lateral Acceleration		mean	-0.009	no
		minimum	-0.081	no
		maximum	0.038	no
Acceleration	$\dot{v}_{\text{ped}}$	mean	0.155	yes
Pedal Position		minimum	0.004	no
		maximum	0.089	yes
Battery Temperature	$\theta$	mean	0.161	yes
		minimum	0.159	no
		maximum	0.161	no
Geodetic Altitude	$l_{\text{alt}}$	mean	-0.056	no
		minimum	-0.051	yes
		maximum	-0.018	yes
		difference	-0.042	yes

To condense the information provided by these signals, different aggregation functions are applied to the originally 10-Hz-sampled signals. These aggregation func-

---

<sup>4</sup>The energy consumption is established from a fleet of development vehicles and does not necessarily represent the behavior of the same EVs in production condition.

tions comprise i) *mean* value, ii) *maximum* value, iii) *minimum* value, and iv) the *difference* throughout a trip  $j$ , where applicable, see second-left column in [Table 3.1](#).

To get an overview on univariate feature relevance, the SPEARMAN-correlation coefficient (see [Section A.1.1](#)) in relation to  $\Gamma$  is calculated for all samples, see second-right column in [Table 3.1](#). Furthermore, the rightmost column outlines, whether a combination of signal and aggregation function has been selected as feature for the subsequent training process of the models. Using this systematic approach, the feature vector

$$\mathbf{x}_{\Gamma,j} = \begin{pmatrix} \Delta o \\ \text{mean}\{v\} \\ \text{max}\{v\} \\ \text{mean}\{\dot{v}\} \\ \text{min}\{\dot{v}\} \\ \text{mean}\{\dot{v}_{\text{ped}}\} \\ \text{max}\{\dot{v}_{\text{ped}}\} \\ \text{mean}\{\theta\} \\ \text{min}\{l_{\text{alt}}\} \\ \text{max}\{l_{\text{alt}}\} \\ \Delta l_{\text{alt}} \end{pmatrix}_j \in \mathbb{R}^{11}, \quad \forall j \in \mathcal{J}, \quad (3.3)$$

is obtained, whose individual components are describe in the following:

- The trip distance  $\Delta o$  is a basic feature that is also used by the Baseline Model.
- The average vehicle speed  $\text{mean}\{v\}$  and the maximum vehicle speed  $\text{max}\{v\}$  characterize the type of trip and the user's individual driving style. High values of both features are assumed to increase the specific energy consumption, as the positive SPEARMAN-correlation coefficient in [Table 3.1](#) indicates.
- Similar to vehicle speed, the average and minimum longitudinal acceleration  $\text{mean}\{\dot{v}\}$  and  $\text{min}\{\dot{v}\}$  also characterize driving style, particularly harsh braking maneuvers. To achieve sufficient deceleration in this situation, the EV's braking system must perform hydraulic braking instead of energy-saving recuperation.
- Aggressive driving style can also be quantified using the driver's input on the acceleration pedal. In particular, the desired mean and maximum propulsion torque represented by  $\text{mean}\{\dot{v}_{\text{ped}}\}$  and  $\text{max}\{\dot{v}_{\text{ped}}\}$ . The higher these values are, the less efficient the electrical power train components can be operated, see the positive SPEARMAN-correlation coefficient in [Table 3.1](#).

- The average battery temperature  $\text{mean}\{\theta\}$  throughout a trip needs to be estimated separately using a thermal battery model, see also [Section 3.2.2](#). It considers e.g., environmental conditions such as the ambient temperature. Furthermore, aggressive and thus inefficient driving increases the battery temperature and lowers the efficiency of the energy storage itself.
- The features based on the geodetic altitude  $l_{\text{alt}}$  introduce additional information about the driving environment. While  $\Delta l_{\text{alt}}$  represents the sole potential energy difference of a trip,  $\min\{l_{\text{alt}}\}$ , and  $\max\{l_{\text{alt}}\}$  also provides information about peaks and valleys visited throughout the trip. Prospectively, the integral over the absolute change in geodetic altitude could help to also capture multiple peaks and valleys along a trip.

Although the features characterizing driving style may be correlated to each other, the combination of these features potentially exhibits additional information. Furthermore, features based on the EV's lateral acceleration—i.e., features characterizing driving style on curvy roads—have low correlation to  $\Gamma$ . Therefore, none of these features has been selected, see [Table 3.1](#).

After cleaning the data, i.e., removing invalid values and outliers, the total amount of trips recorded throughout two years is  $|\mathcal{J}| = 2146$ . The data is randomly shuffled and 80% thereof are selected as training set  $\mathcal{J}_{\text{train}}$ , with a cardinality of  $|\mathcal{J}_{\text{train}}| = 1717$ . Accordingly, 20% of the data is used as independent test set  $\mathcal{J}_{\text{test}}$ , with  $|\mathcal{J}_{\text{test}}| = 429$ . To guarantee an efficient training process of the Data-driven Model, mean and variance normalization is performed on the data using *SciKit-Learn* [254].

### 3.1.2 Data-driven Models to Estimate Energy Consumption

The proceeded data exploration as described in [Section 3.1.1](#) revealed the importance of single features over others. However, the specific dependency between the input features  $\mathbf{x}_\Gamma$  and the output label  $\Gamma$  is still unknown. For this reason, three machine learning models of different complexity are presented in this section to estimate the specific energy consumption  $\Gamma$ .

### 3.1.2.1 Regression using a Linear Approximation

For the initial approach, a linear regression model, e.g., as described in [255], is used. To this end, the input vector  $\mathbf{x}_{\Gamma,j}$  is multiplied by the weight matrix  $\mathbf{W} \in \mathbb{R}^{1 \times 11}$ , and added to the bias  $b \in \mathbb{R}$ . Hence, the estimated energy consumption

$$\hat{\Gamma}_{\text{DD,LR}} = \mathbf{W} \cdot \mathbf{x}_{\Gamma} + b, \quad (3.4)$$

is obtained. The components of the weight matrix  $\mathbf{W}$  and the value of  $b$  are determined from the training data  $\mathcal{J}_{\text{train}}$ , such that

$$\min_{\mathbf{W} \in \mathbb{R}^{1 \times 11}, b \in \mathbb{R}} \{q\text{MSE}(\Gamma_j, \hat{\Gamma}_{\text{DD,LR},j})\}, \quad \forall j \in \mathcal{J}_{\text{train}}. \quad (3.5)$$

In particular, the target is to minimize the Mean Squared Error (MSE) (see Section A.2.4) of actual energy consumption  $\Gamma$  and estimated energy consumption  $\hat{\Gamma}_{\text{DD,LR}}$  for all training samples  $j \in \mathcal{J}_{\text{train}}$ . The model is implemented in *Python* [256] using *SciKit-Learn* [254].

### 3.1.2.2 Regression using an Artificial Neural Network

The interdependence between the input features  $\mathbf{x}_{\Gamma}$  and the energy consumption  $\Gamma$  may not be linear. Hence, an ANN regression model is designed, as ANNs are able to represent highly non-linear behavior. For the according fundamentals of ANNs in engineering applications, see [257].

To ensure a proper model performance, different combinations of hyperparameters are evaluated and tuned accordingly. In Table 3.2, all tested and finally selected values of hyperparameters are listed. Preliminary tests especially focusing on the

**Table 3.2:** Tested and selected hyperparameters for ANN regression model to calculate  $\hat{\Gamma}_{\text{DD,MLP}}$ .

Hyperparameter	Tested Values	Selected Value
Number of Hidden Layers	1, 2, 3, 4, 5	3
Number of Neurons per Hidden Layer	5, 10, 20, 50	20
Activation Function	sigmoid, rectified linear unit	sigmoid
Batch Size for Model Training	32, 64	32
Epochs of Model Training	up to 5000	200 (with <i>early stopping</i> )
Loss Function for Model Training	MSE	MSE
Learning Rate of Optimizer	0.001	0.001

number of hidden layers and the number of neurons per hidden layer showed that

advancing the interior model architecture may improve the model’s performance, see [Table 3.3](#).

**Table 3.3:** Result of preliminary tests to examine different architectures of the ANN model to calculate  $\hat{I}_{DD,MLP}$ .

Test Run	Hidden Layers	Neurons	$q_{MSE}$	$q_{RMSE}$	$q_{MAE}$	$q_{RMAE}$
1	1	50	0.084	0.290	0.138	0.610
1	3	20	0.071	0.267	0.078	0.257
2	1	50	0.060	0.246	0.110	0.419
2	3	20	0.047	0.217	0.085	0.328

As a consequence, a Multi Layer Perceptron (MLP) with three hidden layers is used to calculate  $\hat{I}_{DD,MLP}$ . Each hidden layer consists of 20 neurons that possess a sigmoid activation function. The output layer only consists of one neuron activated by a linear function. For model training, batches of 32 trips are fed to the model. To avoid overfitting the model, the gradient of the model’s loss function is monitored in each training epoch; in this way, 200 epochs of model training appeared to be sufficient (*early stopping*). Using the MSE as loss function and the *Adam* optimizer [258], the incremental learning process is proceeded with a learning rate of 0.001. The model is implemented in *Python* [256] using *Tensorflow* [259].

### 3.1.2.3 Regression using a Random Forest

Besides ANNs, random forests can also be used for modeling complex relations between input features and output label. The advantage of random forests over ANNs, however, is given by less computational effort and smaller amounts of data required for training. To train a random forest regressor, multiple decision trees are trained with randomly selected subsets of the training data  $\mathcal{J}_{train}$ . Each decision tree creates an individual estimation of the desired output  $\Gamma$ . Then, the average of these estimations is used as ensemble estimation  $\hat{I}_{DD,RF}$ . Further details on the fundamentals of random forest regression can be found in [260].

The training process can be automated using a programming framework. However, hyperparameters also need to be adapted for a proper model performance. [Table 3.4](#) presents the manually tuned parameters used in this study. The entire random forest consists in total of 100 separate decision trees. Each decision tree is allowed to have a maximum depth of ten decision nodes. Furthermore, each internal node requires at least two training samples to be split. For training each

**Table 3.4:** Selected hyperparameters for random forest regression model to calculate  $\hat{\Gamma}_{DD,RF}$ .

Hyperparameter	Selected Value
Total Number of Decision Trees	100
Maximum Depth of Decision Trees	10 nodes
Samples to Split Internal Node	2
Loss Function for Model Training	MSE

individual decision tree, the **MSE** is used as loss metric. The model is implemented in *Python* [256] using *SciKit-Learn* [254].

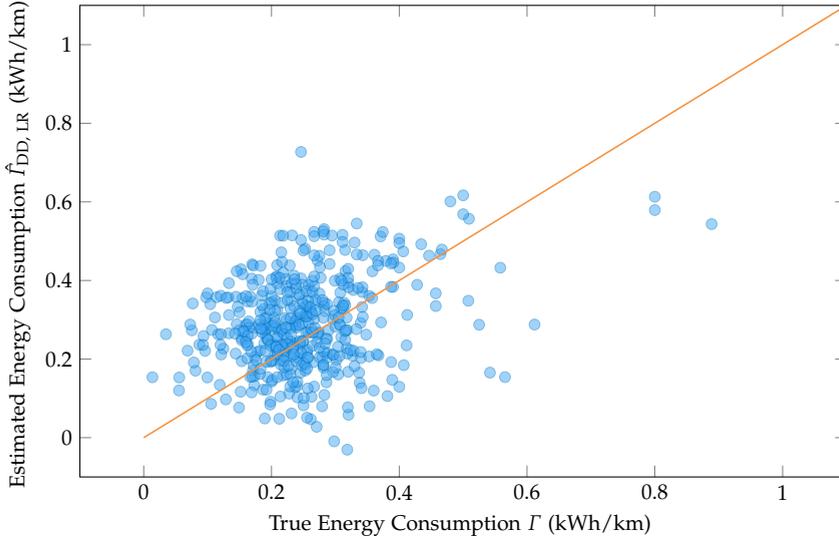
### 3.1.3 Validation of Consumption Model

After training the data-driven machine learning models (see [Section 3.1.2.1](#), [Section 3.1.2.2](#), and [Section 3.1.2.3](#)) on the training data set  $\mathcal{J}_{\text{train}}$ , the independent test data set  $\mathcal{J}_{\text{test}}$  is used for validation. To quantify the model performance, the error metrics (A.7)-(A.10) are calculated. For comparison, the performance of the baseline model and all data-driven models is listed in [Table 3.5](#).

**Table 3.5:** Comparison of baseline model and data-driven models according to the error metrics **MSE**  $q_{\text{MSE}}$  (in  $(\text{kWh}/\text{km})^2$ ), **Root Mean Squared Error (RMSE)**  $q_{\text{RMSE}}$  (in  $\text{kWh}/\text{km}$ ), **Mean Absolute Error (MAE)**  $q_{\text{MAE}}$  (in  $\text{kWh}/\text{km}$ ), and **Relative Mean Absolute Error (RMAE)**  $q_{\text{RMAE}}$ , on the independent test data set  $\mathcal{J}_{\text{test}}$ .

Model		$q_{\text{MSE}}$	$q_{\text{RMSE}}$	$q_{\text{MAE}}$	$q_{\text{RMAE}}$
Baseline	$\hat{\Gamma}_{\text{BL}}$	0.209	0.457	0.172	0.482
Linear Regression	$\hat{\Gamma}_{\text{DD,LR}}$	0.180	0.424	0.142	0.532
<b>ANN</b> Regression	$\hat{\Gamma}_{\text{DD,MLP}}$	0.071	0.267	0.078	0.257
Random Forest Regression	$\hat{\Gamma}_{\text{DD,RF}}$	0.044	0.210	0.062	0.214

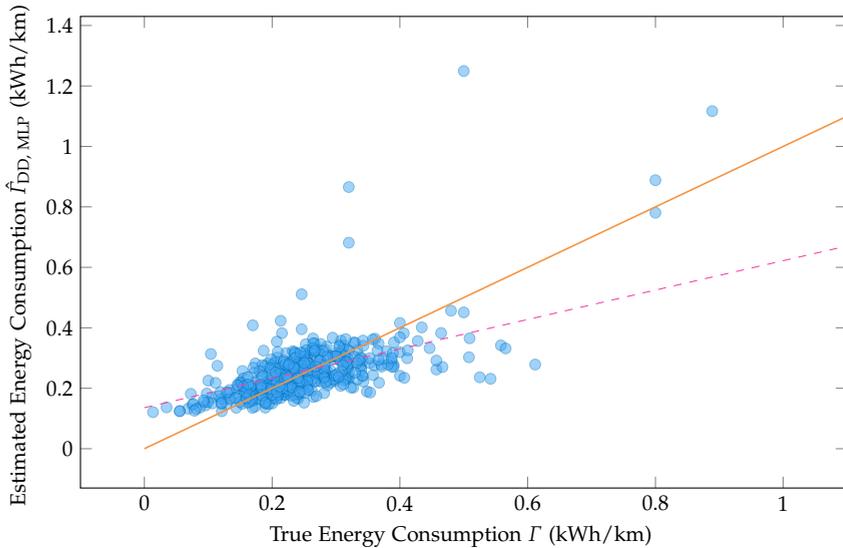
Using the baseline model as benchmark, the linear regression model slightly improves the estimation error. For instance, the Mean Absolute Error (**MAE**) of the linear regression model is improved by 17.44% compared with the baseline model, see [Table 3.5](#). Yet, the linear regression model estimates most of the test samples incorrectly, as [Figure 3.4](#) indicates. Furthermore, neither systematic overestimation nor underestimation of values can be observed. Thus, the relation between input features  $\mathbf{x}_T$  and the true energy consumption  $\Gamma$  is assumed to be non-linear. Compared with the linear regression model, the **ANN** model yields significantly more accurate estimations  $\hat{\Gamma}_{\text{DD,MLP}}$ . In particular, the **ANN** model reduces the **MAE** of



**Figure 3.4:** Comparison of energy consumption estimation  $\hat{\Gamma}_{DD,LR}$  of the linear regression model (see Section 3.1.2.1) with the true energy consumption  $\Gamma$  of recorded EV trips in independent test set  $\mathcal{J}_{test}$ ; the orange line indicates ideal model behavior.

estimated energy consumption by 54.65% compared with the baseline model, see Table 3.5. In Figure 3.5 the ANN model estimations on the independent test set  $\mathcal{J}_{test}$  are also visualized. This result supports the assumption of present non-linearity between  $x_\Gamma$  and the true energy consumption  $\Gamma$ . The performance of both random forest regression and ANN regression is comparable. Yet, for the chosen test data set  $\mathcal{J}_{test}$ , the random forest regression outperforms all other models, see Table 3.5 and Figure 3.6. The estimation error can significantly be improved compared with both the linear regression and baseline model. For instance, the random forest regression cuts the MAE by 63.95% compared with the baseline model. Overall, the results show that all data-driven models to estimate the specific energy consumption  $\Gamma$  outperform the baseline model. Hence, the hypothesis that known driving behavior of the EV user enhances consumption estimations, can be confirmed. The developed random forest regression model seems to most adequately represent the non-linear relation between  $x_\Gamma$  and  $\Gamma$ , and should therefore be used. For this purpose, however, both the availability of the required data and the EV user's consent to the use of this data are essential prerequisites.

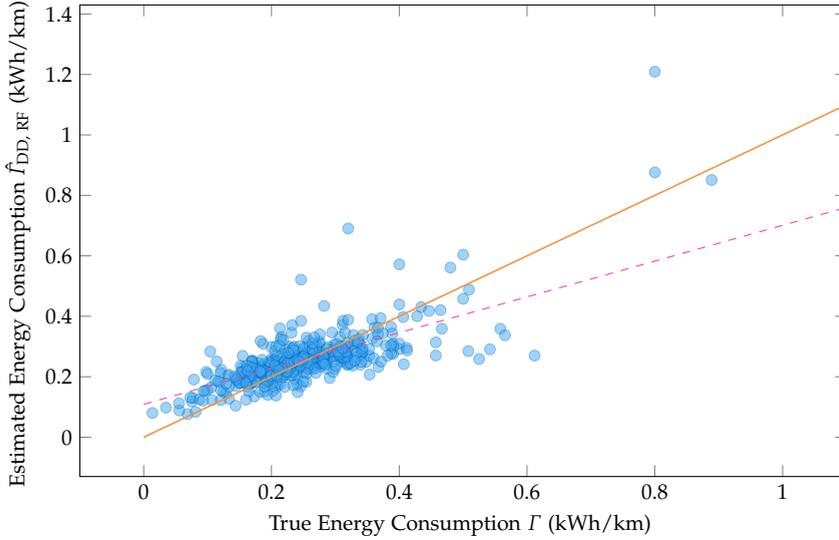
On the upper end of the independent test set  $\mathcal{J}_{test}$  some outliers occur for  $\Gamma$ . In real applications of the model, such high values of  $\Gamma$  usually occur for very short



**Figure 3.5:** Comparison of energy consumption estimation  $\hat{\Gamma}_{DD,MLP}$  of the ANN regression model (see Section 3.1.2.2) with the true energy consumption  $\Gamma$  of recorded EV trips in independent test set  $\mathcal{J}_{test}$ ; the orange line indicates ideal model behavior; the dashed pink line indicates a systematic misrepresentation.

trips, e.g., taken in cold ambient conditions. Hence, the absolute deviation of a trip’s total energy consumption is comparably small. Compared with the linear regression model, both the random forest regression the ANN model more accurately estimates these outliers, cf., Figure 3.4, Figure 3.5 and Figure 3.6. Furthermore, the random forest model also more accurately estimates outliers of  $\hat{\Gamma}_{DD,MLP}$  produced by the ANN model, cf., Figure 3.5 and Figure 3.6.

Figure 3.5 shows that the ANN model appears to possess a systematic misrepresentation. Particularly, the energy consumption is slightly overestimated for trips with a true energy consumption of less than approx. 0.27 kWh/km. On the other hand, for trips with true energy consumption higher than approx. 0.27 kWh/km, the energy consumption is slightly underestimated. To illustrate this behavior, the dashed pink line in Figure 3.5 indicates the trend of the estimations  $\hat{\Gamma}_{DD,MLP}$ . Similar to the ANN model, the estimations  $\hat{\Gamma}_{DD,RF}$  of the random forest model show a slight systematic misrepresentation; again, see the dashed pink line in Figure 3.6. This behavior could be explained by the fact that high values of  $\Gamma$  mostly occur for short trips with a high amount of auxiliary energy consumption; this comprises e.g., active battery temperature adaption, air conditioning, or interior and seat heat-



**Figure 3.6:** Comparison of energy consumption estimation  $\hat{\Gamma}_{DD,RF}$  of the random forest regression model (see Section 3.1.2.3) with the true energy consumption  $\Gamma$  of recorded EV trips in independent test set  $\mathcal{J}_{test}$ ; the orange line indicates ideal model behavior; the dashed pink line indicates a systematic misrepresentation.

ing. In the present setup, however, there is no feature in  $\mathbf{x}_\Gamma$  that represents the energy consumption of these auxiliary consumers. By introducing an additional feature for auxiliary energy consumption, a further improvement of the estimation accuracy of the data-driven models is expected. However, this feature also has to be predicted for future trips. To this end, an additional model would be required, possibly introducing further estimation errors. Furthermore, data on auxiliary energy consumption in sufficient amount and quality would be required to both design and validate such a model.

In addition to specific energy consumption, prospectively the deviation of total energy consumption per trip could also be employed as error metric. In this way, very short—and thus irrelevant—trips would be effectively suppressed for model training. As a consequence, the overall model performance is expected to be improved.

### 3.2 Battery Model

The models presented in this section help to estimate the state of the EV battery throughout a charging event  $k$ . Each charging event that starts at arrival time  $t_{k,0}$  and ends at departure time  $t_{k,N_k}$  is discretized in time with  $\Delta t$ . In this way, the charging event's time horizon  $[t_{k,0}, t_{k,N_k}]$  is divided into  $N_k$  time intervals of duration  $\Delta t$  and  $N_k + 1$  states. Accordingly, the set of time intervals is defined as

$$\mathcal{N}_k = [0, N_k - 1] \subset \mathbb{N}. \quad (3.6)$$

All time intervals  $n \in \mathcal{N}_k$  start at time  $t_n$  and end at time  $t_{n+1}$ . The battery state at time  $t_n$  is characterized by the battery energy content  $e_n$ , the internal battery temperature  $\theta_n$ , and the maximum battery capacity  $e_{\max,n}$ . For a general quantification of the battery state—e.g., for different types of batteries—,  $e$  is normalized as SOC  $S_C$  according to (1.1); similarly,  $e_{\max}$  is normalized as SOH  $S_H$  according to (1.2). To calculate the evolution of  $e$ ,  $\theta$ , and  $e_{\max}$  over time, the battery's dynamic behavior is represented in three models, see Figure 3.7. First, an **electrical model** estimates the

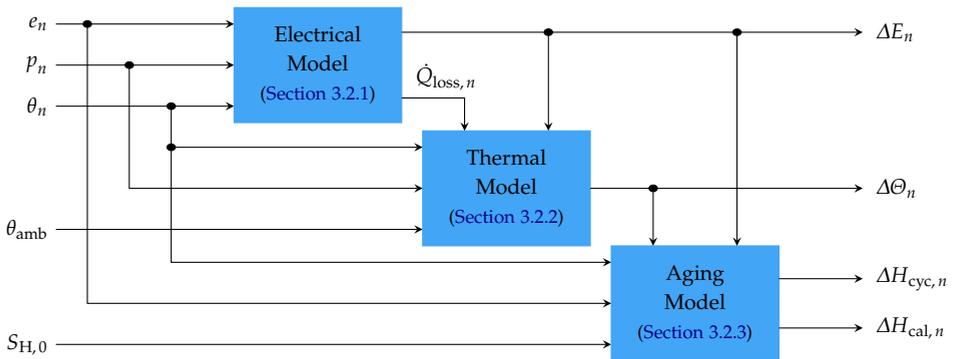


Figure 3.7: Schematic block diagram of EV battery model and its components.

change of battery energy  $\Delta E_n$  and the internal losses  $\dot{Q}_{\text{loss},n}$  in time interval  $n$ . Then, a **thermal model** estimates the change of battery temperature  $\Delta\theta_n$  in time interval  $n$ . Finally, an **aging model** estimates both cyclic battery aging increment  $\Delta H_{\text{cyc},n}$  and calendar aging increment  $\Delta H_{\text{cal},n}$  in time interval  $n$ . The combination of these three models is referred to as the **battery model**. All individual components are described in Section 3.2.1, Section 3.2.2, and Section 3.2.3. Note that these models can be vehicle-, vehicle-type-, battery-, or battery-type-specific; a generic reuse is thus limited. General initialization followed by incremental adaption, however, is conceivable.

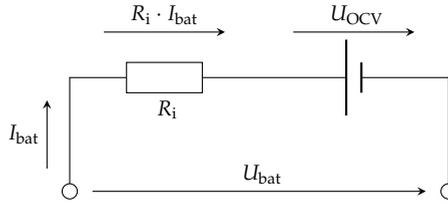
### 3.2.1 Electric Battery Model

The energy level  $e$  of the EV battery—and thus the SOC  $S_C$ —changes with surrounding influences, especially the charging power  $p$ . An electrical battery model helps to calculate the evolution of the battery energy. For this purpose, the electrical battery model, as shown in Figure 3.7, estimates the energy throughput of the battery

$$\Delta E_n = e_{n+1} - e_n, \quad \forall n \in \mathcal{N}_k, \quad (3.7)$$

for a given time interval  $n$ , battery temperature  $\theta_n$ , battery energy  $e_n$ , and charging power  $p_n$ . This model might represent dedicated power electronics, hence it is vehicle-specific.

The EV battery usually consists of several battery cells connected to a combined series-parallel circuit. Using THÉVENIN'S theorem [261], the Equivalent Circuit Model (ECM) as presented in Figure 3.8 is designed.<sup>5</sup> The model consists of the



**Figure 3.8:** Equivalent circuit model of an EV battery for low-dynamic operation with internal resistance  $R_i$  and voltage source  $U_{\text{OCV}}$

voltage source  $U_{\text{OCV}}$  serially connected with the internal resistance  $R_i$ . Both battery current  $I_{\text{bat}}$  and battery voltage  $U_{\text{bat}}$  are quantities that can be measured at the battery's terminals. Following KIRCHHOFF'S voltage law (also known as KIRCHHOFF'S loop rule), the term

$$U_{\text{OCV},n} = U_{\text{bat},n} - R_{i,n} \cdot I_{\text{bat},n}, \quad (3.8)$$

mathematically describes the ECM in Figure 3.8. Using the convention that the battery current  $I_{\text{bat}}$  is greater than zero during charging,  $U_{\text{bat}} > U_{\text{OCV}}$  for charging

<sup>5</sup>THÉVENIN'S theorem states that any linear electrical network consisting of resistances, current and voltage sources may be replaced by an equivalent combination of a voltage source serially connected to a resistance [261]. Hence, the name Equivalent Circuit Model.

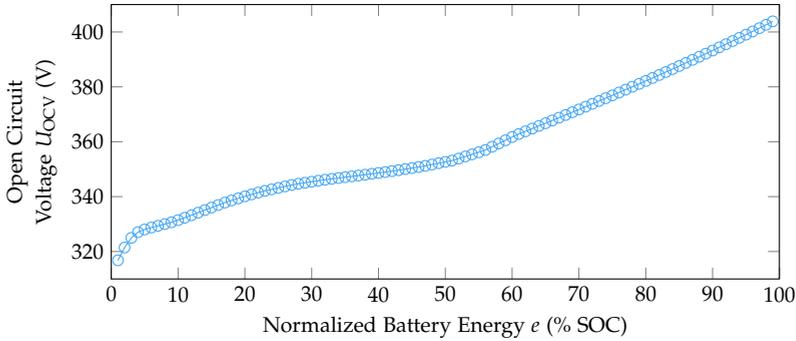
and  $U_{\text{bat}} < U_{\text{OCV}}$  for discharging; accordingly,  $U_{\text{bat}} = U_{\text{OCV}}$ , if no external load is connected, i.e., if  $I_{\text{bat}} = 0$ . Substituting the terminal voltage  $U_{\text{bat}}$  in (3.8) with

$$U_{\text{bat},n} = \frac{p_n}{I_{\text{bat},n}}, \quad (3.9)$$

and solving (3.8) for the battery current  $I_{\text{bat}}$  yields

$$I_{\text{bat},n} = \frac{-U_{\text{OCV},n} + \sqrt{U_{\text{OCV},n}^2 + 4R_{i,n} \cdot p_n}}{2R_{i,n}}; \quad (3.10)$$

despite two possible solutions only the greater one is physically feasible, see also [262]. The gross charging power  $p_n$  consumed from the charging station is assumed to remain constant throughout a single time interval  $n, \forall n \in \mathcal{N}$ . In (3.10), the value of  $U_{\text{OCV},n}$  depends on the battery energy  $e_n$  and is obtained from the characteristic curve in Figure 3.9. In a similar way,  $R_{i,n}$  depends on the battery temperature  $\theta_n$

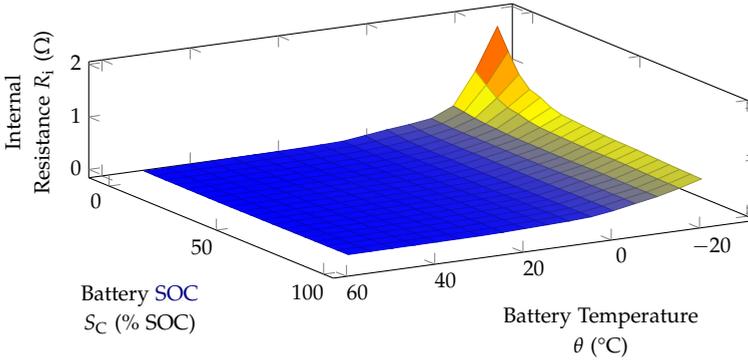


**Figure 3.9:** Characteristic curve of open circuit voltage  $U_{\text{OCV}}$  with respect to normalized battery energy  $e$ .

and the battery energy  $e_n$ . The according values of  $R_{i,n}$  are obtained from the characteristic curve in Figure 3.10. Similar to  $p_n$ , the values of both  $U_{\text{OCV},n}$  and  $R_{i,n}$  are also assumed to be constant throughout a single time interval  $n$ . Hence, the ECM as shown in Figure 3.8 is only suitable to represent quasi-stationary operation such as charging the EV battery. For high-dynamic operation, such as discharging the EV battery while driving, an advanced ECM, e.g., with additional resistor-capacitors-pairs would be required [263]. Note that the characteristic curves in Figure 3.9 and Figure 3.10 are battery-type-specific and may be replaced for different types of batteries.

Once the battery current  $I_{\text{bat}}$  is calculated, the OHmic [264] losses

$$\dot{Q}_{\text{loss},n} = R_{i,n} \cdot I_{\text{bat},n}^2 \quad (3.11)$$



**Figure 3.10:** Characteristic curve of internal battery resistance  $R_i$  with respect to battery SOC  $S_C$  and battery temperature  $\theta$ .

within the battery may also be calculated. Given the charging power  $p_n$ , finally the energy throughput

$$\Delta \hat{E}_n = \Delta t \cdot (p_n - \dot{Q}_{\text{loss},n}), \quad (3.12)$$

for each time interval  $n$  is obtained.

Note that  $\dot{Q}_{\text{loss},n} > 0$  occurs both while charging and discharging. Hence, it decreases  $|\Delta E_n|$  during charging and increases  $|\Delta E_n|$  during discharging. Furthermore, the losses  $\dot{Q}_{\text{loss}}$  are dissipated into heat and thus may increase the battery temperature  $\theta$  both while charging and discharging, see also [Section 3.2.2](#).

### 3.2.2 Thermal Battery Model

The internal temperature  $\theta$  of the EV battery affects both battery parameters, such as  $R_i$ , and battery aging, see Section 3.2.3. Hence, a thermal battery model is required to estimate the change of battery temperature

$$\Delta\theta_n = \theta_{n+1} - \theta_n, \quad \forall n \in \mathcal{N}_k, \quad (3.13)$$

for a given time interval  $n$ , battery temperature  $\theta_n$ , battery energy  $e_n$ , ambient temperature  $\theta_{\text{amb}}$ , and charging power  $p_n$ . Due to heat exchange with surrounding components, these models are mostly vehicle-type-specific.

In existing applications found in the literature (see Section 1.1.2.3), mostly reduced-order models are used to estimate  $\Delta\theta$ . Such models describe the simplified thermal battery behavior on a macroscopic level. For this purpose, heat flows into and out of the battery are balanced, based on the first principle of thermodynamics. Particularly, the differential formulation of the battery's thermal energy balance

$$\frac{d\theta}{dt} = \frac{\dot{Q}_{\text{loss}} + \dot{Q}_{\text{amb}} + \dot{Q}_{\text{BMS}}}{c_h}, \quad (3.14)$$

is consulted. Here,  $c_h$  represents the heat capacity of the battery. The power loss  $\dot{Q}_{\text{loss}}$  is calculated beforehand using (3.11) (see Section 3.2.1) and is assumed to be completely dissipated into heat. Furthermore, the heat flow

$$\dot{Q}_{\text{amb}} = -\alpha \cdot (\theta - \theta_{\text{amb}}), \quad (3.15)$$

describes conductive heat exchange between the battery and the environment;  $\alpha$  represents the specific heat transition coefficient of the battery system. The direction of  $\dot{Q}_{\text{amb}}$  depends on the difference between  $\theta$  and  $\theta_{\text{amb}}$ . If  $\theta > \theta_{\text{amb}}$ ,  $\dot{Q}_{\text{amb}}$  is negative, i.e., heat flows out of the battery; accordingly, if  $\theta < \theta_{\text{amb}}$ ,  $\dot{Q}_{\text{amb}}$  is positive, i.e., heat flows into the battery. To maintain the battery temperature within a range that guarantees efficient and safe operation, the BMS can actively add the heat flow  $\dot{Q}_{\text{BMS}}$ ; heat may therefore either be injected for heating the battery, i.e.,  $\dot{Q}_{\text{BMS}} > 0$ , or withdrawn for cooling the battery, i.e.,  $\dot{Q}_{\text{BMS}} < 0$ .

Applying discrete-time Forward-EULER integration [265] to (3.14) yields the battery temperature change

$$\Delta\hat{\theta}_{\text{RO},n} = \Delta t \cdot \frac{\dot{Q}_{\text{loss},n} + \dot{Q}_{\text{amb},n} + \dot{Q}_{\text{BMS},n}}{c_h}, \quad (3.16)$$

of the reduced-order thermal battery model. Preliminary tests, however, show that the reduced-order model inadequately represents the battery's thermal behavior for

some cases. Therefore, the design of a data-driven thermal model is described in the following. The results of the reduced-order thermal model are consulted as benchmark for validation purposes in [Section 3.2.4.2](#).

### 3.2.2.1 Selection and Preprocessing of Training Data

To improve the estimations of reduced-order thermal models, data-driven approaches use additional information to estimate the change of battery temperature  $\Delta\Theta_{\text{DD}}$ . In this way, hidden thermodynamic processes, e.g., electro-chemical heat sources or sinks can be captured. The required training data for these models is extracted from 279 real-world charging events that are recorded with a sampling rate of approx. 30 s from batteries installed and operated in EVs. Events containing gaps greater than 90 s are screened out and resampled to the regular 30 s sampling rate using linear interpolation. Then, a rolling time window of duration  $\Delta t = 5$  min is applied to obtain a set of individual samples  $\mathcal{B}_k$  for each charging event  $k \in \mathcal{K}$ . For model design and validation, the samples of all charging events  $k \in \mathcal{K}$  are united as set of historical battery data samples

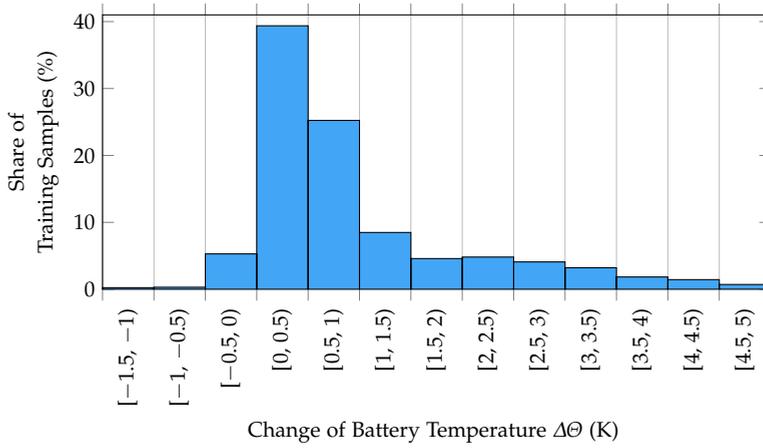
$$\mathcal{B} = \bigcup_{\forall k \in \mathcal{K}} \mathcal{B}_k. \quad (3.17)$$

Subsequently, the difference of battery temperature

$$\Delta\Theta_b = \theta_{b+1} - \theta_b, \quad \forall b \in \mathcal{B}, \quad (3.18)$$

throughout each training sample  $b$  is calculated as regression target for the data-driven models. [Figure 3.11](#) shows a histogram of  $\Delta\Theta$  for all samples in  $\mathcal{B}$ . On average, the change of battery temperature amounts to 0.961 K for each time interval of 5 min. In most cases,  $\Delta\Theta$  is positive, i.e., the battery temperature rises, e.g., due to internal charging losses. Only a few samples exist in which the battery temperature decreases, i.e.,  $\Delta\Theta$  is negative. These cases may occur, e.g., after long trips in which  $\theta$  is high; given a lower ambient temperature  $\theta_{\text{amb}}$  compared with the battery temperature  $\theta$ , the heat exchange with the environment may thus cool down the battery slowly.

To the end of relevant features, all available signals are manually screened based on basic knowledge of EV battery modeling, see [Section 1.1.2](#). The leftmost column of [Table 3.6](#) presents the most promising features. The values of the features  $e$ ,  $\theta$ ,  $I_{\text{bat}}$ , and  $p_b$  are taken at the beginning of each sample  $b$  and assumed to be constant for the time interval of  $\Delta t = 5$  min. For simplicity, the ambient temperature  $\theta_{\text{amb}}$  is assumed to remain constant throughout one charging event  $k$  and thus also for each



**Figure 3.11:** Histogram of battery temperature change  $\Delta\Theta_b$  for all data samples  $b$  in battery data set  $\mathcal{B}$ .

**Table 3.6:** Selection of thermal battery model features according to univariate feature relevance based on the SPEARMAN-correlation coefficient (see Section A.1.1) in relation to the change of battery temperature  $\Delta\Theta$  and manual screening.

Feature		SPEARMAN-Correlation $\rho_{\text{Sp}}(\bullet, \Delta\Theta)$	Selected
Battery Energy	$e_b$	-0.48	no
Battery Temperature	$\theta_b$	-0.32	yes
Battery Current	$I_{\text{bat},b}$	0.46	no
Gross Charging Power	$p_b$	0.43	yes
Ambient Temperature	$\theta_{\text{amb}}$	-0.18	no
Difference to Ambient Temperature	$\theta_b - \theta_{\text{amb}}$	-0.12	no
Battery Internal Resistance	$R_{i,b}$	0.24	no
Open Circuit Voltage	$U_{\text{OCV},b}$	-0.48	no
Energy Throughput	$\Delta E_b$	0.41	yes
(Estimated) Internal Losses	$\dot{Q}_{\text{loss},b}$	0.52	yes

training sample  $b \in \mathcal{B}_k$ ; advanced models could include a forecast of  $\theta_{\text{amb}}$  for the charging event's time window  $[t_{k,0}, t_{k,N_k}]$ . Note that  $R_{i,b}$  and  $U_{\text{OCV},b}$  are obtained from the characteristic curves in [Figure 3.9](#) and [Figure 3.10](#); furthermore,  $\Delta E_b$  and  $\dot{Q}_{\text{loss},b}$  are calculated beforehand using the ECM as presented in [Section 3.2.1](#). The developed data-driven thermal model is thus a *gray-box* approach.

To initially analyze the univariate relevance of these features to the label  $\Delta\Theta$ , the SPEARMAN-correlation coefficient (see [Section A.1.1](#)) is consulted, as presented in the middle column of [Table 3.6](#). Although the battery energy  $e$  shows a relatively strong correlation to  $\Delta\Theta$ , it is screened out due to a parallel correlation. In particular, the higher  $e$ , the less charging power  $p$  is applied and thus less internal losses occur to heat up the battery; instead,  $p$  and the energy throughput  $\Delta E$  are selected as features. The same principle applies to  $U_{\text{OCV}}$ , as it is calculated from  $e$ . In a similar way, the battery current  $I_{\text{bat}}$  is screened out, as it strongly correlates to the charging power  $p$  and the internal losses  $\dot{Q}_{\text{loss}}$ ; Hence, only  $\dot{Q}_{\text{loss}}$  is selected as feature. The internal resistance  $R_i$ , on the other hand, is screened out as it directly influences  $\dot{Q}_{\text{loss}}$  via the ECM. All features related to the ambient temperature  $\theta_{\text{amb}}$  show low correlations, and thus are not selected. Altogether, the input features of the data-driven thermal model can be represented by the feature vector

$$\mathbf{x}_{\Theta,b} = \begin{pmatrix} p_b \\ \dot{Q}_{\text{loss},b} \\ \Delta E_b \\ \theta_b \end{pmatrix}_b \in \mathbb{R}^4, \quad \forall b \in \mathcal{B}. \quad (3.19)$$

Since the EVs from which the training data is taken do not support bi-directional charging, discharging, e.g., for V2G applications as described in [Section 1.1.1](#), is underrepresented. The characteristic curves of both  $R_i$  and  $U_{\text{OCV}}$ , however, show similar values for charging and discharging. The battery's thermal behavior is thus assumed to be independent on the direction of  $p$  and  $\Delta E$ ; accordingly, absolute input values are used for these features.

After data cleaning, i.e., removing invalid values and outliers, the total number of battery data samples amounts to  $|\mathcal{B}| = 18868$ . This data is split in a training set  $\mathcal{B}_{\text{train}}$ , with a cardinality of  $|\mathcal{B}_{\text{train}}| = 15094$ . The remaining samples, i.e., the independent test set  $\mathcal{B}_{\text{test}}$  with  $|\mathcal{B}_{\text{test}}| = 3774$ , is used for validation. To guarantee an efficient training process of the data-driven models, the data is normalized using *SciKit-Learn* [254].

### 3.2.2.2 Data-driven Models to Estimate the Change of Battery Temperature

To the end of creating a data-driven model from the training data  $\mathcal{B}_{\text{train}}$ , two machine learning approaches are tested. First, a LR model is formulated that estimates

$$\Delta\hat{\Theta}_{\text{DD,LR}} = \mathbf{W} \cdot \mathbf{x}_{\Theta} + b, \quad (3.20)$$

based on the feature vector  $\mathbf{x}_{\Theta}$ . The determination of the weight matrix  $\mathbf{W}$  and the bias  $b$  follows the equivalent procedure as in Section 3.1.2.1. A more detailed description of the design and training of LR models can also be found in [255]. The model is implemented in *Python* [256] using the *SciKit-Learn* [254] framework.

Second, a more complex ANN model is designed, to also model possibly non-linear thermal behavior and hidden electro-chemical processes of the EV battery. The according fundamentals of ANNs in engineering applications can be found in [257]. To establish the best performing values for the ANN model hyperparameters, a *grid-search* is run that evaluates all combinations of hyperparameters as listed in Table 3.7; the rightmost column also reveals the actually selected value for each hyperparameter. To particularly evaluate the interior ANN architecture, preliminary

**Table 3.7:** Tested and selected hyperparameters for ANN regression model to calculate  $\Delta\hat{\Theta}_{\text{DD,MLP}}$ .

Hyperparameter	Tested Values	Selected Value
Number of Hidden Layers	1, 2, 3	2
Number of Neurons per Hidden Layer	5, 10, 20, 30, 40, 50	10
Activation Function	sigmoid, rectified linear unit	sigmoid
Batch Size for Model Training	32, 64, 128	128
Epochs of Model Training	up to 10000	3000 (with <i>early stopping</i> )
Loss Function for Model Training	MSE	MSE
Learning Rate of Optimizer	0.001	0.001

tests are run in which the number of hidden layers and the number of neurons per hidden layer is varied. The corresponding results are shown in Table 3.8.

It can be seen that for the given problem, ANN models with two hidden layers generally seem to outperform models with only one hidden layer. Furthermore, the number of neurons per hidden layer exhibits an optimum in the region of ten. Consequently, an MLP with two hidden layers is used to calculate  $\Delta\hat{\Theta}_{\text{DD,MLP}}$ . Each hidden layer consists of ten neurons that possess a sigmoid activation function; the output layer consists of only one neuron activated by a linear function. For model training, batches of 128 samples are fed to the model. To avoid overfitting the model, the gradient of the model's loss function is monitored in each training epoch; in this

**Table 3.8:** Result of three preliminary test runs to examine different architectures of the ANN model to calculate  $\Delta\hat{\Theta}_{DD,MLP}$ ; for evaluation, the  $R^2$ -Score  $q_{R^2}$  of estimated and true values (see (A.6)) is used.

Hidden Layers	Neurons	$q_{R^2}$ , Run 1	$q_{R^2}$ , Run 2	$q_{R^2}$ , Run 3
1	5	0.525	0.670	0.640
1	10	0.577	0.721	0.693
1	20	0.472	0.709	0.682
2	5	0.768	0.688	0.665
2	10	0.802	0.724	0.713
2	20	0.749	0.659	0.648

way, 3000 epochs of model training appeared to be sufficient (*early stopping*). Using the MSE as loss function and the Adam optimizer [258], the incremental learning process is proceeded with a learning rate of 0.001. The model is implemented in Python [256] using the Tensorflow [259] framework.

### 3.2.3 Battery Aging Model

Throughout the operation of EV batteries, irreversible physical and electro-chemical processes decrease the battery’s usable storage capacity  $e_{\max}$ . These degradation processes—also called *battery aging*—in turn cause the EV’s usable driving range and monetary value to decline. A model to quantify battery degradation is thus essential to design a charging assistant, see also Section 1.1.2.4. For this purpose, a semi-empirical degradation model is applied to calculate the progression of the SOH  $S_H$ , as given in (1.2). Particularly,

$$S_{H,n+1} = S_{H,n} + \Delta H_{\text{cyc},n} + \Delta H_{\text{cal},n}, \quad (3.21)$$

needs to be estimated for each time interval  $n \in \mathcal{N}_k$  of a charging event  $k$ . Here, one differentiates the *cyclic aging* increment  $\Delta H_{\text{cyc}}$  and *calendar aging* increment  $\Delta H_{\text{cal}}$ ; both are described in the following.

*Cyclic aging* is mainly caused by charging or discharging the EV battery. Among other processes, a loss of active lithium material occurs due to mechanical stress, see also [125]. For the battery cells used in this study, the cyclic aging increment

$$\Delta H_{\text{cyc},n} = \beta_A \cdot |\Delta E_n|^{\beta_B}, \quad (3.22)$$

only depends on the absolute energy throughput  $\Delta E_n$ . The degradation history, i.e., the SOH of previous time steps, is negligible. Thus,  $\Delta H_{\text{cyc}}$  is independent from  $S_H$ .

Additionally, high battery temperature, high SOC, over-voltage, and mechanical vibrations cause degradation of both active and inactive battery components, see also [125]. Hence, the battery capacity fades over time, regardless of the energy throughput (*calendar aging*). For the battery cells used in this work, the calendar aging increment

$$\Delta H_{\text{cal},n} = 1 - S_{H,0} + \beta_C \exp\left(\frac{\beta_D}{273\text{K} + \theta_n} + \beta_E e_n\right) \cdot (\Delta t + \tau_n)^{\beta_F}, \quad (3.23)$$

can be described based on a set of ARRHENIUS curves [266]. Here,  $S_{H,0}$  is the SOH at the beginning of the charging event  $k$ , i.e., at time  $t_{k,0}$ ;  $S_{H,0}$  serves as a reference for all time steps of charging event  $k$ , as calendar aging occurs on larger time scales (years) than charging (hours). Furthermore,

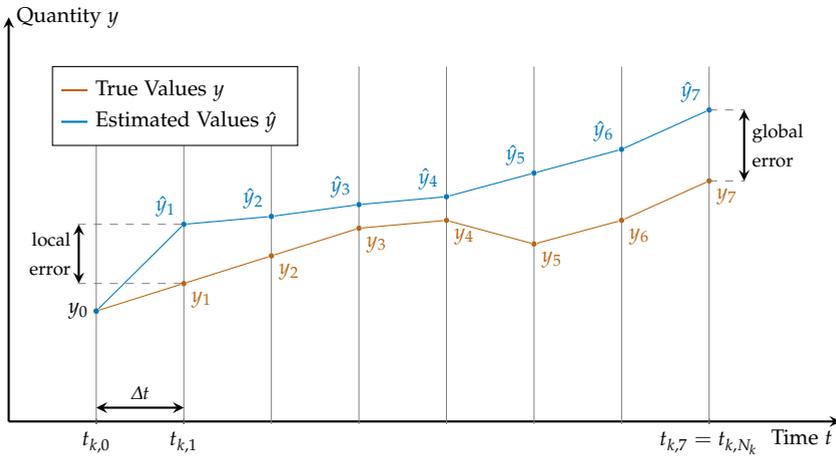
$$\tau_n = \left( \frac{S_{H,0} - 1}{\beta_C \exp\left(\frac{\beta_D}{273\text{K} + \theta_n} + \beta_E e_n\right)} \right)^{1/\beta_F}, \quad (3.24)$$

represents the equivalent battery age for each time interval  $n \in \mathcal{N}_k$  as a function of  $S_{H,0}$ . Thus,  $\Delta H_{\text{cal}}$  depends on the degradation history.

Both model characteristics and parameters are estimated from extensive cell tests at varying conditions, e.g., battery energy and temperature. Hence, the detailed parameters  $\beta_{A..F}$  represent battery-type-specific aging characteristics that are confidential. As these parameters may differ for different types of battery cells, the aging model may need to be replaced accordingly, e.g., with models as in [267]. Furthermore, the aging model is scaled from individual battery cells to the EV battery consisting of several interconnected cells. For this, a linear scaling is assumed. Production variance of battery cells, however, may yield deviating behavior as fast degrading cells can determine the SOH of the entire EV battery. Therefore, prospectively an implicit representation, e.g., via machine learning approaches as in [268] is also conceivable.

### 3.2.4 Validation of Battery Model

For validation of the battery model, measured time series data of 279 unidirectional charging events  $k \in \mathcal{K}$  is obtained from a fleet of ten real-world EVs. Here,  $\mathcal{K}$  denotes the set of available charging events  $k$ . To the end of data collection, all EVs are equipped with cloud-connected data loggers and utilize the framework as described in Section 2.3.1. After charging events are recorded with a sampling rate of approx. 30s, they are resampled to a time interval of  $\Delta t = 5$  min; each time interval equates to a battery data sample  $b$ . In this way, a set  $\mathcal{B}_k$  of historical battery data samples  $b$  is obtained for each charging event  $k$ . To quantify the estimation accuracy of the single battery model components, two measures are used. For the sake of comprehensibility, Figure 3.12 shows exemplary profiles of a quantity's true values  $y$  and estimated values  $\hat{y}$  over the time window  $[t_{k,0}, t_{k,N_k}]$  of charging event  $k$ ; the quantity  $y$  functions as a placeholder for battery state values, such as  $e$  or  $\theta$ .



**Figure 3.12:** Exemplary profiles of a quantity's true values  $y$  and estimated values  $\hat{y}$  over the time window  $[t_{k,0}, t_{k,N_k}]$  of charging event  $k$  to explain the error measures of the battery model (see Section 3.2).

The *local error* corresponds to the deviation between the estimated value  $\hat{y}$  and the actual value  $y$  in a single time step. To determine the local error, the battery model is applied to each data sample  $b \in \mathcal{B}_k$  of all charging events  $k \in \mathcal{K}$ .<sup>6</sup> Subsequently,

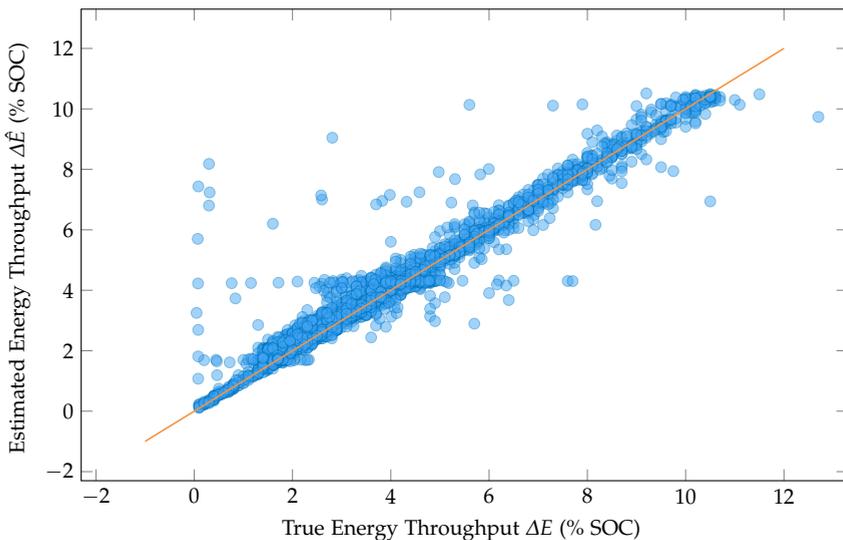
<sup>6</sup>For the data-driven thermal model (see Section 3.2.2) only the independent test data set  $\mathcal{B}_{\text{test}} = \bigcup_{k \in \mathcal{K}_{\text{test}}} \mathcal{B}_k$  is used for validation.

the estimated values are compared with the actual values of each sample using the Root Mean Squared Error (RMSE).

The *global error* mirrors the deviation  $\hat{y}_{N_k} - y_{N_k}$  of the estimated value  $\hat{y}_{N_k}$  and the actual value  $y_{N_k}$  at the end of a charging event  $k$  that consists of  $N_k$  time intervals. For this purpose, the battery model is repeatedly applied to all samples  $b$  out of the (ordered) set  $\mathcal{B}_k$ . This is, the estimation  $\hat{y}_b$  of one data sample  $b$  is used to estimate  $\hat{y}_{b+1}$  of the subsequent data sample  $b + 1$ . Finally, the deviation after the last sample equals the global error. For a general statement, the MAE is calculated from the global error of all charging events  $k \in \mathcal{K}$ .

### 3.2.4.1 Validation of Electrical Model

The electrical battery model as described in Section 3.2.1 yields an RMSE of 0.35 % SOC for single battery data samples  $b \in \mathcal{B}$  (local error). To illustrate this result, Figure 3.13 shows a scatter plot of estimated values  $\Delta\hat{E}$  over the true values  $\Delta E$ ; in addition, the orange line indicates ideal model behavior.

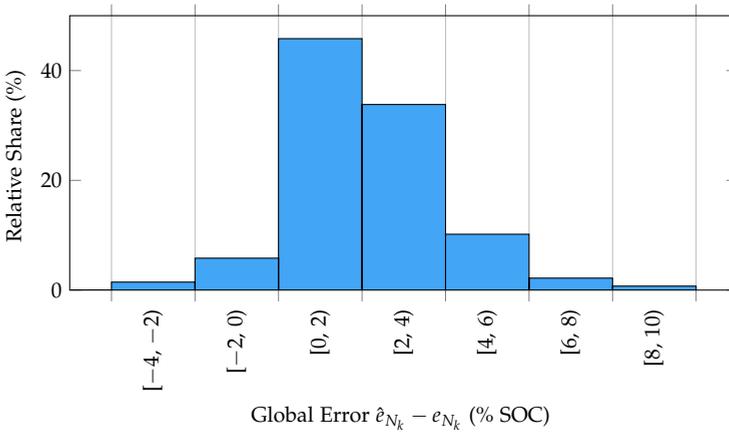


**Figure 3.13:** True and estimated energy throughput for individual samples  $b \in \mathcal{B}_k$  for all  $k \in \mathcal{K}$  (local error, normalized as SOC) using the electrical battery model (Section 3.2.1), the orange line indicates ideal model behavior.

It can be seen that the model estimations  $\Delta\hat{E}$  mostly fit the actual values  $\Delta E$ . This can be achieved, as both  $U_{OCV,n}$  and  $R_{i,n}$  are chosen from a characteristic curve for

each time interval  $n$  depending on  $e_n$  and  $\theta_n$ . Furthermore, high-dynamic changes of the battery energy are leveled out, as they occur on smaller time scales than the chosen  $\Delta t = 5 \text{ min}$  [263]. If charging the battery to  $e_{\max}$ , the BMS corrects the characteristic SOC curve towards the end of the charging event; hence, a few outliers occur, e.g., with  $\Delta E \cong 0$  and  $\Delta E \ll \Delta \hat{E}$ . Prospectively, these outliers could also be handled by the electrical battery model, if the update policy of the BMS is included in the estimation algorithm.

Figure 3.14 presents a histogram of the deviation  $\hat{e}_{N_k} - e_{N_k}$  (global error) at the end of all charging events  $k \in \mathcal{K}$ .



**Figure 3.14:** Deviation  $\hat{e}_{N_k} - e_{N_k}$  of true and estimated SOC at the end of all charging events  $k \in \mathcal{K}$  (global error) using the electrical battery model (Section 3.2.1).

The deviations are distributed off-centric around 0% SOC that would represent ideal model behavior. Hence, the ECM tends to overestimate a charging event's final SOC. However, the MAE calculated from these deviations only amounts to 1.896 kWh, or a normalized error of 2.37% SOC, respectively. This equals an acceptable driving range deviation of approx. 7.6 km. The accuracy of the electrical model is thus considered as sufficient. Furthermore, both the chosen time interval  $\Delta t = 5 \text{ min}$  and the ECM appear to be suitable to properly represent an EV's charging process.

### 3.2.4.2 Validation of Thermal Model

In [Section 3.2.2](#), different models to estimate the change of battery temperature are described. Accordingly, battery temperature profiles are calculated with these models and compared with actual measured temperature profiles of real charging events  $k \in \mathcal{K}_{\text{test}}$ .<sup>7</sup> [Table 3.9](#) presents the validation results by means of local and global error metrics as described at the beginning of [Section 3.2.4](#).

**Table 3.9:** Local and global error of thermal battery models ([Section 3.2](#)).

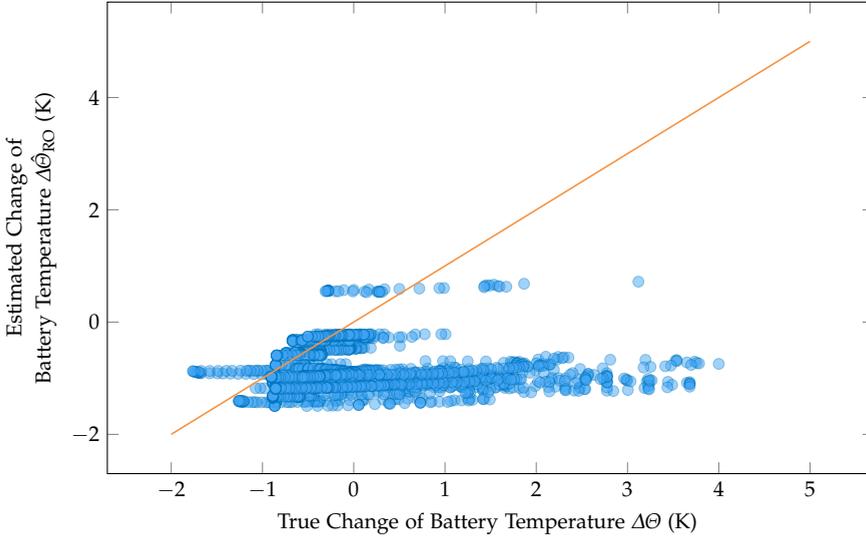
Model	Local Error (RMSE)	Global Error (MAE)
Constant Battery Temperature	0.72 K	7.57 K
Reduced-Order Thermal Model	1.56 K	4.78 K
LR Thermal Model	0.76 K	4.18 K
ANN Thermal Model	0.29 K	1.96 K

To benchmark the different thermal modeling approaches, constant battery temperature is assumed as naive baseline approach. Particularly, the battery temperature  $\theta_n$  is assumed to equal a charging event’s initial battery temperature  $\theta_0$  in all time intervals  $n \in \mathcal{N}$ . Accordingly, the change of battery temperature  $\Delta\theta_n = 0.0\text{K}$ , for all time intervals  $n \in \mathcal{N}$ . Note that the ambient temperature  $\theta_{\text{amb}}$  does not have any effect in this model. In comparison with the actual measured battery temperature profiles of real charging events, assuming constant battery temperature yields a local error [RMSE](#) of 0.72 K for single time intervals, see [Table 3.9](#). The global error [MAE](#), i.e., the mean temperature deviation at the end of a charging event amounts to 7.57 K. For estimating the battery temperature in single time intervals, assuming constant battery temperature suffices as the local error of 0.72 K indicates. However,  $\theta$  generally rises with the progression of a charging event. The longer a charging event lasts, the greater the propagated error of assuming constant battery temperature will thus be. This is also reflected in the significant global error of 7.57 K that may result in false conclusions, e.g., regarding temperature-dependent battery aging, see also [\(3.23\)](#) and [Section 1.1.2.4](#).

Validating the reduced-order thermal battery model as described at the beginning of [Section 3.2.2](#) reveals a slightly different result, see [Table 3.9](#). The battery temperature estimations per time interval (local error) yield an [RMSE](#) of 1.56 K—a lower accuracy

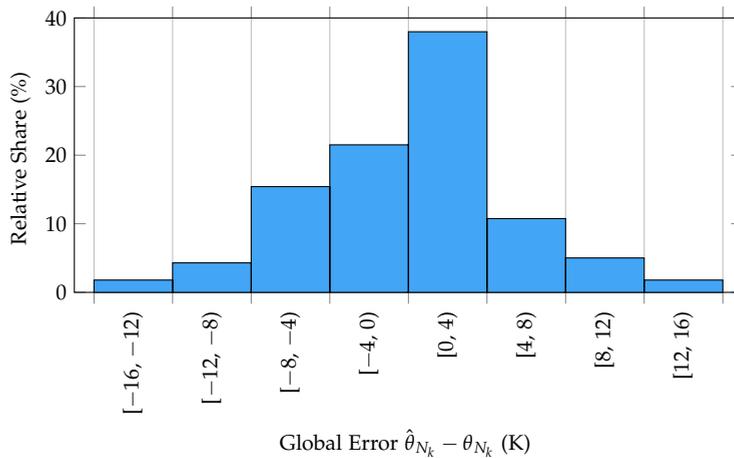
<sup>7</sup>For the data-driven approaches, a subset  $\mathcal{K}_{\text{train}} \subset \mathcal{K}$  is used for model training; for this reason, only the independent test set  $\mathcal{K}_{\text{test}} = \mathcal{K} \setminus \mathcal{K}_{\text{train}}$  is used for validation here.

compared with assuming constant battery temperature. Figure 3.15 shows a scatter plot of true and estimated battery temperature change per time interval.



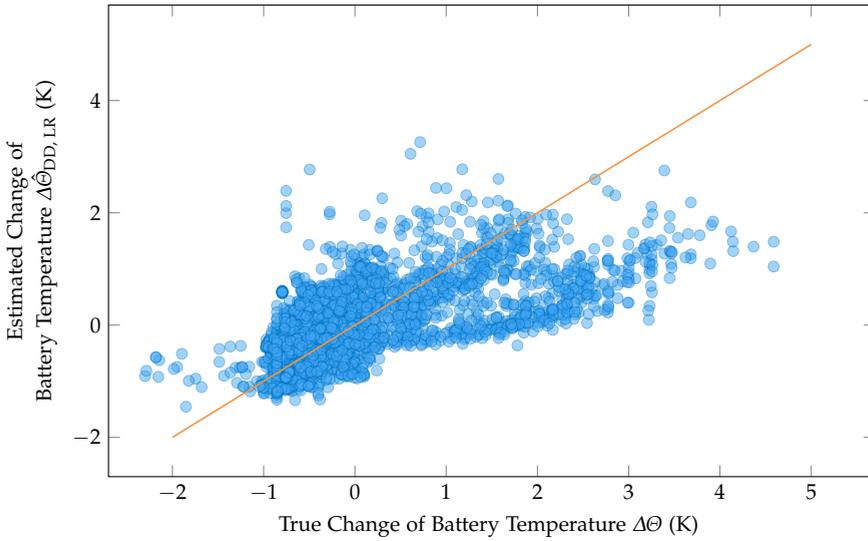
**Figure 3.15:** True and estimated change of battery temperature for individual samples  $b \in \mathcal{B}_k, \forall k \in \mathcal{K}$  (local error) using the reduced-order thermal battery model (Section 3.2.2), the orange line indicates ideal model behavior.

It can be seen directly that the reduced-order model mostly underestimates the real change in battery temperature  $\Delta\Theta$ . When repeatedly applying the reduced-order model, the mean temperature deviation at the end of a charging event, i.e., the global error MAE is 4.783 K. In Figure 3.16 presents a histogram of the reduced-order model global error. Here, it can be seen that despite the lower mean deviation compared with the naive benchmark, the battery temperature at the end of a charging event may significantly deviate. Although the global error of the reduced-order model on average outperforms the naive benchmark, the significant local error mostly precludes application of this model in a charging assistant. Particularly, assuming constant parameters  $c_h$  and  $\alpha$  to represent heat flows into and out of the battery, see (3.16), seems to misrepresent the battery’s real thermal behavior.

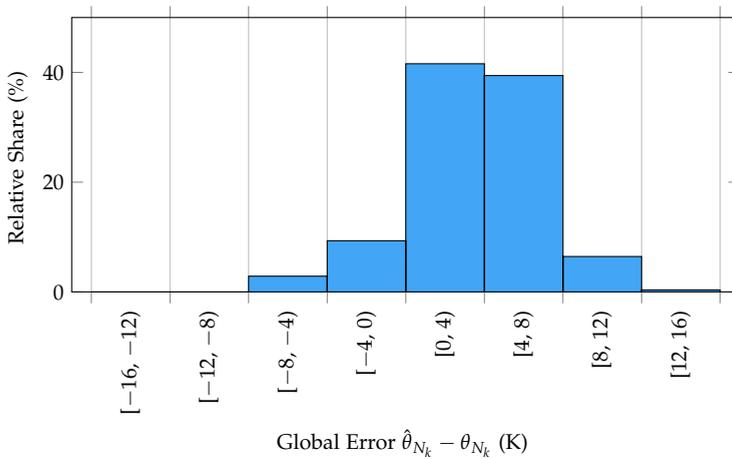


**Figure 3.16:** Deviation  $\hat{\theta}_{N_k} - \theta_{N_k}$  of true and estimated battery temperature  $\theta$  at the end of all charging events  $k \in \mathcal{K}$  (global error) using the reduced-order thermal battery model (Section 3.2.2).

As this result emphasizes the need for more advanced modeling, different data-driven approaches are presented in Section 3.2.2.2. First, a LR model is tested to estimate  $\Delta\hat{\Theta}_{\text{DD,LR}}$ . With a local error RMSE of 0.76 K, the LR yields an accuracy per time interval that lies in between the reduced-order model and the naive benchmark, see Table 3.9. Figure 3.17 presents the according scatter plot to visualize the LR model's local error. Especially for high values  $\Delta\theta$ , the LR model underestimates the true change in battery temperature per time interval. For entire charging events, however, the battery temperature is mostly overestimated by the LR model. This can be seen when contemplating the histogram of temperature deviation at the end of all charging events, see Figure 3.18. Considering the significant global error MAE of 4.18 K, the LR model seems to inadequately represent the battery's thermal behavior. The reason could be hidden electro-chemical processes that introduce non-linearity to the battery's thermal behavior. By design, the LR model is unable to capture this non-linearity.

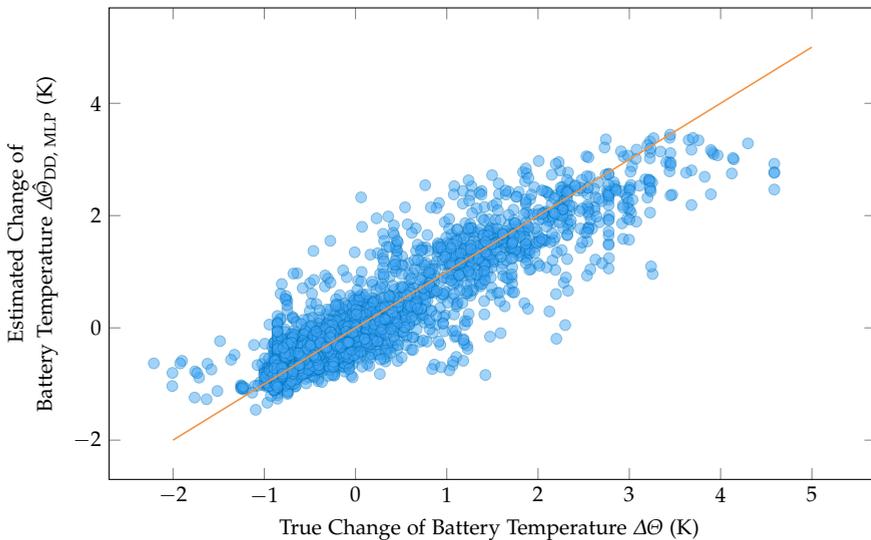


**Figure 3.17:** True and estimated change in battery temperature for individual samples  $b \in \mathcal{B}_{\text{test}}$  (local error) using the LR thermal battery model (Section 3.2.2.2), the orange line indicates ideal model behavior.



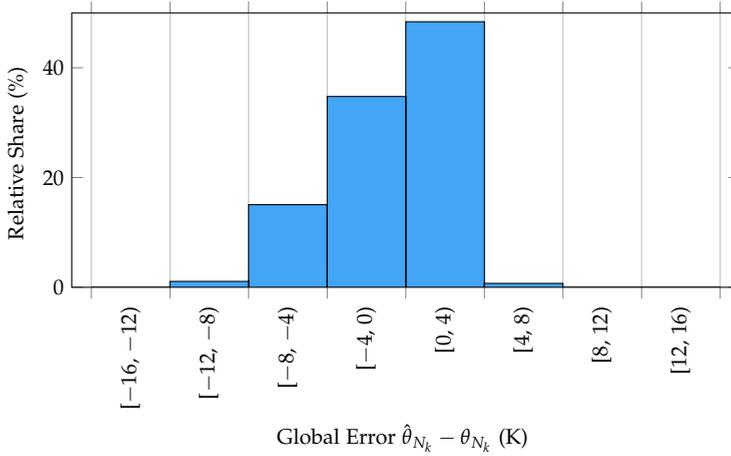
**Figure 3.18:** Deviation  $\hat{\theta}_{N_k} - \theta_{N_k}$  of true and estimated battery temperature  $\theta$  at the end of all charging events  $k \in \mathcal{K}_{\text{test}}$  (global error) using the LR thermal battery model (Section 3.2.2.2).

To overcome this, the design of more advanced ANN models is also described in Section 3.2.2.2. Although having tested several ANN hyperparameters, see Table 3.7, for the sake of brevity only the best-performing model is presented here. Particularly, an MLP is used that possesses two hidden layers with ten neurons each; each neuron is activated by a sigmoid function. The output layer consists of a single neuron activated by a linear function. Using the distinct test data set  $\mathcal{B}_{\text{test}}$ , the ANN thermal model outperforms all other models in local error evaluations, see Table 3.9. Collectively, the single time interval estimations of the ANN model yield an RMSE of 0.29 K. Furthermore, the ANN model neither systematically overestimates nor underestimates the change in battery temperature. To support this, see the scatter plot of true and estimated change of battery temperature in Figure 3.19. The mean temperature deviation at the end of a charging event, i.e., the global error



**Figure 3.19:** True and estimated change in battery temperature for individual samples  $b \in \mathcal{B}_{\text{test}}$  (local error) using the ANN thermal battery model (2 hidden layers, 10 neurons each, Section 3.2.2.2), the orange line indicates ideal model behavior.

MAE is 1.96 K. Similar to the local error, the ANN model's global error neither systematically over nor underestimates the propagated battery temperature. This can be seen when contemplating the global error histogram in Figure 3.20.

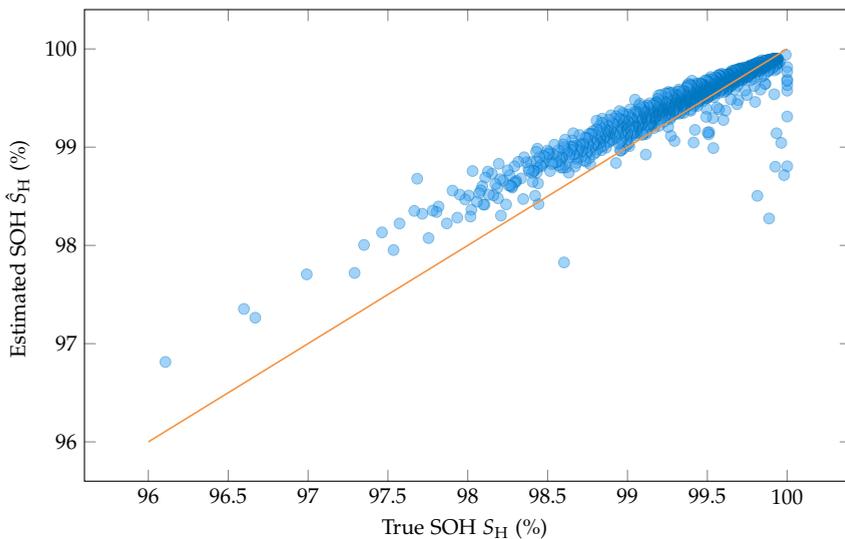


**Figure 3.20:** Deviation  $\hat{\theta}_{N_k} - \theta_{N_k}$  of true and estimated battery temperature  $\theta$  at the end of all charging events  $k \in \mathcal{K}_{\text{test}}$  (global error) using the ANN thermal battery model (2 hidden layers, 10 neurons each, Section 3.2.2.2).

Although the ANN model's performance outperforms all other models, the interpretation of this result requires some caution. ANN models are prone to overfitting, i.e., their ability to extrapolate on unseen data is often poor. For these cases, the LR model could deliver more robust results when considering the local error. It is thus suggested to apply a novelty detection to decide if the more accurate ANN model, or the more robust LR model should be used. To this end, a one-class support vector machine [269] with a radial basis function kernel could be fit onto the training data  $\mathcal{B}_{\text{train}}$ . Thus, all input data tuples that are similar to the training data contained within  $\mathcal{B}_{\text{train}}$  will be considered as *inliers*, i.e., trusted data; estimations for this data will be performed by the ANN model. All other input data tuples will be considered as *outliers*, i.e., unseen data, and will be handled by the LR model. Note that the transition between inliers and outliers could also be designed in a continuous manner (*fuzzy transition*), e.g., as in [240].

### 3.2.4.3 Validation of Aging Model

The battery aging model as described in Section 3.2.3 separately estimates the battery's capacity loss caused by cyclic and calendar aging. Aboard the EV, however, the true SOH  $S_H$  can only be measured as aggregated value, which in turn may be subject to measurement errors; for the present use case, however, this error is assumed to be negligible. To ensure a consistent validation, accordingly also aggregated aging estimations  $\hat{S}_H$  are used; these are calculated by repeatedly applying the aging model according to (3.21) throughout the lifetime of each EV. Figure 3.21 shows a scatter plot of estimated and true SOH values from real operated EVs.



**Figure 3.21:** True and estimated battery SOH values using the battery aging model as describes in Section 3.2.3, the orange line indicates ideal model behavior.

Estimating the aggregated SOH yields an RMSE of 0.159 % SOH compared with measured SOH values. For the application of the proposed charging assistant, the aging model is mainly used within optimization models to inhibit battery aging effects while charging. Regardless of the aging model's actual accuracy, battery aging can be reduced. Thus, the RMSE of 0.159 % SOH seems acceptable for the scope of this thesis.

On the upper end of true SOH, i.e., close to  $S_H = 100\%$  SOH, some outliers can be detected. These occur for new EVs that have batteries with slightly higher initial capacity than the assumed nominal storage capacity  $e_{\text{nom}}$ . For these cases, the BMS

may adapt  $e_{\text{nom}}$  to avoid measured values of  $S_{\text{H}}$  exceeding 100% SOH. The offline calculation of  $\hat{S}_{\text{H}}$  based on the aging model, however, misses the information of adapting  $e_{\text{nom}}$ . Thus, the estimated values  $\hat{S}_{\text{H}}$  can deviate from the true values  $S_{\text{H}}$ . Prospectively, the initial storage capacity would have to be considered directly in the offline aging calculation, to avoid these inaccuracies.

Furthermore, it can be seen that with advancing battery aging, i.e., lower  $S_{\text{H}}$ , the aging model estimates too high values for  $\hat{S}_{\text{H}}$ . Hence, the aging model tends to slightly underestimate battery degradation with progressing real degradation. This appears to be a systematic model deficiency that in order to be resolved would require remodeling, once according data is sufficiently available.

Note that all batteries from which the validation data is obtained, lost less than approx. 4% of their nominal storage capacity  $e_{\text{nom}}$ . Accordingly, none of the data samples shown in [Figure 3.21](#) occurs below true  $S_{\text{H}}$  of approx. 96% SOH. This result seems reasonable, as battery degradation processes occur slowly—on a time scale of years—and the underlying EVs only exist for a period of approx. two years. At the moment, the battery aging model can thus only be validated for the initial phase of degradation. In future work, progressing degradation, i.e.,  $S_{\text{H}}$  values of less than 96% SOH, may also be validated and remodeled if necessary. For this, however, additional data from further degraded EV batteries is required.

### 3.3 Conclusion

Chapter 3 deals with characterizing driving and charging behavior of EVs. For this purpose, separate models represent the EV's energy consumption while driving and the EV battery's state progression while charging.

To the end of consumption modeling (Section 3.1), a physical model is used as baseline estimator. To include user-specific driving style, data-driven consumption models are designed with additional features. For validation, an independent test set is taken from measurement data of real EV trips. Due to low accuracy of an LR model, non-linearity is assumed. This can be confirmed, as ANN and random forest models are able to properly represent user-specific driving style and outperform the baseline model. Hence, it is concluded that to estimate an EV's energy consumption more precisely, user's driving style needs to be considered. In turn, according user-related data needs to be available. If this data is not or only sparsely available, physical baseline model and user-specific data-driven model may be dynamically weighted and combined. In this way, the high accuracy of the data-driven model and the robustness of the baseline model can be leveraged.

For employing user-individual consumption models—such as the presented data-driven model—characteristic driving styles of EV users are required. To efficiently provide such information for future trips, in turn, calls for a user-individual map containing aggregated data of historical trips; such a map, however, is outside the scope of this thesis and is therefore subject to future work. In following research, the presented consumption model should also consider effects of battery aging. Furthermore, cross-validation over several independent test sets and/or vehicles should be proceeded to benchmark generalization of the developed models.

To the end of battery modeling, a three-part model is used to estimate the battery's SOC, temperature, and SOH. First, to estimate the SOC, an ECM is used as electrical model. For validation, discretized time series data measured from real EV charging events is used. Then, a local error metric—per time step—and a global error metric for entire charging events is calculated. The ECM appears to be suitable for EV charging applications, i.e., for low-dynamic battery operation. For other, possibly high-dynamic applications such as driving, a more complex model is likely to be required.

Second, a state-of-the-art reduced-order thermal model is tested and seems to inadequately represent the battery's thermal behavior. Thus, data-driven models are engineered and validated similar to the electrical model. Using additional input features, an ANN model can outperform the reduced-order model. Sparse training

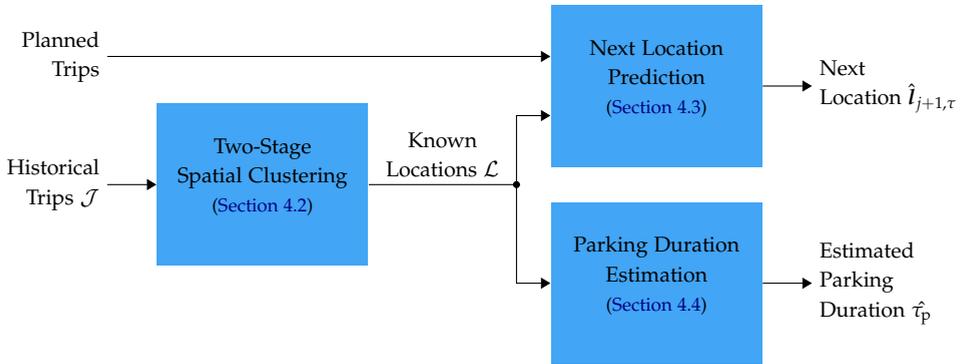
data, however, may result in over-fitting the ANN model. To overcome this problem in future work, a two-model application may be helpful in which novelty detection decides whether to use a more accurate ANN model (for trusted input data) or a more robust LR model (for unseen input data). Furthermore, the input features of the data-driven thermal models seem to lack further influences on the battery's thermal behavior. In future work, the input features could thus be enhanced with information such as internal cell temperatures or ambient conditions, e.g., sun radiation and wind. Typical EV on-board data, however, does not (yet) provide this information, and these quantities would need to be forecast for the time window of each charging event.

Third, a semi-empirical aging model from battery cell tests is used to estimate the battery's SOH. For validation, SOH measurements are used, obtained from real operated EVs that lost less than approx. 4% of their nominal battery storage capacity. For this initial phase of degradation, the aging model yields sufficiently precise results. In future work, progressing degradation, should also be considered in modeling and validation, once data from further degraded EV batteries is available. Furthermore, production variance of battery cells may invalidate the aging model results, as fast degrading cells can determine the SOH of the entire EV battery. Therefore, prospectively an implicit representation, e.g., via machine learning approaches is also suggested.

In summary, data-driven modeling is especially required for thermal and aging models of the EV battery; for the EV battery's electrical behavior while charging, a less complex model—such as the presented ECM—suffices. The fact that all battery model components are either battery- or battery-type-specific enables a generic usage in customer EVs, decoupled from user-individual behavior. Although the data to train and validate the battery model is obtained from real-world operated EVs, exceptional situations may still be underrepresented; these could comprise very high or low ambient temperatures, and bidirectional charging events, e.g., required for V2G services (see Section 1.1.1.2). A corresponding investigation of the EV models' behavior in exceptional situations is thus strongly recommended for future work.

## 4 Mobility Prediction

To ensure sufficient user acceptance of an automatically generated charging strategy, the EV user’s individual mobility requirements need to be considered. Hence, the **mobility prediction** module shown in Figure 4.1 uses an ensemble of different machine learning techniques to pursue the following three targets.



**Figure 4.1:** Schematic block diagram of the mobility prediction module and its components; inputs of the EV user, e.g., via a calendar interface, may provide information on planned trips.

First, a two-stage spatial clustering algorithm is applied to the end locations of the EV user’s historical trips. In this way, locations the user frequently visits, i.e., *known locations*, can be determined. Each of these known locations is characterized by a geodetic center location surrounded by a radius of stay. Subsequently to clustering, each known location may be enhanced with additional context information. The detailed approach can be found in Section 4.2 and is partly based on SCHWENK et al. [270]. Second, the sequence of known locations together with their context information is used to train a random forest classification model. When departing from a known location, this model predicts the next known location that the user may visit and the corresponding probability. In addition, planned trips—e.g., extracted from the user’s calendar as in SCHWENK et al. [271]—can be used to

enhance the prediction accuracy (see also [272]). All belonging methods and evaluations are described in Section 4.3. Third, different ensembles of regression models are generated from the sequence of known locations and their context information. Upon arrival at a known location, these models estimate the typical idle time (i.e., parking duration) of the EV. This approach is partly based on SCHWENK et al. [270]; all details can be found in Section 4.4. Note that all three modeling components require the collection of user data over an extended period of time before a valid output can be generated. A suitable initialization—e.g., user-independent models that can be transfer-learned—is therefore essential to ensure seamless functionality of the charging assistant immediately after activation by the EV user (*cold start* setup, see also Section 2.2.2). Before introducing the detailed algorithms and presenting the results, however, Section 4.1 provides a description of the available data used for model training and evaluation.

## 4.1 Available Data of Users' Mobility Behavior

The available data on individual mobility behavior is generally limited due to data privacy; particularly, users must explicitly consent to the storage and purposeful use of their personal data, see also [273]. In the following, two data sets are described that are used for model training and evaluation in the subsequent Sections 4.2, 4.3, and 4.4.

### 4.1.1 Data of Real Mobility Behavior

The first data set of historical trips  $\mathcal{J}_{\text{real}}$  is recorded from a personal vehicle over the two years 2018 and 2019. During these two years, the vehicle was only used by a single person and was the prioritized means of transport during this time.<sup>1</sup> The vehicle was equipped with an on-board computer to record trip data and parking duration. This data was then transferred via an Internet-of-Things system using the MQTT protocol to a backend database for storage; for further details see also Section 2.3. The final data set in the database consists of  $|\mathcal{J}_{\text{real}}| = 2906$  trips. Each trip  $j \in \mathcal{J}_{\text{real}}$  starts at departure time  $t_{j,0}$  at the start location  $l_{j,0} = (l_{\text{lat},j,0}, l_{\text{lon},j,0})$ . Accordingly, each trip ends at arrival time  $t_{j,\tau}$  at the parking location  $l_{j,\tau} = (l_{\text{lat},j,\tau}, l_{\text{lon},j,\tau})$ . Furthermore, all trips in  $\mathcal{J}_{\text{real}}$  are ordered in time, i.e.,  $l_{j,\tau} = l_{j+1,0}$ . Note that this data set only mirrors the behavior of one individual user and is therefore not necessarily representative of other EV users. In the remainder of this chapter, this data set will be referred to as *real data* or  $\mathcal{J}_{\text{real}}$ , respectively.

### 4.1.2 Generation of Semi-Synthetic Data

Besides the real data recorded from a single user, an additional semi-synthetic data set  $\mathcal{J}_{\text{syn}}$  is generated. This data set aims to represent typical and predictable electric vehicle user behavior with a defined amount of random and unpredictable events. As real user behavior is likely to be less predictable, the data set functions as a qualitative benchmark data set for the predictive models in Section 4.3 and Section 4.4.

Generating the semi-synthetic data set follows the procedure given in Algorithm 4.1 (see page 89) and is briefly described here. For the sake of comparability, the same

---

<sup>1</sup>Note that the vehicle user was formally notified and agreed to the recording and use of their data for scientific purposes.

time frame is used as for the real data, i.e., the years 2018 and 2019. Starting from a *home* location, for each day within this time frame trips are taken in between the eight locations

$$\mathcal{L}_{\text{syn}} = \left\{ l_{\text{home}}, l_{\text{work}}, l_{\text{orchestra}}, l_{\text{gym}}, l_{\text{parents}}, l_{\text{grocery}}, l_{\text{pool}}, l_{\text{backoffice}} \right\}, \quad (4.1)$$

which represent the most commonly visited places for a real EV user. To account for spatial variations of the parking position, each location is offset by a normally distributed variable  $\iota \sim \mathcal{N}(0, 0.0005)$  when generating the sequence of trips.<sup>2</sup> In [Algorithm 4.1](#), each new trip that is added to  $\mathcal{J}_{\text{syn}}$  is denoted by a blue line following the scheme:

[departure time] $\pm$ [standard deviation]: [start location]  $\rightarrow$  [end location] ([mean travel time])

The trip sequences include recurrences on four levels: daily, weekly, monthly, and seasonally. Besides the eight known locations in  $\mathcal{L}_{\text{syn}}$ , no further, unknown locations are visited. All departure times are assumed to be normally distributed; the according standard deviation is directly indicated with each departure time. An external routing service provides mean values for all travel times, neglecting variations due to different start times of a trip; to account for stochastic fluctuations in travel times, they are multiplied with a normally distributed random factor  $\zeta \sim \mathcal{N}(1, 0.05)$ .<sup>3</sup> Twice a week, a grocery store is occasionally visited in between *work* and *home* location, see [Algorithm 4.1](#), lines 19-21 and lines 38-40. The probability that each of these visits occurs is defined by the BERNOULLI-distributed random variable  $\xi \sim \mathcal{B}(0.5)$ . In addition, all trips made Monday through Friday occur only on regular workdays, not holidays, see [Algorithm 4.1](#), line 2. The required list of public and school holidays “holidayList” is obtained from a holiday calendar, or optionally also from the EV user’s personal calendar, if available. In this way, the semi-synthetic data set  $\mathcal{J}_{\text{syn}}$  is obtained, which in the remainder of this chapter will be referred to as *synthetic data*.<sup>4</sup>

---

<sup>2</sup>At the equator, the standard deviation 0.0005 corresponds to a radius of approx. 55m around the target location.

<sup>3</sup>Overlaps caused by the normally distributed times (e.g., departure before arrival time) are handled in the data cleaning process of the subsequent models, see [Section 4.3](#) and [Section 4.4](#).

<sup>4</sup>To support future research on predicting human mobility, the data set  $\mathcal{L}_{\text{syn}}$  and source code of [Algorithm 4.1](#) is online available at: <https://github.com/KarlSchwenk/mobility-data-creator>

**Algorithm 4.1:** Procedure to generate semi-synthetic mobility data set  $\mathcal{J}_{\text{syn}}$  based on set of known locations  $\mathcal{L}_{\text{syn}}$  and real mobility behavior of a single EV user (note the time format *hours:minutes* (local time zone) and the date format *day.month.year*).

---

**Input:**  $\mathcal{L}_{\text{syn}}$ , holidayList

```

1: for each day  $\in \{1.1.2018, 2.1.2018, \dots, 30.12.2019, 31.12.2019\}$  :
    | # check if current day is public/school holiday
2:   if day  $\notin$  holidayList :
3:     | if day is Monday :
4:       | if day is first Monday of month :
5:         | # monthly visit of backoffice
6:         | 08:00  $\pm$  5 min:  $I_{\text{home}} \rightarrow I_{\text{backoffice}}$  (70 min)
7:         | 15:45  $\pm$  45 min:  $I_{\text{backoffice}} \rightarrow I_{\text{home}}$  (85 min)
8:       | else:
9:         | 06:00  $\pm$  5 min:  $I_{\text{home}} \rightarrow I_{\text{work}}$  (22 min)
10:        | 16:30  $\pm$  45 min:  $I_{\text{work}} \rightarrow I_{\text{home}}$  (25 min)
11:     | if day is Tuesday :
12:       | 06:00  $\pm$  5 min:  $I_{\text{home}} \rightarrow I_{\text{work}}$  (22 min)
13:       | if day  $\in \{1.5.2018, \dots, 30.9.2018, 1.5.2019, \dots, 30.9.2019\}$  :
14:         | # seasonal visit of public swimming pool, only in summer months
15:         | 14:45  $\pm$  45 min:  $I_{\text{work}} \rightarrow I_{\text{pool}}$  (12 min)
16:         | 16:20  $\pm$  20 min:  $I_{\text{pool}} \rightarrow I_{\text{home}}$  (23 min)
17:       | else:
18:         | 16:30  $\pm$  45 min:  $I_{\text{work}} \rightarrow I_{\text{home}}$  (25 min)
19:     | if day is Wednesday :
20:       | 06:00  $\pm$  5 min:  $I_{\text{home}} \rightarrow I_{\text{work}}$  (22 min)
21:       | if  $1 == \zeta \sim \mathcal{B}(0.5)$  :
22:         | # random visit of grocery store with 50% probability
23:         | 16:30  $\pm$  45 min:  $I_{\text{work}} \rightarrow I_{\text{grocery}}$  (20 min)
24:         | 17:15  $\pm$  5 min:  $I_{\text{grocery}} \rightarrow I_{\text{home}}$  (6 min)
25:       | else:
26:         | 16:30  $\pm$  45 min:  $I_{\text{work}} \rightarrow I_{\text{home}}$  (25 min)
27:         | 18:24  $\pm$  30 min:  $I_{\text{home}} \rightarrow I_{\text{gym}}$  (35 min)
28:         | 20:49  $\pm$  10 min:  $I_{\text{gym}} \rightarrow I_{\text{home}}$  (32 min)

```

---

Continuation of Algorithm 4.1.

---

```

# continued for-loop from Algorithm 4.1, line 1
|
| # continued if-statement from Algorithm 4.1, line 4
26: |   if day is Thursday :
27: |       06:15 ± 5 min:  $I_{\text{home}} \rightarrow I_{\text{work}}$  (22 min)
28: |       if day ∈ {1.5.2018, ..., 30.9.2018, 1.5.2019, ..., 30.9.2019} :
|           # seasonal visit of public swimming pool, only in summer months
29: |           15:20 ± 45 min:  $I_{\text{work}} \rightarrow I_{\text{pool}}$  (12 min)
30: |           17:00 ± 20 min:  $I_{\text{pool}} \rightarrow I_{\text{home}}$  (23 min)
31: |       else:
32: |           16:30 ± 45 min:  $I_{\text{work}} \rightarrow I_{\text{home}}$  (25 min)
33: |   if day is Friday :
34: |       06:00 ± 5 min:  $I_{\text{home}} \rightarrow I_{\text{work}}$  (22 min)
35: |       if 1 ==  $\zeta \sim \mathcal{B}(0.5)$  :
|           # random visit of grocery store with 50% probability
36: |           15:30 ± 45 min:  $I_{\text{work}} \rightarrow I_{\text{grocery}}$  (20 min)
37: |           16:15 ± 5 min:  $I_{\text{grocery}} \rightarrow I_{\text{home}}$  (6 min)
38: |       else:
39: |           15:30 ± 45 min:  $I_{\text{work}} \rightarrow I_{\text{home}}$  (25 min)
40: |           19:27 ± 30 min:  $I_{\text{home}} \rightarrow I_{\text{orchestra}}$  (19 min)
41: |           22:15 ± 5 min:  $I_{\text{orchestra}} \rightarrow I_{\text{home}}$  (18 min)
42: |   if day is Saturday :
43: |       10:00 ± 45 min:  $I_{\text{home}} \rightarrow I_{\text{grocery}}$  (6 min)
44: |       11:05 ± 5 min:  $I_{\text{grocery}} \rightarrow I_{\text{home}}$  (6 min)
45: |   if day is Sunday :
46: |       11:55 ± 45 min:  $I_{\text{home}} \rightarrow I_{\text{parents}}$  (24 min)
47: |       15:00 ± 60 min:  $I_{\text{parents}} \rightarrow I_{\text{home}}$  (25 min)

```

**Output:**  $\mathcal{J}_{\text{syn}}$

---

## 4.2 Determination of Users' Known Locations

Generally, the habits of human behavior—in this case **EV** users—are assumed to be coupled with a limited number of specific locations; for instance, the most common locations comprise the user's home address, work place and favorite grocery store. When targeting to generate an **EV** charging strategy that is adapted to these habits, the knowledge of **EV** users' frequently visited locations is crucial. In the following these frequently visited locations will also be referred to as *known locations*  $\mathcal{L}$ .

In the process of determining a user's known locations based on parking locations recorded aboard an **EV** (as in [Section 2.3](#)), two challenges may occur. First, the tolerance inherent in measurements of a **GPS** location may deviate up to approx. 30 m from the actual parking location.<sup>5</sup> Second, the **EV** user may not necessarily park their **EV** at exactly the same **GPS** location, although visiting only a single location.

To overcome these challenges and estimate the true locations an **EV** user frequently visits, a two-stage spatial clustering approach is presented in this section. For this purpose, the real data  $\mathcal{J}_{\text{real}}$  as described in [Section 4.1.1](#) is used. After initially cleaning the data set, a state-of-the-art density-based spatial clustering algorithm is run in the first stage. In the second stage, a joining algorithm combines clusters that are assumed to represent a single known location of the **EV** user. Finally, the proposed scheme is briefly evaluated to demonstrate its functionality.

### 4.2.1 Data Cleaning

To ensure a proper data quality, the raw data is initially processed as follows. First, all trips possessing measurement errors are removed; these may comprise invalid **GPS** values for their parking locations and/or corrupt timestamps. Then, all trips shorter than 15 s are removed, as these are assumed to be unrealistic trip times and thus also measurement errors. After removing all invalid trips, the parking duration  $\tau_p$  is calculated, i.e., the time after each trip the **EV** is stationary before the next trip begins. Subsequently, the data is filtered to only include trips possessing a parking duration between 2 h and 24 h. This time frame appears to be particularly relevant for smart charging applications such as **charging scheduling** and **charging optimization**. For a parking time of less than 2 h it is either unreasonable to plug in

---

<sup>5</sup>For parking locations without **GPS** connection, the last transferred location is used.

the EV at all, or, if plugged in, the user would likely require the EV to be charged at maximum power; hence, there would not be enough flexibility to enable **charging optimization**. On the other hand, if the parking duration exceeds 24 h there is too much flexibility. As a consequence, **charging optimization** either becomes trivial, or the uncertainty inherent in **charging scheduling** exceeds a trustworthy level. More details on the latter case can be found in the according chapters on **charging scheduling** (Chapter 5) and **charging optimization** (Chapter 6).

Note that performing the filtering task in a real-world application calls for an additional model predicting whether the parking duration will be within 2 h to 24 h or not. In preliminary tests using the real data set, a binary random forest classifier achieved a precision of 93.1 % and a recall of 85.5 %.

## 4.2.2 Two-Stage Density-Based Spatial Clustering

Given clean data, the next aspect in determining an EV user’s known locations is a two-stage clustering algorithm. In the first stage, a standard *DBSCAN*-algorithm [274] is applied to the parking locations  $l_{j,\tau} = (l_{\text{lat},j,\tau}, l_{\text{lon},j,\tau})$  of all trips  $j \in \mathcal{J}_{\text{real}}$ .<sup>6</sup> *DBSCAN* is a state-of-the-art density-based spatial clustering algorithm that can be parameterized with little prior knowledge of the data. To level out measurement scatter of parking locations in the first clustering stage, the density parameter (i.e., the maximum spatial distance of two data points within a cluster) is adjusted to GPS measurement tolerance. In addition, specifying a minimum number of data points per cluster allows to designate irregularly visited locations as “noise”. Furthermore, *DBSCAN* is able to detect clusters of arbitrary shapes; a property helpful to handle various parking setups, e.g., both along a street (elongated cluster shape) and in a rectangular parking lot (bulky cluster shape). For more insight into the functionality and properties of *DBSCAN*, see [274].

Although this initial clustering stage generates several suitable clusters, it may also create multiple clusters representing only a single known location. For this reason, the second-stage algorithm joins clusters whose center locations are closer to each other than a predefined threshold. This threshold represents a typical walking distance a user may take. Particularly, users would either walk that distance—without using the EV—or, if they choose to drive, the energy consumption of such short trips is negligible. The second stage algorithm furthermore assigns GPS points primarily labeled as “noise” to existing clusters, if these locations are within a specific

---

<sup>6</sup>*DBSCAN* is short for “Density-Based Spatial Clustering for Applications with Noise”, see also [274].

neighborhood radius. In this way, irregular parking positions at a known location can still be detected by the algorithm. Hence, the size of the neighborhood radius should be chosen to a reasonable distance for users to walk when parking at a known location.

As a final step, the set of inferred known locations  $\mathcal{L}$  is ordered and each location  $l \in \mathcal{L}$  is labeled according to the frequency of its occurrence. The location that contains the most data points is labeled as  $l_1$  (or "Location 1"), that with the second-most data points  $l_2$  (or "Location 2"), and so on. Furthermore, the data points considered as noise will be referred to as "-1". Once determined by the two-stage algorithm, each known location may also be augmented with information, e.g., on charging infrastructure, surrounding amenities and user preferences, see also [275]. In this way, the performance of subsequent predictive models (see Section 4.3 and Section 4.4) can be improved.

### 4.2.3 Evaluation of Clustering

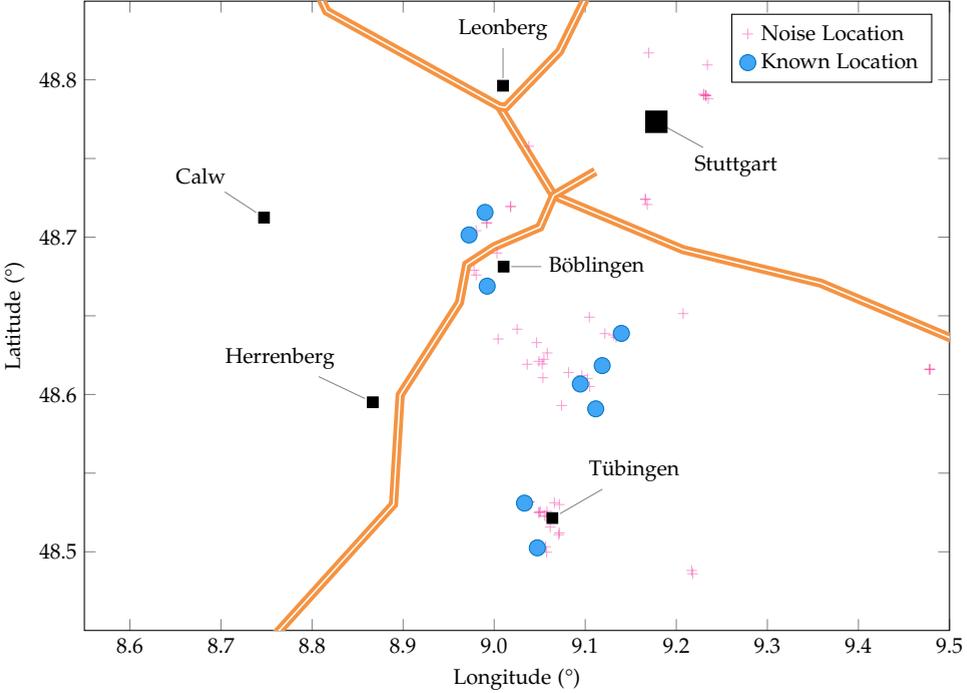
This section aims to demonstrate how the two-stage clustering approach presented in Section 4.2.2 is able to enhance a default *DBSCAN* algorithm. To achieve this, all parts of the algorithm are implemented in *Python* [256] using *Scikit-Learn* [254]. Then, the synthetic data set  $\mathcal{J}_{\text{syn}}$  is used to tune the parameters of the two-stage clustering algorithm (see Table 4.1). In particular, the parameters are manually

**Table 4.1:** Parameters and selected values for two-stage density-based spatial clustering algorithm (see Section 4.2.2) to determine EV users' known locations.

Parameter	Selected Value
Density Parameter of Clusters	100 m
Minimum Number of Data Points per Cluster	5
Maximum Cluster Distance for Joining Clusters	500 m
Neighborhood Radius to Assign Noise	300 m

adjusted in a way that the inferred known locations after clustering are identical with the ones used for initially generating the data set (see Algorithm 4.1). This procedure can be legitimately applied, as no further, unknown locations besides the eight known locations in  $\mathcal{L}_{\text{syn}}$  are visited. Hence, there are no "noise" locations in  $\mathcal{J}_{\text{syn}}$ , i.e., no trips to irregularly visited locations. The eventually chosen values for each parameter are also given in Table 4.1.

For evaluation, both available data sets  $\mathcal{J}_{\text{syn}}$  and  $\mathcal{J}_{\text{real}}$  (see Section 4.1) are evaluated individually. For the real data  $\mathcal{J}_{\text{real}}$ , Figure 4.2 shows a map section of southwestern Germany with the inferred known locations  $\mathcal{L}_{\text{real}}$  drawn as blue circles.<sup>7</sup> Here, the proposed two-stage clustering approach is used. It can be seen



**Figure 4.2:** Map section of southwestern Germany with inferred known locations  $\mathcal{L}_{\text{real}}$  (blue circles) and noise points (pink crosses) using the proposed two-stage clustering approach, based on real data  $\mathcal{J}_{\text{real}}$ , orange double lines indicate major highways.

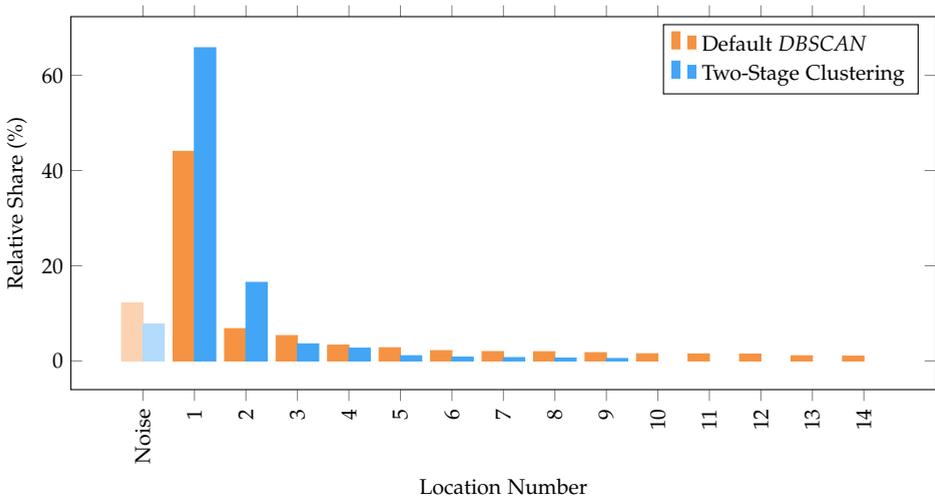
that all nine determined locations are separated locally and have travel distances of less than approx. 50 km in between each other. Furthermore, irregularly visited locations, i.e., “noise” points are drawn as pink crosses.<sup>8</sup> Although many “noise” points are determined, either the density of these points or the minimum number of data points did not suffice to form another cluster. However, if the user would continue to visit one of these “noise” locations, an additional cluster, i.e., known location, could be formed eventually. To always obtain a clustering result that is representative of the EV user’s momentary mobility behavior, a frequent recalcu-

<sup>7</sup>To ensure proper data privacy for the real data set, location numbering is omitted in Figure 4.2.

<sup>8</sup>Further “noise” points outside the scope of Figure 4.2 exist, which are not shown here.

lation (e.g., on a weekly basis) may be required. Alternatively, a stream clustering approach that uses a rolling time window of historical data (e.g., 90 days) as described in [242] could be implemented.

To also perform a quantitative comparison of default *DBSCAN* with the proposed two-stage clustering approach, in Figure 4.3 a bar chart is shown. For each known location that was determined from the real data set  $\mathcal{J}_{\text{real}}$ , vertical bars indicate the relative relevance, i.e., share of data points per cluster; orange bars are used for default *DBSCAN*, blue bars for two-stage clustering. Furthermore, the share of data points labeled as “noise” is shown. Default *DBSCAN* finds 42 locations from



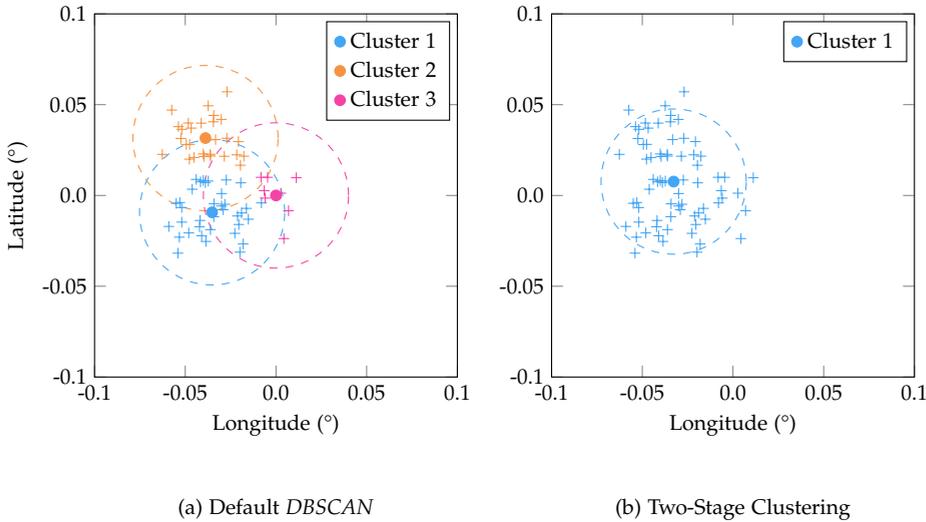
**Figure 4.3:** Inferred known locations  $\mathcal{L}_{\text{real}}$  and their relative relevance to compare default *DBSCAN* with the proposed two-stage clustering approach (see Section 4.2.2) on the real data set  $\mathcal{J}_{\text{real}}$ ; in addition, the share of data points labeled as “Noise” is shown.

which only 14 (as shown in Figure 4.3, orange bars) possess more than 1% of the data points. Hence, especially the locations 15 to 42 are locations the user does not visit on a regular basis. Generating trustworthy predictions on parking duration or next locations to visit becomes accordingly difficult for these locations. Another measure of quality is the share of data points assigned to “noise”, i.e., the share of irregularly visited locations. Default *DBSCAN* labels 12.21% of all data points as “noise”.

However, when using the two-stage clustering approach, i.e., comprising cleaning the data, joining close clusters and assigning “noise” within a neighborhood radius,

only 9 locations are found, see blue bars in Figure 4.3. Hence, the relevance of the locations has been redistributed. Data points, which the default *DBSCAN* algorithm assigned to locations with low relevance, i.e., location 10 to 42, were reassigned to one of the locations 1 to 9. For instance, the relevance of location 1 increases from 44.03 % to 65.81 %, accordingly. Furthermore, the share of data points labeled as “noise” is reduced from 12.21 % to 7.75 %.

The effectiveness of the second stage of the spatial clustering method can also be seen considering an example of the synthetic data set  $\mathcal{J}_{\text{syn}}$ . Figure 4.4 therefore visually compares the results of the default *DBSCAN* (left) and the two-stage clustering (right). All raw data points are depicted as crosses, the inferred clusters



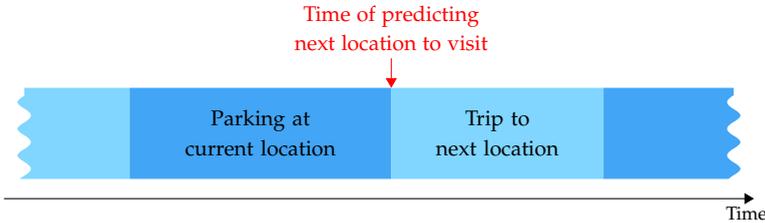
**Figure 4.4:** Comparison of the default *DBSCAN* (a) and the two-stage spatial clustering approach (including data cleaning) (b). The two-stage approach combines clusters that are not clearly separated to ensure well-defined single locations for further assessment.

are shown as colored markers with a surrounding neighborhood radius. Noticeably, the default *DBSCAN* method generates three distinct clusters (Figure 4.4, left). However, these three clusters are very close to each other, and there is no clear separation in between. On the contrary, when applying the two-stage clustering method, only one cluster is generated (Figure 4.4, right), despite the underlying raw data being identical. This single cluster, formed by joining clusters whose center locations are less than 500 m apart, represents a clearly defined single location. Hence, it is far more advantageous for the performance of subsequent predictive models as described in Section 4.3 and Section 4.4. Furthermore, a two-stage ap-

proach allows to flexibly adjust parameters depending on the EV user’s individual mobility setup. For example, a user living in a densely populated city requires a different neighborhood distance than one living in a sparsely populated country region. This accounts for a major advantage of the two-stage clustering approach.

### 4.3 Prediction of Next Location

Adapting an EV’s charging process in an automated manner necessarily calls for knowledge of the EV’s future energy demand. To this end, the next location to visit  $\hat{l}_{j+1,\tau}$  needs to be predicted given a specific context, e.g., the time, and the current and last locations. For this, the point in time of departing from a known location (see Section 4.2) is taken as reference, see Figure 4.5.<sup>9</sup> Based on the



**Figure 4.5:** Schematic sequence of trips (light blue) and parking periods (dark blue) to indicate the point in time (red arrow) of predicting the next location to visit.

next trip’s destination, an external routing service and a consumption model (e.g., as in Section 3.1) can be used to derive the future energy consumption of the EV. Furthermore, probabilistic predictions are applied to obtain all possible destinations of the next trip and their corresponding probability  $\mathbb{P}_t$ . This information allows to also estimate an expected energy consumption and a measure of uncertainty, e.g., the energy consumption’s standard deviation. Depending on the data quality, the distribution of the next trip’s energy consumption may also be estimated directly in a future application. In addition to  $\mathbb{P}_t$ , a residual probability  $\mathbb{P}_r$  is deduced to quantify the chance of the user visiting an unknown (“noise”) location next.

In the remainder of this section, first the engineering of features and labels based on historical trips  $\mathcal{J}$  and the previously determined known locations  $\mathcal{L}$  (see Section 4.2)

<sup>9</sup>Adapting a future charging process (as in Chapter 6) requires to predict the next trip’s destination already upon arrival at the current location. At this point in time, however, some input features (see Section 4.3.1) are unknown and therefore need to be estimated separately in future work.

is described. Then, two classification models are designed to predict the most likely destination  $\hat{l}_{j+1,\tau}$  of the next trip and estimate the transition probabilities  $P_t$ . Finally, all methods are evaluated and the results are presented.

### 4.3.1 Data Engineering for Next Location Prediction

This step combines the historical trip data  $\mathcal{J}$  with the determined known locations  $\mathcal{L}$ .<sup>10</sup> Particularly, the historical sequence of visited (known) locations is considered as time series of discrete events. First, a general preprocessing is conducted. For each trip  $j \in \mathcal{J}$  the start location  $l_{j,0} \in \mathcal{L}$  (i.e., the cluster number) is derived from the end location  $l_{j-1,\tau_p} \in \mathcal{L}$  of the previous trip  $j-1$ . For this, the previously determined known locations  $\mathcal{L}$  are used (see Section 4.2). If data gaps exist, e.g., from data cleaning as described in Section 4.2.1, the start location of the consecutive trip is assigned to “noise”. All these trips starting from an unknown location are removed for training and evaluation of the predictive models (see Section 4.3.2); the resulting models therefore only produce valid predictions for trips starting from a known location. Furthermore, *round trips*, i.e., trips with the same start and end location are excluded in both training and test data. These trips may exist due to removing very short trips (see Section 4.2.1) or joining several identified locations in the second clustering stage (see Section 4.2.2). In this step each trip may also be enriched with supplementary information such as calculated energy consumption or the type of trip, i.e., spontaneous or planned; the latter information e.g., could be acquired from the EV user’s calendar. Together with one-time user inputs such as *home* and *work* location, the accuracy of the predictive models may thus be improved.

Second, each trip’s end location  $l_{j,\tau}$  is defined as one-hot encoded training label for the classification models described in Section 4.3.2. When starting a new trip from a known location, the models are thus trained to predict the most likely location  $\hat{l}_{j+1,\tau}$  at which this trip may end.

To effectively achieve this, additional features are designed from the start time of each trip. All features are then combined in the feature vector  $x_l \in \mathbb{R}^9$ , whose components are briefly described here:

- **Current Location:** The known location from which the next trip starts, given as an integer numeric value; Unknown locations, i.e., “noise” equate to  $-1$ .

---

<sup>10</sup>All data engineering steps are separately proceeded for both data sets  $\mathcal{J}_{\text{syn}}$  and  $\mathcal{J}_{\text{real}}$ , see Section 4.1.

- **Last Location:** The known location from which the last trip started, given as an integer numeric value; Unknown locations, i.e., “noise” equate to  $-1$ .
- **Hour of Day:** The time at which the next trip starts, expressed as duration since midnight and given as a real numeric value, e.g., 09:42 is encoded as 9.7.
- **Time Window of Day:** One-hot encoded variable to indicate different time windows of a day (based on the hour of day), i.e., morning (05:00-08:59), noon (09:00-12:59), afternoon (13:00-16:59), evening (17:00-21:59), and night (22:00-4:59).
- **Day of Week:** Encoding of the day of the week (integer numeric value, 0 for Sunday to 6 for Saturday) to represent frequently reoccurring weekly trips.
- **Month:** Numeric values that indicate the month of the year (1-12 as an integer numeric value) to represent seasonally varying mobility behavior.
- **Is Holiday:** Boolean value based on a holiday calendar indicating which days are public and/or school holidays; in this way, work-free days e.g., without specific trips to work could be mapped.
- **Last Trip Duration:** Travel time of the previous trip before parking at the current location.
- **Parking Duration at Current Location:** The time the EV has been parked at the current location before the next trip started.

To quantify the performance of these features, the GINI importance  $q_{\text{Gini}}$  is used as shown in Table 4.5. This metric is considered appropriate as in Section 4.3.2 a ran-

**Table 4.2:** GINI importance  $q_{\text{Gini}}$  using a random forest model for all features when predicting the next trip destination on the real data  $\mathcal{J}_{\text{real}}$ . Higher numbers indicate that the feature causes a larger mean decrease in impurity and is therefore more relevant to the random forest [276,277].

Feature	Importance $q_{\text{Gini}}$ to Label $\hat{l}_{j+1,\tau}$
Current Location	0.2184
Last Location	0.0419
Hour of Day	0.2386
Time Window of Day	0.0475
Day of Week	0.0945
Month	0.0733
Is Holiday	0.0135
Last Trip Duration	0.1151
Parking Duration at Current Location	0.1573

dom forest classification is used as machine learning model [276,277]. A definition of  $q_{\text{Gini}}$  taken from [276–278] can be found in Section A.1.2.

It can be seen that “Hour of Day” is the most important feature to predict the label  $\hat{l}_{j+1,\tau}$ . Hence, a certain daily rhythm may exist and the EV user may have regular trips based on the time of the day. On the contrary, the “Time Window of Day” feature, which is highly correlated with the hour of the day, is not important. Apparently, the information regarding the daily rhythm is thus sufficiently covered by the “Hour of Day” feature. Furthermore, the “Current Location” is almost as important to predict  $\hat{l}_{j+1,\tau}$  as the “Hour of Day”. A future trip’s destination thus also strongly depends on the location from which it starts. This seems reasonable, as most trips are coupled to a specific purpose, which in turn is usually coupled to a specific location. Both the “Parking Duration at Current Location” and “Last Trip Duration” are also relevant for the label. This suggests repetitive patterns, i.e., the next location may be different if the car was parked for a large amount of time or if the last trip was very short. All remaining features only show minor importance for predicting  $\hat{l}_{j+1,\tau}$ . For “Day of Week”, “Month”, and “Is Holiday” this suggests that the mobility behavior of the evaluated user does not exhibit underlying weekly and seasonal patterns. The low importance of the “Last Location” indicates that the history of visited locations has only a minor impact on the next location to be visited.

### 4.3.2 Models to Predict the Next Location to Visit

For predicting the next location to visit two models are tested. First, a naive baseline model is presented, which only takes the feature “Current Location” into account; it functions as a benchmark. Second, a more advanced random forest model is designed that considers all features  $x_j$ ; the according fundamentals on random forests can be found in [260]. Both models are trained and tested using a five-fold cross-validation on both data sets described in Section 4.1, i.e., the real data set  $\mathcal{J}_{\text{real}}$  and the synthetic data set  $\mathcal{J}_{\text{syn}}$ . The according train and test sets are denominated using the index “train” or “test”, respectively.

#### 4.3.2.1 Most-Common Next Location Model

To design a naive baseline model, the most frequently visited next location is determined for each known location. For this, the training data set  $\mathcal{J}_{\text{train}}$  is filtered

to only contain trips departing from a single known location  $l$ . Thus, a filtered training data set of trips

$$\mathcal{J}_{\text{train}}^{(l)} = \{j \in \mathcal{J}_{\text{train}} | l_{j,0} = l\}, \quad (4.2)$$

is obtained for each known location  $l \in \mathcal{L}$ . When departing from a known location  $l \in \mathcal{L}$ , the next location to visit is predicted as

$$\hat{l}_{j+1,\tau} = \underset{\forall k \in \mathcal{L}}{\operatorname{argmax}} \left\{ \left| \{j \in \mathcal{J}_{\text{train}}^{(l)} | l_{j,\tau} = k\} \right| \right\}, \quad (4.3)$$

regardless of previously visited locations or time. To better understand the characteristics of this model, consider the following example. For an arbitrary EV user, the majority of historical trips starting from the *home* location possess the end location *work*. Then, the predicted next place to visit for all future trips starting from *home* would always be the *work* location. Despite the current location, the predictions of this model are thus independent from all other features in  $\mathbf{x}_j$ ; all resulting predictions are time-invariant, accordingly.

In the following, this model will be referred to as *Baseline* model. It is used to benchmark the more advanced random forest model described in Section 4.3.2.2.

#### 4.3.2.2 Random Forest Classification Model

As second model, a random forest classification model as described in [260] is used. It consists of a maximum of 100 decision trees, which are individually trained with randomly selected subsets of the training data. For the training process, the GINI importance is used as loss metric (see also [276,277]). All other hyperparameters of the model are used with their default values as given by *Scikit-Learn* [254].

Primarily, the random forest model aims to predict a one-hot encoded vector of all next locations to visit. In addition to these deterministic predictions, however, the conditional transition probabilities  $\mathbb{P}_t(l|\mathbf{x}_l), \forall l \in \mathcal{L}$  are also estimated. Given a context expressed by the feature vector  $\mathbf{x}_l$ , this probability represents the chance that the EV user may visit the location  $l$  next. Note that the summed transition probabilities for all known locations may be less than one. Accordingly, the residual probability

$$\mathbb{P}_r(\mathbf{x}_l) = 1 - \sum_{\forall l \in \mathcal{L}} \mathbb{P}_t(l|\mathbf{x}_l), \quad (4.4)$$

can be deduced as a measure of uncertainty for each context  $\mathbf{x}_l$ . The residual probability  $\mathbb{P}_r$  thus represents the chance that the EV user may visit an unknown location next.

In the following, this model will be referred to as *Random Forest* model.

### 4.3.3 Validation and Results

To evaluate the models described in Section 4.3.2, a five-fold cross-validation on both the real data set  $\mathcal{J}_{\text{real}}$  and the synthetic data set  $\mathcal{J}_{\text{syn}}$  (see Section 4.1) is performed. For this, all models and evaluation schemes are implemented in *Python* [256] using *Scikit-Learn* [254]. Only the eight most frequently visited locations are considered for both training and evaluation. All trips with unknown start and end location, and all *round trips*, i.e., trips with the same start and end location, are removed. In doing so, the real data set  $\mathcal{J}_{\text{real}}$  can directly be compared with the synthetic data set  $\mathcal{J}_{\text{syn}}$ . Furthermore, the present study focuses on predicting the next locations to visit in order to estimate the EVs future energy demand. If either the start or end location is unknown (or both are identical), this cannot be achieved. Therefore, the next location is only predicted for trips starting from and going to commonly visited, i.e., known locations.

To quantify model performance, the mean accuracy  $q_{\text{ACC}}$  (see (A.5)) is calculated. This quantity represents the portion of next locations a model predicts correctly, i.e., predicted locations that equal the true location visited next. The evaluation is also proceeded for location-specific subsets (as defined in (4.2)) of the test data, to evaluate the model for a specific scope, see Section 4.3.3.1.

Furthermore, the probabilistic output of the random forest model, i.e., the probability of each location to be visited next, is evaluated in Section 4.3.3.2. For this, a matrix is created that contains all possible transition probabilities  $\mathbb{P}_t$  based on the set of known locations  $\mathcal{L}$ . Based on the real data set, finally, a one-day simulation is used to qualitatively evaluate the transition probabilities over time.

#### 4.3.3.1 General Performance of Models to Predict the Next Location

In terms of the average accuracy  $q_{\text{ACC}}$ , the random forest classification model is generally able to outperform a naive baseline model (see Section 4.3.2). As shown in Table 4.3, the random forest model yields an average accuracy  $q_{\text{ACC}} = 65.12\%$  on the real data set  $\mathcal{J}_{\text{real}}$ ; the random forest hence correctly predicts approx. two thirds of all next locations to visit. This level of accuracy is considered sufficient for use cases such as **charging scheduling** described in Chapter 5. The baseline model, on the other hand, only achieves an (likely insufficient) average accuracy of  $q_{\text{ACC}} = 43.64\%$ .

**Table 4.3:** Comparison of baseline and random forest model (see Section 4.3.2) accuracy  $q_{ACC}$  for different scopes (start locations) when predicting the next location to visit on the real data set  $\mathcal{J}_{real}$  using a five-fold cross-validation.

Scope	Baseline Model	Random Forest
Location 1	20.59%	59.83%
Location 2	67.72%	65.27%
Location 3	81.68%	76.56%
Location 4	66.67%	41.50%
Location 5	66.02%	66.01%
Location 6	76.17%	70.88%
Location 7	51.0%	38.0%
Location 8	91.67%	88.81%
Average	43.64%	65.12%

The location-specific evaluation in Table 4.3 shows that especially for Location 1, i.e., the *home* location, the random forest model yields more precise results. For all other locations, however, the baseline model either achieves similar or even higher accuracy compared with the random forest model. This can be explained by the fact that trips starting from any location other than *home* mostly possess the end location *home*; the naive baseline model, which always predicts the most frequent trip destination per location, can easily handle those trips. Trips starting from *home*, on the contrary, go to many different locations. Here, the baseline model is only able to correctly predict the next location for about one fifth of all trips. The random forest model, however, has the ability to leverage features such as time and previously visited locations to achieve much higher accuracy ( $q_{ACC} = 59.83\%$ ) for Location 1. In the future, the predictions of both models could be combined to achieve a higher average accuracy. Alternatively, an individual random forest model could be trained for each start location. In this way, the underlying classification problem would simplify and improved accuracy can be expected. This, however, also comes with increased computational complexity for both model training and predictions.

As shown in Table 4.4, evaluating the models on the synthetic data set  $\mathcal{J}_{syn}$  exhibits similar results. While the random forest model yields an average accuracy of  $q_{ACC} = 94.0\%$ , the average accuracy of the baseline model only amounts to  $q_{ACC} = 69.25\%$ . A result that can be expected when considering the rule-based creation of the synthetic data as described in Section 4.1.2. As stated in Algorithm 4.1 (lines 19 and 35), the only trips whose end locations are subject to randomness start from Location 2, i.e., the *work* location. Those trips possess a 50% chance of randomly visiting a grocery store twice a week. Contemplating the location-specific

**Table 4.4:** Comparison of baseline and random forest model (see Section 4.3.2) accuracy  $q_{\text{ACC}}$  for different scopes (start locations) when predicting the next location to visit on the synthetic data set  $\mathcal{J}_{\text{syn}}$  using a five-fold cross-validation.

Scope	Baseline Model	Random Forest
Location 1	52.93%	99.35%
Location 2	65.31%	76.69%
Location 3	89.32%	98.61%
Location 4	100.0%	100.0%
Location 5	100.0%	100.0%
Location 6	100.0%	100.0%
Location 7	98.57%	100.0%
Location 8	100.0%	100.0%
Average	69.25%	94.0%

results in Table 4.4 reveals this characteristic of the data set  $\mathcal{J}_{\text{syn}}$ . While the random forest model almost perfectly ( $q_{\text{ACC}} \cong 100.0\%$ ) predicts the next location for most start locations, Location 2 is the only location at which the accuracy drops to 76.69%. In this situation, however, even ideal models could not exceed 80% accuracy.<sup>11</sup>

Although the levels of accuracy for the synthetic data  $\mathcal{J}_{\text{syn}}$  are significantly beyond the levels of the real data  $\mathcal{J}_{\text{real}}$ , the results still provide a useful benchmark. The synthetic data features an inherent structure that arises from rule-based (see Algorithm 4.1) sequences of trips without irregular trips to unknown locations, e.g., vacation journeys. Regardless of the underlying user that the real data is taken from, it is very unlikely to be as structured as the synthetic data  $\mathcal{J}_{\text{syn}}$ . In terms of next locations to visit, any real data set is likely to be less predictable than  $\mathcal{J}_{\text{syn}}$ , accordingly (see also [279]). The random forest model's accuracy on  $\mathcal{J}_{\text{syn}}$  thus constitutes an qualitative upper bound for next place prediction.

#### 4.3.3.2 Probabilistic Evaluation of Random Forest Model

A point prediction of the next location to visit may not be helpful if all possible next locations and their corresponding probability are of interest. For this reason, the random forest model also estimates the conditional transition probabilities

---

<sup>11</sup>Five trips per week, three of which are completely predictable and two with 50%-uncertain destinations; hence,  $q_{\text{ACC}} \leq \frac{1}{5}(3 \cdot 100\% + 2 \cdot 50\%) = 80\%$ .

$\mathbb{P}_t(I|\mathbf{x}_I), \forall I \in \mathcal{L}$ . For a given context  $\mathbf{x}_I$ , the transition probability  $\mathbb{P}_t(I|\mathbf{x}_I)$  represents the chance to visit the location  $I$  next. When considering all combinations of start and end locations, i.e., current and next locations,  $\mathbb{P}_t$  may be presented in a *transition matrix*. Figure 4.6 shows the transition matrix of the random forest model trained on the real data  $\mathcal{J}_{\text{real}}$  for an arbitrary Monday morning at 05:00 o'clock.

		Next Location								$\Sigma$	$\mathbb{P}_r$
		1	2	3	4	5	6	7	8		
Current Location	1	0%	92%	2%	1%	0%	1%	2%	0%	98%	2%
	2	21%	58%	0%	15%	0%	0%	1%	1%	96%	4%
	3	35%	31%	2%	4%	0%	0%	4%	2%	78%	22%
	4	36%	27%	4%	0%	0%	0%	2%	2%	71%	29%
	5	33%	29%	2%	0%	1%	0%	1%	2%	68%	32%
	6	33%	30%	1%	1%	2%	0%	2%	3%	72%	28%
	7	31%	30%	1%	1%	6%	0%	2%	4%	75%	25%
	8	31%	29%	1%	0%	3%	0%	1%	1%	66%	34%

**Figure 4.6:** Transition matrix containing transition probabilities  $\mathbb{P}_t$  for all combinations of current and next location on a Monday morning at 05:00 o'clock; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1).

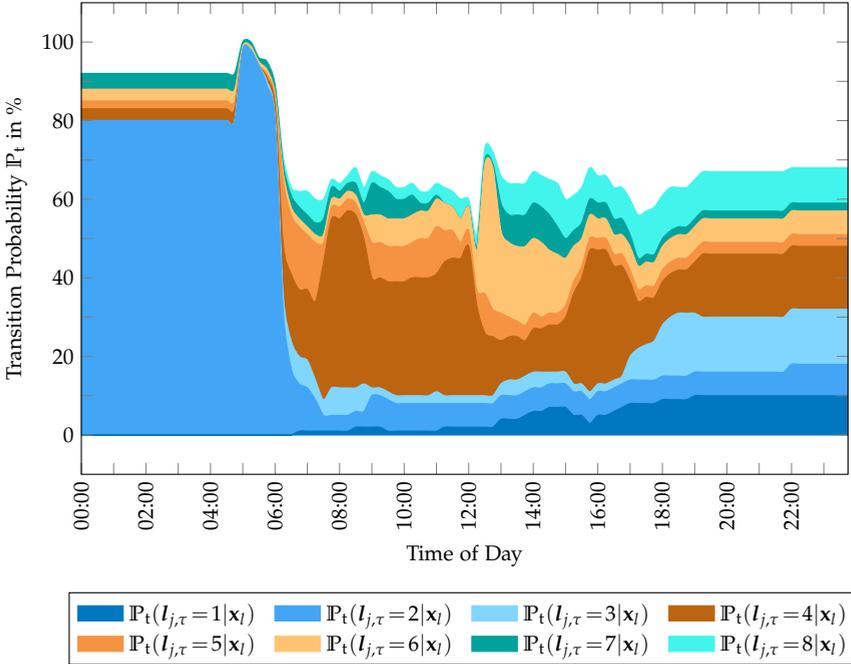
In addition to the transition probabilities, the summed probability for each current location is also calculated and given in a separate column; the sum may be less than one (or 100 %, respectively), i.e., if the next trip may have an unknown end location. For this case, the residual probability  $\mathbb{P}_r$  is deduced according to (4.4) and also given in another column.

It can be seen that on Monday morning the most likely trip will occur from Location 1 (*home*) to Location 2 (*work*) with a probability of  $\mathbb{P}_t = 92\%$ . Furthermore, the residual probability  $\mathbb{P}_r$  for trips starting from *home* only amounts to 2%; given this context, hence only 2 % of trips are predicted to have an unknown end location.

Note that the transition matrix as shown in Figure 4.6 is only valid for a single point in time.<sup>12</sup> To also evaluate the transition probabilities over time, a one-day simulation is performed. In particular, the random forest model's input vector  $\mathbf{x}_I$  is varied over the time of one day and for all known locations to obtain a set of  $\mathbb{P}_t$ -

<sup>12</sup>The baseline model (see Section 4.3.2.1) would be the simplest approach that yields a time-invariant transition matrix.

profiles. The result for an arbitrary Monday when starting a trip from Location 1 (*home*) is drawn in Figure 4.7.



**Figure 4.7:** Stacked transition probabilities  $\mathbb{P}_t$  for a trip starting from Location 1 (*home*) to all possible end locations on an arbitrary Monday; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1).

Here, a significant peak of the transition probability for trips from *home* to *work* can be detected between 05:00 and 06:00 o'clock. A result that can be expected when considering the mobility behavior of a typical commuter. At approx. 08:00 o'clock another peak can be observed for the probability of trips from Location 1 to Location 4. This could indicate a conditionally recurring trip to Location 4, e.g., dropping off children at school when working from home. The transition probabilities of going to any other location are relatively small ( $\mathbb{P}_t < 15\%$ ) for the majority of the time. Furthermore, after the “commute”-peak between 05:00 and 06:00 o'clock, the residual probability  $\mathbb{P}_r$  (represented by the white space between the stacked curves and 100%) always amounts to approx. 40%. The chance for a trip starting at *home* to end at an unknown location significantly rises after 06:00 o'clock, accordingly.

Similarly to [Figure 4.6](#), the transition matrix at 16:00 o'clock on an arbitrary Monday is shown in [Figure 4.8](#); for this, the random forest model was also trained on the real data  $\mathcal{J}_{\text{real}}$ .

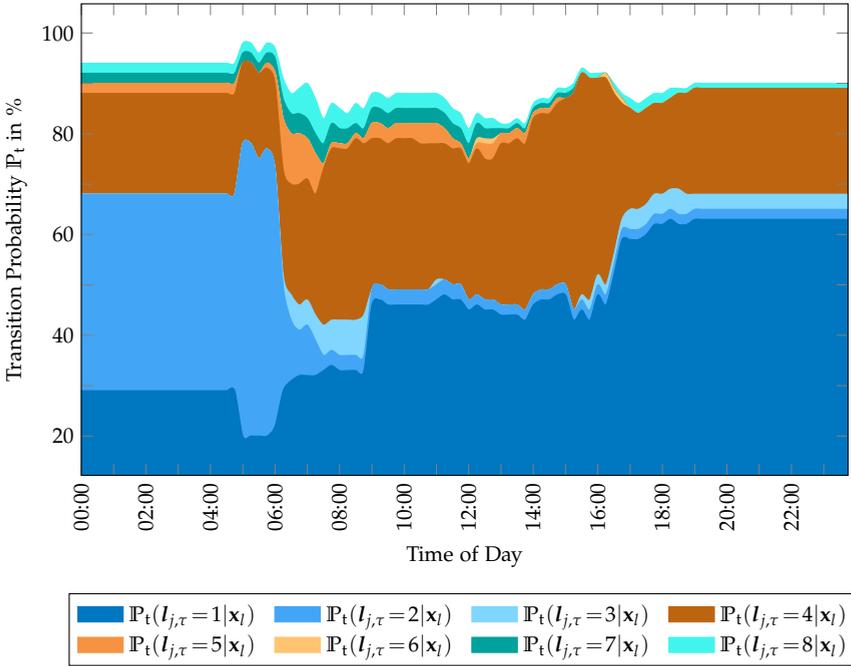
		Next Location								$\Sigma$	$\mathbb{P}_r$
		1	2	3	4	5	6	7	8		
Current Location	1	2%	7%	2%	41%	3%	5%	3%	9%	72%	28%
	2	44%	2%	0%	48%	0%	0%	0%	1%	95%	5%
	3	53%	0%	3%	27%	0%	3%	0%	2%	88%	12%
	4	71%	0%	8%	3%	1%	1%	0%	1%	85%	15%
	5	49%	0%	6%	10%	1%	2%	1%	1%	70%	30%
	6	62%	0%	5%	13%	0%	3%	1%	2%	86%	14%
	7	58%	0%	2%	13%	4%	4%	1%	3%	85%	15%
	8	48%	0%	4%	10%	4%	1%	1%	1%	69%	31%

**Figure 4.8:** Transition matrix containing transition probabilities  $\mathbb{P}_t$  for all combinations of current and next location on a Monday afternoon at 16:00 o'clock; random forest model (see [Section 4.3.2.2](#)) trained on real data set  $\mathcal{J}_{\text{real}}$  (see [Section 4.1.1](#)).

Here, especially trips starting from Location 2 (*work*) seem to be relevant as the summed probability of 95 % indicates; the residual probabilities  $\mathbb{P}_r$  for all other locations are relatively high. The end location for trips starting at *work* is predicted to either be Location 1 (*home*, 44 %) or Location 4 (48 %). Apparently, the *EV* user from which the real data set is taken, has two almost equally important trip destinations when leaving work on Monday afternoon. This can also be discovered when contemplating [Figure 4.9](#) in which the simulated  $\mathbb{P}_t$ -profiles for trips starting at *work* are shown.

After approx. 09:00 o'clock, both transition probabilities for trips to Location 1 (dark blue) and Location 4 (brown) steadily amount to values between 40 % and 60 %. Note that the residual probability throughout the entire day is always less than 20 %. In general, the chance for trips to end at an unknown location when starting at Location 2 is smaller than for trips starting at Location 1.

Although *round trips*, i.e., trips with the same start and end location were removed from the training and test data, on Monday morning still a 58 % chance of departing from Location 2 and also driving to Location 2 is predicted (see [Figure 4.6](#)). This observation is consistent with the falsely predicted transition probability for trips ending at Location 2, shown as the blue area in [Figure 4.9](#). This can be explained by the fact that on Monday morning barely any trips start from the *work* location.



**Figure 4.9:** Stacked transition probabilities  $\mathbb{P}_t$  for a trip starting from Location 2 (*work*) to all possible end locations on an arbitrary Monday; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1).

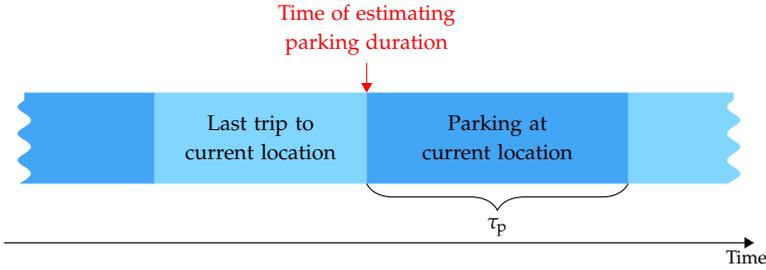
The training data for this situation is sparse and the prediction quality is expected to be poor, accordingly. In future work, the model may be engineered in a way such that—by design—identical current and next locations are excluded and thus cannot be predicted. Additionally, the prediction quality should also be estimated for each input sample, e.g., as additional model output. In this way, trustworthy predictions obtained from inputs sufficiently represented in the training data could be distinguished from predictions whose confidence is low.

Note that the  $\mathbb{P}_t$ -profiles as shown in Figure 4.7 and Figure 4.9 can be used to directly derive the expected energy demand over time. In particular, the energy consumption of each single trip in between two known locations can first be estimated; either using an external routing service and/or a user-individual consumption model as described in Section 3.1. Then, each trip’s energy consumption can be weighted by the transition probability  $\mathbb{P}_t$  of this trip occurring and summed up for each point in time afterwards; the resulting profile represents the expected energy

demand over time. In addition to that, a measure of prediction quality could be used together with the residual probability  $\mathbb{P}_r$  to quantify the uncertainty inherent in the expected energy profile; for this, a general energy consumption, e.g., the average, for all (historical) trips to unknown locations could be consulted. Both expected energy demand and the quantified uncertainty can subsequently provide valuable information for the robustness of **charging scheduling** (Chapter 5) and **charging optimization** (Chapter 6) schemes. The last mentioned aspects, however, are outside the scope of this thesis and may be subject to future work.

## 4.4 Estimation of Parking Duration at Known Locations

As evaluations on EV usage reveal (see e.g., Section 1.1.3.1), EVs (and also cars in general) are parked the majority of time. Hence, especially services that generate added value for EV customers while the EV is parked hold large economic potential. For example, **charging optimization** (see Chapter 6), which aims to improve an EVs charging process in terms of operating cost and sustainability. A sufficient knowledge of time flexibility—i.e., the excess time not required for charging when the EV is parked—is thus inevitable. For this, in turn, the *parking duration*  $\tau_p$  of an EV needs to be estimated upon arrival at a known location (see Section 4.2); Figure 4.10 illustrates the time at which the estimation is made.



**Figure 4.10:** Schematic sequence of trips (light blue) and parking periods (dark blue) to indicate the point in time (red arrow) of estimating the duration  $\tau_p$  of the next parking period.

In this section, first the engineering of features and labels based on historical trips  $\mathcal{J}$  and the previously determined known locations  $\mathcal{L}$  (see Section 4.2) is described. Then, different regression models to estimate the parking duration  $\tau_p$  are designed. Finally, all methods are evaluated with a custom metric and the results are presented.

### 4.4.1 Data Engineering for Parking Duration Estimation

Data engineering aims to combine the historical trip data  $\mathcal{J}$  with the determined known locations  $\mathcal{L}$  to fit the requirements of the predictive models described in Section 4.4.2.<sup>13</sup> First, for each trip  $j \in \mathcal{J}$  the start location  $l_{j,0} \in \mathcal{L}$  (i.e., the cluster number) is derived from the end location  $l_{j-1,\tau_p} \in \mathcal{L}$  of the previous trip  $j - 1$ . For

<sup>13</sup>All data engineering steps are separately proceeded for both data sets  $\mathcal{J}_{\text{syn}}$  and  $\mathcal{J}_{\text{real}}$ , see Section 4.1.

this, the previously determined known locations  $\mathcal{L}$  are used (see [Section 4.2](#)). If data gaps exist, e.g., from data cleaning as described in [Section 4.2.1](#), the start location of the consecutive trip is assigned to “noise”. At this step, each trip can also be enriched with additional information such as calculated energy consumption or the type of trip, i.e., spontaneous or planned; the latter information e.g., could be acquired from the user’s calendar. Together with one-time user inputs such as *home* and *work* location, the predictive models’ accuracy may thus be improved.

Second, the regression target, i.e., the estimated parking duration  $\hat{\tau}_{p,j}$  following each trip  $j \in \mathcal{J}$ , is approached via two different labels:

- **Label A:** Upon arrival of the EV at a known location, the duration of the time period until the next departure is estimated; this corresponds to directly estimating  $\tau_{p,j}$ . This label will also be referred to as *parking duration*.
- **Label B:** Upon arrival of the EV at a known location, the point in time the next departure occurs is estimated; subsequently,  $\hat{\tau}_{p,j}$  is calculated by subtracting the given arrival time from the estimated departure time. This label will also be referred to as *departure time*.

For the scope of this study, separate models are trained for both labels and evaluated separately. In future work, however, the estimations of Label A and Label B may be merged to enhance the overall accuracy. Ideas for this are given in [Section 4.4.3.2](#).

Third, additional features are designed from the end time of each trip to provide useful information to estimate  $\hat{\tau}_{p,j}$ . Similar to  $\mathbf{x}_l$  (see [Section 4.3.1](#)), all features are then combined in a feature vector  $\mathbf{x}_{\tau_p} \in \mathbb{R}^8$ ; the differences between  $\mathbf{x}_l$  and  $\mathbf{x}_{\tau_p}$  are described here:

- **Current Location:** The known location where the EV is currently parked; the feature “Last Location” is removed.
- All absolute time features (e.g., “Hour of Day”) refer to the end time of each trip.
- **Last Parking Duration at Current Location:** Alike to “Parking Duration at Current Location” but referring to the last visit of the current location.

To quantify the performance of the features in  $\mathbf{x}_{\tau_p}$ , the GINI importance  $q_{\text{Gini}}$  is used as shown in [Table 4.5](#).

The GINI importance (see its definition in [Section A.1.2](#)) is considered appropriate since the machine learning models described in [Section 4.4.2](#) are based on random forests [[276,277](#)].

**Table 4.5:** GINI importance  $q_{\text{Gini}}$  using a default random forest for all features when considering both Label A and Label B on the real data  $\mathcal{J}_{\text{real}}$ . Higher numbers indicate that the feature causes a larger mean decrease in impurity and is therefore more relevant to the random forest [276,277].

Feature	Importance $q_{\text{Gini}}$ to Label A	Importance $q_{\text{Gini}}$ to Label B
Current Location	0.0682	0.1750
Hour of Day	0.4390	0.3094
Time Window of Day	0.0076	0.0076
Day of Week	0.0587	0.0879
Month	0.0659	0.0736
Is Holiday	0.0152	0.0195
Last Parking Duration at Current Location	0.1672	0.1611
Last Trip Duration	0.1781	0.1658

Apparently, “Hour of Day” is the most important feature for both Label A and Label B. This suggests that a certain daily rhythm exists and that the EV user has regular trips based on the time of the day. Furthermore, the “Time Window of Day” feature, which is highly correlated with the hour of the day, is not important. This suggests that the information regarding the daily rhythm is sufficiently covered by the “Hour of Day” feature. The “Last Parking Duration at Current Location” and “Last Trip Duration” are also relevant for both labels, and thus, suggest repetitive patterns. Particularly, when the EV user visited a location for a certain period, they will likely stay for a similar duration the next time visiting this location. Furthermore, after a long trip the EV is likely to be parked longer than if the last trip was very short. Interestingly, the feature “Current Location” is not important for Label A but relatively important for Label B. The reason could be that the departure time (Label B) is more strongly coupled to the current location than the parking duration. For example, a typical user is expected to leave their *work* location at a similar time every workday regardless of when they arrived. The departure time can thus be accurately predicted based on the location; the parking duration, however, could vary broadly. Surprisingly “Day of Week”, “Month”, and “Is Holiday” only show low importance for both labels. This suggests that for the evaluated user underlying weekly and seasonal patterns are either nonexistent or not learned properly. This could be explained, however, by the fact that the real data set is obtained from only one user, and it cannot be ensured that they exhibit seasonal characteristics in their mobility behavior.

## 4.4.2 Models to Estimate Parking Duration

In this section, first a simple parking duration estimation model is presented. Then, three different machine learning models based on one or more random forests are designed; the according fundamentals can be found in [260]. All models are trained and tested for both labels described in Section 4.4.1, i.e., Label A (parking duration) and Label B (departure time). Furthermore, a five-fold cross-validation is proceeded on both data sets described in Section 4.1, i.e., the real data set  $\mathcal{J}_{\text{real}}$  and the synthetic data set  $\mathcal{J}_{\text{syn}}$ . The according train and test sets are denominated using the index “train” or “test”, respectively.

### 4.4.2.1 Mean Parking Duration per Known Location

As a naive baseline model, the mean parking duration of every known location is used as parking duration estimation. To achieve this, the training data set  $\mathcal{J}_{\text{train}}$  is filtered to only contain trips of a single known location  $l$ . Thus, a filtered training data set of trips

$$\mathcal{J}_{\text{train}}^{(l)} = \{j \in \mathcal{J}_{\text{train}} \mid l_{j,\tau} = l\}, \quad (4.5)$$

is obtained for each known location  $l \in \mathcal{L}$ . When arriving at a known location  $l \in \mathcal{L}$ , the parking duration can then be estimated as

$$\hat{\tau}_p = \frac{1}{|\mathcal{J}_{\text{train}}^{(l)}|} \sum_{\forall j \in \mathcal{J}_{\text{train}}^{(l)}} \tau_{p,j}, \quad \forall l \in \mathcal{L}. \quad (4.6)$$

Note that apart from the current location, the estimations of this model are independent from all other features in  $\mathbf{x}_{\tau_p}$ . This model is used to benchmark the more advanced machine learning models; in the following, it will be referred to as *Baseline* model.

### 4.4.2.2 Random Forest Regression Model

As a basic machine learning technique, a single random forest regression model as proposed by BREIMAN [260] is implemented. The random forest model consists of a maximum of 100 decision trees, which are individually trained with randomly selected subsets of the training data. For the training process, the MSE is used as loss metric (see also Appendix A). All other hyperparameters of the model are set to their default values as given by *Scikit-Learn* [254]. The resulting estimation  $\hat{\tau}_p$  is

obtained by averaging the individual estimations from each decision tree. For all details and fundamentals of random forests please refer to [260].

In the following, this model will be referred to as *Default RF* model.

#### 4.4.2.3 Location-Based Ensemble Model

The semantic parking location of the EV such as *home* or *work* is assumed to noticeably impact the parking duration. For this reason, the training data  $\mathcal{J}_{\text{train}}$  is split into subsets  $\mathcal{J}_{\text{train}}^{(l)}$  only containing trips with the same parking location  $l$ , as already described in (4.5). Then, a random forest model is trained for each subset, i.e., for each known location  $l \in \mathcal{L}$ . The problem complexity can thus be reduced and a set of less complex, location-specific random forest models is obtained. Accordingly, the overall estimation accuracy is also expected to be improved. This, however, also comes with higher computational complexity increasing with the number of known locations. In the present case, the number of known locations is relatively small and therefore the additional computation time is negligible.

This model will be referred to as *Location Ensemble* model in the following.

#### 4.4.2.4 Time-Based Ensemble Model

The stochastic nature of human mobility behavior causes both parking duration (Label A) and departure times (Label B) to vary randomly to some extent. Therefore, a regression approach that allows for a certain tolerance is conceivable. For example, instead of estimating a precise departure time such as 9:10am, the hour of day in which the departure takes place, i.e., between 9am and 10am, could be estimated. To achieve this calls for a multi-class classification model predicting in which hour the next departure occurs. However, the classes are interdependent and therefore an *ordered classification* problem exists, see [280]. For the present case, a manual approach to an ordered classification model in the form of a time-based ensemble is engineered as follows. Instead of training a single random forest model, a binary random forest *classifier* is trained for each hour in the relevant time window, to determine whether a departure occurs within this hour or not. The relevant time window is either the parking duration range for Label A, or each hour of a day for Label B, respectively. Based on this approach, the first rising edge, i.e., the change in the classification from “no departure” to “departure” is used as the estimated parking duration. Prospectively, the output of all classifiers could also be used to estimate a (time-discrete) probability density function of the parking duration.

To better understand this method, consider the following example of training and predicting for Label A. In the training process, the filtered parking duration is between 2 h to 24 h. Hence, 22 random forest classifiers are trained, beginning with a classifier between 2 h to 3 h. For a data sample with a parking duration of 7.35 h, the first five classifiers (for 2 h to 6 h) receive the label “no departure”. The remaining 17 classifiers are trained with the label “departure”. In the prediction process, the parking duration is calculated based on the first classifier to return the label “departure” given the input. Therefore, if the first five classifiers (for 2 h to 6 h) return “no departure” and the sixth classifier returns “departure”, the predicted parking duration would be 7 h. The time-based ensemble model thus tends to round off the true parking duration.

Note that the computational complexity increases linearly with the number of time slots defined; e.g., if instead of one-hour time slots, 15 min slots would be used, the computation time would increase by approx. a factor of four. In the following, this model will be referred to as *Time Ensemble* model.

### 4.4.3 Validation and Results

This section first presents the case study used to evaluate the predictive models described in Section 4.4.2. The evaluation is performed in a five-fold cross-validation on both the real data set  $\mathcal{J}_{\text{real}}$  and the synthetic data set  $\mathcal{J}_{\text{syn}}$  (see Section 4.1). To this end, all models and evaluation schemes are implemented in *Python* [256] using *Scikit-Learn* [254]. Only the eight most frequently visited locations are considered for both the training and evaluation of the parking duration estimation. All trips with unknown end location, i.e., those trips assigned to the “noise” cluster, are removed. In this way, the real data set  $\mathcal{J}_{\text{real}}$  can directly be compared with the synthetic data set  $\mathcal{J}_{\text{syn}}$ . Furthermore, the present study focuses on the parking duration estimation to improve the EV users’ experience of **charging optimization** (see Chapter 6). This optimization is not possible if the location—and as a result of this, the available charging infrastructure—is unknown. Therefore, the parking duration is only estimated for commonly visited, i.e., known locations.

The general model performance is quantified in a specific way, based on the later **charging optimization** use case of the predictive models. In this use case with the objective of energy cost reduction, only electricity prices with a time resolution of either 15 min or 60 min are available. Therefore, fluctuations within this threshold

do not affect the final optimization result (see also [Chapter 6](#)). Thus, the custom accuracy metric

$$q_{p,\varepsilon_{\tau_p}} = \frac{\left| \left\{ j \in \mathcal{J}_{\text{test}} : |\tau_{p,j} - \hat{\tau}_{p,j}| \leq \varepsilon_{\tau_p} \right\} \right|}{|\mathcal{J}_{\text{test}}|}, \quad (4.7)$$

is defined with  $\tau_{p,j}$  being the true parking duration, and  $\hat{\tau}_{p,j}$  the estimated parking duration. Equation (4.7) thus represents the percentage of estimated parking periods where the absolute difference between the estimated and actual parking duration is less than an error threshold  $\varepsilon_{\tau_p}$ . To evaluate different levels of quality, four values are considered for

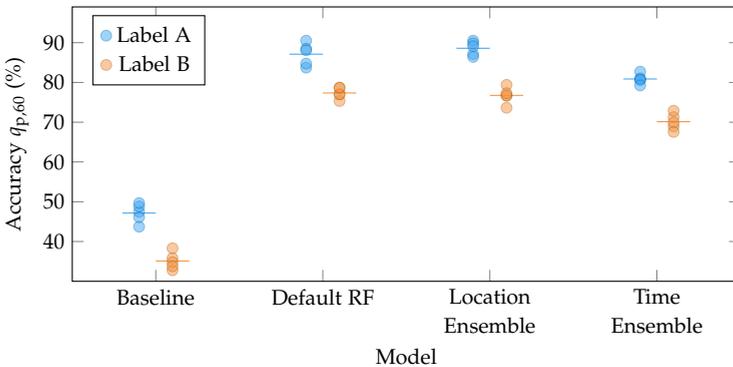
$$\varepsilon_{\tau_p} \in \{15, 30, 45, 60\}, \quad (4.8)$$

which are given in minutes.

Based on this case study, in the following, all major findings are presented and set into the broader context of this thesis.

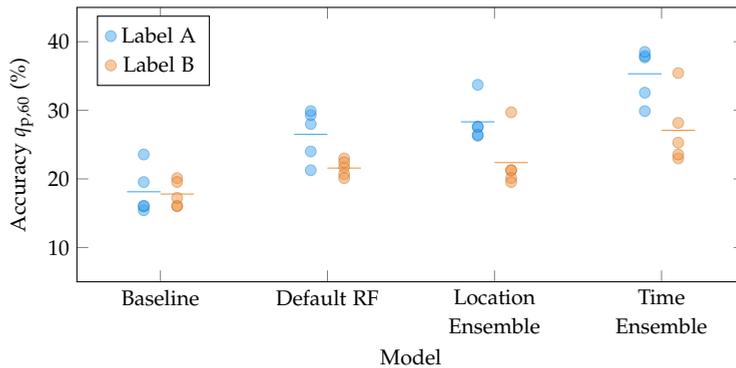
#### 4.4.3.1 Benchmark through Semi-Synthetic Data

To evaluate the general performance of the predictive models, the accuracy  $q_{p,60}$  is consulted. The results of the five-fold cross-validation are visualized in [Figure 4.11](#) and [Figure 4.12](#) for both Label A and Label B (see [Section 4.4.1](#)).



**Figure 4.11:** Accuracy ( $q_{p,60}$ ) of different models to estimate parking duration (Label A) and next departure time (Label B) on the synthetic data set  $\mathcal{J}_{\text{syn}}$  with five-fold cross-validation, horizontal line indicates mean accuracy.

For the synthetic data  $\mathcal{J}_{\text{syn}}$ , all three machine learning models vastly outperform the baseline, with accuracy nearing  $q_{p,60} = 90\%$  for Label A in some cases. Thus, even simple predictive models based on random forests seem to efficiently outperform the baseline given well-chosen features. The parking duration of the real data appears to be more challenging to predict. Although the performance is noticeably worse on the real data, even in this case, the default random forest model can outperform the baseline. Furthermore, note that regardless of the data set, the estimations for Label B are in general less accurate than those for Label A.



**Figure 4.12:** Accuracy ( $q_{p,60}$ ) of different models to estimate parking duration (Label A) and next departure time (Label B) on the real data set  $\mathcal{J}_{\text{real}}$  with five-fold cross-validation, horizontal line indicates mean accuracy.

These results clearly show that the synthetic data set  $\mathcal{J}_{\text{syn}}$  can be used as a benchmark for future analyses on EV users' mobility behavior. In particular, the synthetic data set does not contain “noise” or measurement errors (see Section 4.1). However, real data sets such as  $\mathcal{J}_{\text{real}}$  are likely to contain further stochastic influences, measurement errors, and irregular trips resulting in “noise” locations. Such real mobility behavior is expected to be less predictable than the synthetic data in any case (see also [279]). Accordingly,  $\mathcal{J}_{\text{syn}}$  is considered as benchmark data set, i.e., an upper bound for the accuracy of parking duration estimation.

#### 4.4.3.2 Benefit of Two-Model Regression Approach

Although the accuracy for Label B is always slightly lower than the one for Label A (see Figure 4.11 and Figure 4.12), including models for both regression targets may be beneficial. Particularly, the performance of the models for different labels seems

to depend on the current location. To illustrate this, consider the accuracy  $q_{p,60}$  for both Label A and Label B evaluated at three different locations as shown in Table 4.6. It can be seen that estimations made for the *home* location are more accurate

**Table 4.6:** Per-location accuracy ( $q_{p,60}$ , mean of five-fold cross-validation) of estimated parking duration (Label A) and next departure time (Label B), on the real data set  $\mathcal{J}_{\text{real}}$  using the default random forest model.

Regression Target	Home Location	Work Location	Back Office Location
Label A	25.60%	31.98%	34.0%
Label B	16.86%	33.28%	56.4%

when using the parking duration (Label A). The performance at the *work* location is similar for both labels. The reason could be that the *work* location is usually characterized by both regular departure times and steady parking duration (e.g., regular working day). Considering the “Back Office” location, on the other hand, estimating the departure time (Label B) is noticeably improved. This improvement can be explained by the fact that departure times from such a location are often determined by following appointments. However, the time spent at the “Back Office” location, i.e., the parking duration, may vary broadly. Accordingly, the predictive models’ performance differs for Label A and Label B.

This difference in performance may be leveraged to enhance the overall accuracy of estimating parking duration. Models for both labels exhibit better performance in different scenarios, and therefore both sources of information should be considered. To effectively decide, which regression target, i.e., Label A or Label B, is more accurate, however, a well-designed estimation merger is vital. The details of such a merger module has not been examined yet and should be subject to future work. A dependence on the current location, however, is strongly expected. Furthermore, user-specific dependencies may also be important, since mobility patterns may be diverse. For example, a person who works rotating shifts always stays at work for a relatively consistent amount of time. The arrival times at the *work* location, however, change frequently. In this case, estimating the parking duration (Label A) will likely outperform estimating departure times (Label B). In contrast, when contemplating a user working in a store always opening at a specific time, estimating departure times (Label B) may likely perform better than estimating the parking duration (Label A).

#### 4.4.3.3 Evaluation of Ensemble Models

In addition to a default random forest model, time- and location-based ensemble models were also designed, see Section 4.4.2. For comparison, the mean accuracy of all predictive models and various tolerances  $\varepsilon_{\tau_p}$  are used. Table 4.7 shows these results for estimating parking duration (Label A); Table 4.8 shows the values for estimating departure times (Label B).

**Table 4.7:** Comparison of mean error metrics for predictive models to estimate parking duration (Label A) based on the real data set  $\mathcal{J}_{\text{real}}$  with a five-fold cross-validation.

Model	MSE	$q_{p,15}$	$q_{p,30}$	$q_{p,45}$	$q_{p,60}$
Baseline	27.03 h <sup>2</sup>	5.39%	10.21%	15.03%	18.12%
Default RF	22.04 h <sup>2</sup>	5.73%	12.84%	20.41%	26.49%
Location Ensemble	21.32 h <sup>2</sup>	8.25%	15.94%	22.47%	29.01%
Time Ensemble	23.8 h <sup>2</sup>	8.26%	18.92%	27.52%	35.32%

**Table 4.8:** Comparison of mean error metrics for predictive models to estimate the next departure time (Label B) based on the real data set  $\mathcal{J}_{\text{real}}$  with a five-fold cross-validation.

Model	MSE	$q_{p,15}$	$q_{p,30}$	$q_{p,45}$	$q_{p,60}$
Baseline	16.02 h <sup>2</sup>	4.82%	8.50%	13.20%	17.80%
Default RF	14.18 h <sup>2</sup>	6.20%	11.48%	16.53%	21.24%
Location Ensemble	14.11 h <sup>2</sup>	6.65%	12.05%	18.02%	23.30%
Time Ensemble	16.52 h <sup>2</sup>	7.23%	15.26%	23.30%	27.89%

It can be seen that the location ensemble leads to increased accuracy compared with both the baseline and the default random forest model for Label A and Label B. This could be explained by the fact that the complexity of each individual, location-specific random forest model decreases. Each random forest can thus more easily learn existing dependencies related to parking duration. While combining several location-specific models in this way enhances the overall accuracy, the computational complexity—and thus the training and prediction time—also rises. For the present case with considering eight locations, however, this is a manageable problem.

In general, the performance of the time ensemble is even superior to that of the location ensemble. Compared with the default random forest model, the significant improvement in estimation accuracy could be explained by the adapted problem formulation that allows for a tolerance. The model can more easily handle stochastic

fluctuations of the parking duration (or the departure time), as it no longer tries to estimate the exact duration (or point in time, respectively). Instead, a time window of one hour is estimated in which the next departure will likely happen. However, the computational complexity of the time ensemble model increases with the time resolution selected for the individual classification models. In this study, an hourly resolution is chosen, but higher resolutions (e.g., 15 min) are conceivable. Although increasing the time resolution could improve estimation accuracy, it also increases computational cost. Hence, a trade-off exists, which should be examined further in future work.

As previously discussed in [Section 4.4.3.2](#), there might be a benefit of combining several models—in this case location and time ensemble—in an intelligent way. To achieve this, however, a suitable estimation merger would be required that should be subject to future work.

#### 4.4.3.4 Enhanced Service Quality via Estimation Confidence Filtering

As stated in the previous [Sections 4.4.3.1](#) to [4.4.3.3](#), the implemented machine learning models have proven to outperform a naive benchmark in estimating parking duration. Yet, the model accuracy  $q_{p,60}$  only amounts to 35.32 % for Label A using the real data set  $\mathcal{J}_{\text{real}}$ . The proportion of parking periods whose duration can be estimated with a tolerance of less than one hour is therefore only about one third on average. This level of accuracy is unlikely to be sufficient for EV users to accept smart charging applications such as **charging scheduling** ([Chapter 5](#)) and **charging optimization** ([Chapter 6](#)). The reason is that incorrect estimations could cause the EV operating cost to increase or even cause the EV not to be charged sufficiently when the user demands it.

To solve this problem, the confidence of each parking duration estimation may be determined and used. In this way, only trustworthy estimations could be used for subsequent smart charging; estimations assessed to be incorrect would then be withdrawn and default charging (i.e., plug in and charge with maximum power) could be used. To the end of determining estimation confidence, several approaches are conceivable. Similar to [\[281\]](#), e.g., two quantile regression models could be used to estimate the interquartile range; the smaller this range, the higher the confidence of the estimation. The difficulty here is to find a characteristic threshold for the interquartile range to effectively distinguish trustworthy from incorrect estimations.

**Table 4.9:** Comparison of time ensemble model accuracy for different scopes (known locations with corresponding number of data samples) when estimating parking duration (Label A) on the real data set  $\mathcal{J}_{\text{real}}$  with a five-fold cross-validation.

Scope	Samples	$q_{p,15}$	$q_{p,30}$	$q_{p,45}$	$q_{p,60}$
Location 1	1324	6.39%	15.68%	23.71%	31.74%
Location 2	332	6.58%	11.53%	25.37%	33.45%
Location 3	72	2.86%	12.86%	24.05%	24.05%
Location 4	54	21.33%	48.67%	82.0%	94.44%
Location 5	22	11.11%	11.11%	11.11%	11.11%
Location 6	16	30.0%	56.67%	66.67%	86.67%
Location 7	14	33.33%	33.33%	50.0%	50.0%
Location 8	12	0.0%	33.33%	44.44%	44.44%
Average	1846	8.26%	18.92%	27.52%	35.32%

A more practical approach using the existing models is presented here. Table 4.9 shows the results of the time ensemble (see Section 4.4.2) based on the real data set  $\mathcal{J}_{\text{real}}$  evaluated for different known locations. Although the average accuracy  $q_{p,60} = 35.32\%$  might not yield a satisfactory result when applied in **charging optimization**, the isolated accuracy for Location 4 ( $q_{p,60} = 94.44\%$ ) and Location 6 ( $q_{p,60} = 86.67\%$ ) could be sufficient. Thus, the current location where the vehicle is parked functions as an indicator for estimation confidence. For this data set, the **charging optimization** service could e.g., only be prompted if the **EV** is detected to be parked at either Location 4 or Location 6. However, these locations are only visited in approx. 4% of all cases. When parking at another location, default charging would be applied. Besides the current location, other features, e.g., the day of week, could be also used to evaluate the isolated model performance and deduce estimation confidence. This aspect, however, should be examined in future work.

## 4.5 Conclusion

In [Chapter 4](#) the mobility behavior of EV users is modeled in order to provide useful information for smart charging applications such as **charging scheduling** ([Chapter 5](#)) and **charging optimization** ([Chapter 6](#)). For this, real data from a single EV user is utilized. As a benchmark, furthermore, a synthetic data set is generated based on realistic user behavior, see [Section 4.1](#).

To determine frequently visited locations, i.e., the user's *known locations*, a two-stage clustering algorithm is designed in [Section 4.2](#). The first stage uses a state-of-the-art density-based spatial clustering algorithm and helps to level out GPS measurement scatter. In the second stage, determined clusters are combined to semantic locations based on given neighborhood parameters. In this way, the two-stage algorithm yields more robust results compared with standard density-based clustering. In addition, more realistic results can be achieved via the two-stage clustering as it can be parameterized to user-individual mobility habits; this accounts for a major advantage of the two-stage approach. In future work, the two-stage algorithm could be adapted to clustering of data streams, e.g., as described in [\[242\]](#).

Subsequently to clustering, models to predict the next location that the user may visit when departing from a known location are designed in [Section 4.3](#). To this end, a heuristic baseline model is compared with a random forest classification model. The use of a synthetic data set allows for a qualitative benchmark regarding the accuracy of next location prediction models. For the real data set, the random forest correctly predicts approx. two thirds of all locations visited next and thus achieves sufficient accuracy for smart charging applications. Although the random forest outperforms the baseline model on average, for some scenarios, e.g., specific locations, the baseline model yields similar or even better results. In the future, the predictions of both models could therefore be combined to achieve higher overall accuracy. Alternatively, different modeling approaches could be tested for improved accuracy, e.g., an individual random forest model for each known location. Furthermore, probabilistic predictions of the random forest model can be used to estimate an EV's expected energy demand in the future. In addition, a measure of uncertainty can be derived, which favors the robust design of smart charging applications based on next location prediction.

At third, different machine learning ensembles based on random forests are designed to estimate an EV's parking duration when arriving at a known location. Similar to the next place prediction, the use of a synthetic data set allows to suggest a qualitative benchmark for parking duration estimation. For this, both the parking

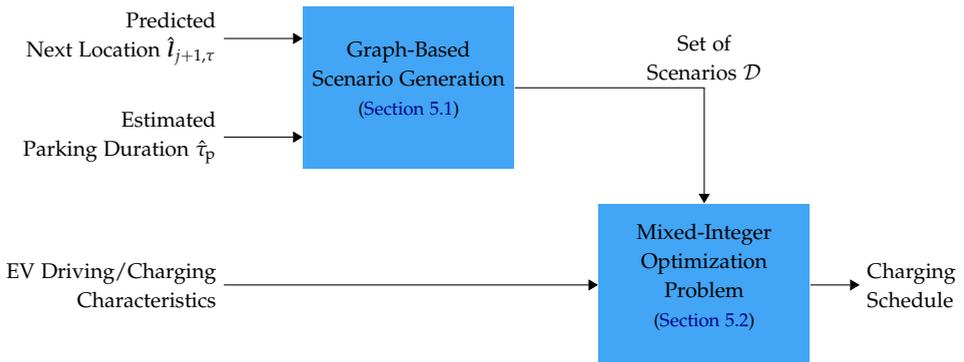
duration directly and also the absolute point in time at which the next departure occurs are estimated with separate models. Using a custom metric that allows for an application-specific tolerance, this two-model regression approach appears to be beneficial in terms of situation-dependent accuracy. Compared with a default random forest regression model, both a time- and location-based ensemble model exhibit improved performance. On the real data set, however, the best-performing model only estimates the duration of approx. one third of all parking phases with a deviation of less than one hour. This level of precision is assumed to be insufficient for most consecutive smart charging applications. Evaluating the models for specific known locations, on the contrary, shows sufficiently accurate estimations. Besides the current location, thus also other features, e.g., the day of week, could be used to deduce estimation confidence in the future. Follow-up work should also consider probabilistic estimations of the parking duration, e.g., using quantile regression models.

To prove general functionality, a validation of all algorithms on a variety of user types is inevitable. Furthermore, all models are user-individual and the available data is sparse, accordingly. To obtain a customer-ready service, a *ramp-up* concept is required, which allows for trustworthy predictions with a minimum of data at the beginning. For this, the problem formulation could be generalized to leverage data of several EV users in creating general predictive models. To incrementally adapt these models to individual users, a suitable adjustment of model parameters, i.e., *transfer learning*, could be applied afterwards.



# 5 Charging Scheduling

The requirements of humans' individual mobility in most cases can be very diverse (see Section 1.1.3). For instance, both planned and especially spontaneous trips possess an uncertainty inherent in their start time, destination, travel time, and energy consumption. It is therefore both difficult and inconvenient for EV users to keep track and maintain a sufficiently charged battery to stay mobile.<sup>1</sup> The **charging scheduling** module as proposed in Figure 5.1 supports EV users to i) keep an overview of upcoming charging events that are required to stay mobile, ii) find convenient charging opportunities, and iii) ensure a battery reserve to be able to reach any desired destination given a user-individual security level. Rather than



**Figure 5.1:** Block diagram of the charging scheduling scheme; mobility prediction (Chapter 4) determines predicted next location  $\hat{l}_{j+1,\tau}$  and estimated parking duration  $\hat{t}_p$ , EV driving and charging characteristics are derived from the corresponding EV models (Chapter 3).

adapting single EV charging processes, thus, the aim is to determine adequate locations, time windows, and amounts of energy for multiple charging opportunities in a future time horizon.

<sup>1</sup>The term “staying mobile” refers to EV trips that do not require additional, unplanned charging stops along the way.

To achieve this, the outputs of the **mobility prediction** module as described in [Chapter 4](#) are leveraged. First, a set of scenarios  $\mathcal{D}$ , i.e., possible sequences of trips in between all known locations  $\mathcal{L}$ , is generated based on the predicted next location  $\hat{l}_{j+1,\tau}$ , the corresponding estimated probability  $\mathbb{P}_t$ , and estimated parking duration  $\hat{\tau}_p$ . In this way, the uncertainty inherent in the user's actions can be represented. All details can be found in [Section 5.1](#).

To model the decision problem of where and how much energy to charge, a mixed-integer optimization problem is designed in [Section 5.2](#). Besides the previously generated set of scenarios  $\mathcal{D}$ , the EV's characteristics on energy consumption, charging, and battery aging (see [Section 2.2.1](#)) are also consulted. Solving the optimization problem then yields a *Charging Schedule* for a given future horizon (e.g., five days) that contains the following information:

- Required user actions, i.e., park and plug in the EV at a close-by charging station when approaching a targeted location at a specific time.
- The amount of energy to charge at a specific location.

Once the EV is connected to a charging station, a subordinate scheme, e.g., as in [Chapter 6](#), may process the amount of energy to charge together with the expected parking duration to calculate a single charging event's power trajectory.

To prove and quantify the functionality of the **charging scheduling** module, finally a MONTE CARLO simulation is designed and run. The simulation setup and corresponding results are outlined in [Section 5.3](#). Note that the contents of this chapter are partly based on and further develop the concepts presented by SCHWENK et al. in [[271](#),[282](#),[283](#)].

## 5.1 Scenario Generation

Although the sources of uncertainty in humans' individual mobility are manifold (see [Section 1.1.3](#)), the highest impact is assumed to arise from the user's decision of where to go next. The uncertainty inherent in these decisions needs to be quantified in order to formulate a computationally tractable decision problem (see [Section 5.2](#)), which is able to handle stochastic influences. Thus, all possible mobility scenarios a user may take—i.e., all sequences of trips and visits of a known location—are determined (see also [\[284\]](#)). For this purpose, a graph-based approach is used that implements the following recursive procedure.<sup>2</sup>

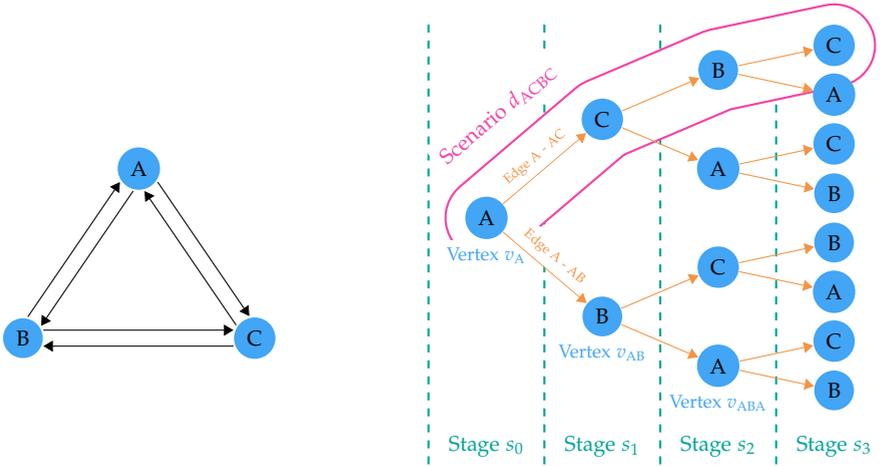
Each time the **charging scheduling** module is triggered, an out-tree  $\mathcal{T}$  with a set of vertices  $\mathcal{V}$  is created based on the set of known locations  $\mathcal{L}$  that was determined beforehand, see [Section 4.2](#). The root vertex  $v_0$  represents the current location  $l_0$  of the EV at the time of calculation. For each possible next location  $l \in \mathcal{L} \setminus l_0$  despite the current location, a successor vertex  $v_l$  is then connected to  $v_0$  via a directed edge. The procedure is repeated, i.e., the successor vertices are searched for each previously appended vertex until a termination criterion is met; e.g., a defined future time horizon, a number of locations visited, or a minimum occurrence probability.

To better understand this procedure and introduce further nomenclature, consider the example as shown in [Figure 5.2](#). Here, an exemplary set of known locations  $\mathcal{L} = \{A, B, C\}$  (see [Figure 5.2a](#)) is taken as a basis on which the tree  $\mathcal{T}$  (see [Figure 5.2b](#)) is constructed. The start location A is represented by vertex  $v_A$  from which only the remaining locations B and C may be visited. Accordingly, the vertices  $v_{AB}$  and  $v_{AC}$  are connected to  $v_A$  via the edges A-AC and A-AB, respectively. This procedure is then repeated for each vertex until—in this example—three consecutive trips have been taken from the start location.

Each vertex  $v \in \mathcal{V}$  (see [Figure 5.2](#), drawn in blue) represents a visit of a known location  $l \in \mathcal{L}$  at a specific time and is therefore associated with the properties:

- Start time  $t_{v,0}$ , i.e., the point in time at which the EV arrives at vertex  $v$ .
- End time  $t_{v,\tau_p,v} = t_{v,0} + \hat{\tau}_{p,v}$ , i.e., the point in time the EV is assumed to depart from vertex  $v$ ; the estimated parking duration  $\hat{\tau}_{p,v}$  is determined using a model as described in [Section 4.4](#).

<sup>2</sup>For the fundamentals of graph-based concepts and algorithms, see e.g., [\[285\]](#).



(a) Set of Known Locations  $\mathcal{L}$

(b) Constructed Scenario Tree  $\mathcal{T}$

**Figure 5.2:** Exemplary generation of scenarios based on a set  $\mathcal{L} = \{A, B, C\}$  of three known locations (a) and the constructed out-tree  $\mathcal{T}$  (b) to demonstrate the scenario generation procedure and nomenclature.

- Occurrence probability  $\mathbb{P}(v)$ , i.e., the probability that  $v$  is actually visited; determined by multiplying all transition probabilities  $\mathbb{P}_t$  of the tree edges required to reach  $v$ . The root vertex—in this case  $v_A$ —always possesses an occurrence probability of 100%, as it is already being visited at the time of calculation.
- Charging power limitations  $\underline{p}_v$  and  $\bar{p}_v$ , i.e., the maximum and—in case of bidirectional charging negative—minimum available charging power; both are equal to zero if there is no charging station available at this location.
- Walking distance  $w_v$ , i.e., the distance between the location the user intends to visit and the closest charging station; this quantity is used for convenience aspects later, see [Section 5.2.1](#), and may also correspond to other factors such as user preferences and availability of charging stations.

- Electricity price  $\epsilon$ , i.e., the specific price at which electric energy can be purchased at the closest charging station; this quantity is assumed to be constant throughout a single stay at each vertex.<sup>3</sup>

All edges (see Figure 5.2b, drawn in orange) of the tree represent trips of the EV, e.g., connecting two tree vertices  $v_A, v_{AB} \in \mathcal{V}$ . Therefore, edges are associated with the properties:

- Transition probability  $\mathbb{P}_t(v_{AB}|v_A)$ , i.e., the probability of traveling to  $v_{AB}$  when departing from  $v_A$ ; estimated using a model as described in Section 4.3.
- Travel time  $\tau_{(v_A, v_{AB})}$ , i.e., the estimated time required to travel from location A to B at a specific point in time; estimated using an external routing service, or prospectively, from historical trip data of the user.
- Energy consumption  $\Gamma_{(v_A, v_{AB})}$ , i.e., the energy required to travel from location A to B at a specific point in time; e.g., estimated using a model as in Section 3.1.

All vertices reachable within the same number of trips (i.e., edges) starting from the root vertex (here  $v_A$ ) are combined in a set  $\mathcal{V}_s \subset \mathcal{V}$ ; in the following, these sets are called *stages*  $s$  of the tree (see Figure 5.2b, drawn in green). Note that the user might also visit other, unknown locations, which are not in the set of known locations. The summed occurrence probability for all vertices  $v \in \mathcal{V}_s$  of a stage  $s$  can therefore be less than 100%, accordingly. To quantify the chance of the user visiting another location besides the ones represented by vertices, the residual probability

$$\mathbb{P}_r(s) = 1 - \sum_{\forall v \in \mathcal{V}_s} \mathbb{P}(v), \quad (5.1)$$

is introduced for each stage of the tree.<sup>4</sup> For better comprehensibility, an additional “Noise”-vertex can be imagined to exist in each tree stage  $s$  to which  $\mathbb{P}_r(s)$  is assigned to. If  $\mathbb{P}_r = 0$ , all possible next locations the user might visit are known. For  $\mathbb{P}_r = 1$ , on the other hand, none of the known locations is assumed to be visited in a stage; instead, the “Noise”-vertex would be certainly visited. In Section 5.2.4, an approach is presented to account for the uncertainty of visiting an unknown location.

Each path connecting the root vertex (here  $v_A$ ) to a vertex of the final tree stage is called a *scenario*  $d$  (see Figure 5.2b, drawn in pink). All corresponding vertices are

<sup>3</sup>If a dynamic electricity tariff is available (see e.g., [286]), the mean price for the time window at each vertex is used.

<sup>4</sup>See also (4.4) in Section 4.3.

summarized in set  $\mathcal{V}_d \subset \mathcal{V}$ . The set of scenarios  $\mathcal{D}$ , in turn, summarizes all possible scenarios  $d$  of a tree. To avoid a large number of scenarios to be considered in the optimization problem (see Section 5.2), whose occurrence probability is very small, a scenario reduction is applied (see also [271, 287]). In particular,  $\mathcal{D}$  only contains scenarios with an estimated probability higher than a given threshold, e.g., 1%. For consistency, vertices not being visited in any remaining scenario are removed from  $\mathcal{V}$  and their probability weight is added to the residual probability  $\mathbb{P}_r$ . The charging schedule resulting from the subsequent optimization problem will therefore disregard very unlikely events. Note that if no scenarios remain after scenario reduction, the scheduling of charging stops is not considered reasonable. Users may then receive a corresponding notification prompting them to charge conservatively until a (repeated) recalculation of the scheduling scheme provides a valid result again.

Beyond that, users can always manually request a recalculation via a user interface in order to receive an updated charging schedule. For this purpose, already generated scenario trees may be reused to save computation resources and time (*memoization*, see [288]) Prospectively, additional vertices with high confidence, e.g., appointments or planned trips taken from the user's calendar, may be furthermore included in the scenario tree to enhance the robustness of the scheduling scheme, see also [271].

## 5.2 Optimization Problem

Determining possible scenarios, i.e., sequences of trips and visits of known locations, in the previous [Section 5.1](#), allows to identify all available charging opportunities, i.e., locations and time windows in which the EV is parked close to a charging station. The question of where and how much energy to charge best—given a defined objective, e.g., electricity costs—is, however, not trivial for EV users. To support users with this issue, first, a mathematical model of the multi-dimensional decision problem is required. Solving this problem then yields the desired decision policy, i.e., the *charging schedule* that can be provided to the user via an interface, e.g., a smart phone app or the EV's infotainment system.

For this purpose, the following stochastic mixed-integer optimization problem is designed such that

$$\min \quad \mathbb{E}(J_C + J_E) \quad (5.2a)$$

subject to

$$\underline{E}_v \leq E_v \leq \bar{E}_v, \quad E_v \in \mathbb{R}, \forall v \in \mathcal{V}, \quad (5.2b)$$

$$0 \leq C_v \leq 1, \quad C_v \in \mathbb{Z}, \forall v \in \mathcal{V}, \quad (5.2c)$$

$$\underline{E}_v \cdot C_v \leq E_v \leq \bar{E}_v \cdot C_v, \quad \forall v \in \mathcal{V}, \quad (5.2d)$$

$$\underline{e} \leq e_{A,v} \leq \bar{e}, \quad \forall v \in \mathcal{V}, \quad (5.2e)$$

$$\underline{e} \leq e_{D,v} \leq \bar{e}, \quad \forall v \in \mathcal{V}, \quad (5.2f)$$

$$e_{D,v} = e_{A,v} + E_v, \quad \forall v \in \mathcal{V}_d, \forall d \in \mathcal{D}, \quad (5.2g)$$

$$e_{A,v} = e_{D,0} - \sum_{\substack{\forall i, j \in \\ \{\mathcal{V}_d | i \prec j \preceq v\}}} \Gamma_{(i,j)} + \sum_{\substack{\forall k \in \\ \{\mathcal{V}_d | k \prec v\}}} E_k, \quad \forall v \in \mathcal{V}_d, \forall d \in \mathcal{D}, \quad (5.2h)$$

$$\mathbb{P}(e_{A,v} < \tilde{e}_s) \leq \varepsilon_v, \quad \forall v \in \mathcal{V}_s, \forall s \in \mathcal{S}, \quad (5.2i)$$

$$\mathbb{P}(e_{D,v} < \tilde{e}_s) \leq \varepsilon_v, \quad \forall v \in \mathcal{V}_s, \forall s \in \mathcal{S}. \quad (5.2j)$$

Note that by using graph-based scenario generation, the decision problem is decoupled from time. In particular, each parking phase of the EV is represented by a vertex  $v \in \mathcal{V}$ . Furthermore, all vertices  $v \in \mathcal{V}_s$  of a single stage  $s$  may have different start and end times. From a mobility aspect, however, they all represent a potential charging opportunity reached after a specific number of trips with a corresponding energy consumption. Applying a graph-based approach therefore enables to adequately represent this condition in the optimization problem.

The following sections further outline all components of (5.2) in detail.

### 5.2.1 Optimization Objective of Charging Scheduling

Charging scheduling aims to minimize the expected (notional) cost resulting from two components. First, the plug-in costs

$$J_C = \sum_{\forall v \in \mathcal{V}} C_v \cdot (\epsilon_{C,v} + \phi_w \cdot w_v), \quad (5.3)$$

which are meant to represent constant costs  $\epsilon_C$  that might be billed at some charging stations. Furthermore, the user's discomfort arising from plugging in the EV at a charging station is included. In particular, the walking distance  $w_v$  between the users intended destination and the charging station is consulted for this. Depending on a user's preferences, the walking distance can be individually weighted with a factor  $\phi_w$ ; the more important convenience is for a user, the higher  $\phi_w$  should be chosen. Both costs  $\epsilon_{C,v}$  and  $\phi_w \cdot w_v$  only need to be accounted for, if the EV is plugged in; for this reason, the term  $(\epsilon_{C,v} + \phi_w \cdot w_v)$  is multiplied by the plug-in flag  $C_v$ , which is equal to one, if the EV is plugged in, and zero otherwise. Finally, the single costs for each vertex  $v \in \mathcal{V}$  are summed up.

As second cost component, the electricity costs

$$J_E = \sum_{\forall v \in \mathcal{V}} \frac{\epsilon_v \cdot E_v}{\eta_{c,v}} \quad (5.4)$$

are defined to consider the variable costs for charging electric energy. Particularly, the net charged energy  $E_v$  is multiplied with a location- and time-dependent electricity price  $\epsilon_v$ . To account for energy losses in the charging process, the term is divided by the efficiency  $\eta_{c,v}$  associated with vertex  $v$ . Similar, to (5.3), the single costs for all vertices  $v \in \mathcal{V}$  are then summed up.

Note that the energy cost function  $J_E$  implicitly penalizes calendar battery aging. Assuming a globally optimal solution to (5.2), only the minimum required amount of energy is charged. By the end of the scheduling time horizon both the battery energy  $e_A$  upon arrival at and  $e_D$  upon departure from of the last vertex meet the adapted minimum energy  $\tilde{e}_s$ . Accordingly, the average level of the energy stored in the EV battery is reduced. As stated in (3.23) (see Section 3.2.3), calendar aging is proportional to the energy stored in the battery. Consequently, calendar aging also decreases with a reduction of  $J_E$ .

Note that the actual values of both  $J_C$  and  $J_E$  depend on the realized scenario  $d \in \mathcal{D}$ , which in turn, depends on the user's random decisions. At the time of calculation, the realized scenario is therefore unknown. Hence, only the expected value of  $J_C$  and  $J_E$  can be minimized, see (5.2a). To achieve this, the costs associated with

each vertex  $v \in \mathcal{V}$  are multiplied with the corresponding estimated occurrence probability  $\mathbb{P}(v)$ .

## 5.2.2 Decision Variables

To obtain an optimal policy of where and how much energy to charge, two decision variables are defined for each vertex  $v \in \mathcal{V}$ .

First, the net energy  $E_v \in \mathbb{R}$  that is charged into the battery at each vertex  $v \in \mathcal{V}$  is defined. Taking (1.1) as a basis, it can also be contemplated as the difference of the battery's SOC. Due to power limitations, e.g., given by the charging station or the EV power electronics, (5.2b) constrains  $E_v$ . Here, the lower bound

$$\underline{E}_v = \eta_{c,v} \cdot \underline{p}_v \left( t_{v,\tau_{p,v}} - t_{v,0} \right), \quad \forall v \in \mathcal{V}, \quad (5.5)$$

is equal to zero in case of unidirectional charging, as  $\underline{p}_v = 0$  kW. If, on the other hand,  $\underline{p}_v < 0$  kW, i.e., bidirectional charging is available,  $\underline{E}_v$  represents the maximum amount of energy that can be discharged at vertex  $v$ . Similarly, the upper bound

$$\bar{E}_v = \eta_{c,v} \cdot \bar{p}_v \left( t_{v,\tau_{p,v}} - t_{v,0} \right), \quad \forall v \in \mathcal{V}, \quad (5.6)$$

represents the maximum amount of energy that can be charged at vertex  $v$ . Here, the energy efficiency  $\eta_c$  is determined by simulating a charging process using the electrical battery model as described in Section 3.2.1.

Second, a binary variable  $C_v \in \mathbb{Z}$  is defined, which indicates if the EV needs to be plugged into a charging station when parking at vertex  $v \in \mathcal{V}$ . As this variable can only take two values for each vertex, it is constrained by (5.2c). For  $C_v = 1$ , the EV needs to be connected to a charging station at vertex  $v$ ; if  $C_v = 0$ , the EV does not need to be plugged in.

Apparently, the EV cannot be charged if it has not been plugged into a charging station before. To ensure a reasonable combination of plug-in events and the charged energy at each vertex,  $E_v$  is additionally constrained by (5.2d). This constraint ensures in particular that

$$C_v = \begin{cases} 1, & \forall E_v \neq 0, \\ 0, & \forall E_v = 0, \end{cases} \quad (5.7)$$

for all vertices  $v \in \mathcal{V}$ . Thus, the decision to charge or discharge energy at vertex  $v$ , i.e.,  $E_v \neq 0$ , depends on the corresponding plug-in event, i.e.,  $C_v = 1$ .

### 5.2.3 State Variables

In addition to the decision variables described in Section 5.2.2, state variables are defined, which represent the EV battery's energy content (and thus, the SOC) throughout the scheduling time horizon. Particularly, the state variable  $e_{A,v} \in \mathbb{R}$  represents the battery level  $e$  when arriving at vertex  $v$ . Similarly,  $e_{D,v} \in \mathbb{R}$  represents the battery level  $e$  upon departure from vertex  $v$ . The energy level of the EV battery has to be kept within the global battery energy limits  $\underline{e}$  and  $\bar{e}$  at all times. These boundaries are either defined by the battery capacity  $e_{\text{nom}}$  or by user preferences, e.g., a constant battery reserve of 20% SOC. To achieve this, (5.2e) constrains the arrival energy state  $e_{A,v}$ ; in a similar manner,  $e_{D,v}$  is limited via (5.2f).

At each vertex  $v \in \mathcal{V}$ , an amount of energy  $E_v$  may be charged into (or discharged from) the EV battery. Accordingly, the battery level  $e_{D,v}$  upon departure can deviate from the arrival energy  $e_{A,v}$ . Thus, the transition constraint (5.2g) represents the charging of energy.<sup>5</sup> In between two vertices, e.g.,  $v_A$  and  $v_B$ , a trip is taken that is associated with an energy consumption  $\Gamma_{(v_A,v_B)}$ .<sup>6</sup> Thus, the energy  $e_{A,v_B}$  upon arrival at vertex  $v_B$  deviates from the departure energy  $e_{D,v_A}$  of its predecessor vertex  $v_A$ . To represent this condition, the transition constraint (5.2h) is defined for each vertex  $v \in \mathcal{V}_d$  along each scenario  $d \in \mathcal{D}$ . Note the notation

$$\forall k \in \{\mathcal{V}_d | k \prec v\}, \quad \forall d \in \mathcal{D}, \quad (5.8)$$

which represents the set of all predecessors to  $v$  in scenario  $d$ , i.e., all vertices that were visited before reaching  $v$  when starting from the root vertex. All edges, i.e., connected pairs  $i, j$  of vertices, within this set are represented by means of the notation

$$\forall i, j \in \{\mathcal{V}_d | i \prec j \preceq v\}, \quad \forall d \in \mathcal{D}. \quad (5.9)$$

Note that the root vertex  $v_0$  representing the current location at the time of calculation does not have a temporal extension, i.e.,  $t_{v_0,0} = t_{v_0,\tau_p,v}$ ; no energy can thus be charged or discharged at this vertex, i.e.,  $E_{v_0} = 0$  kW. For the initial energy, i.e., the energy content of the battery at the time of calculation, the relation  $e_{A,0} = e_{D,0}$  therefore holds.

<sup>5</sup>The battery of EVs parking for long time periods (several weeks) without charging may be subject to self-discharge. However, the corresponding amounts of energy are negligible for the present use case and are therefore not considered in (5.2g).

<sup>6</sup>For trips whose energy consumption exceeds the (remaining) battery energy, an EV routing service is used to insert further charging stops along the way, see e.g., [289]. The energy consumption  $\Gamma_{(v_A,v_B)}$  is accordingly lowered by the additional energy charged; similarly, the travel time in between  $v_A$  and  $v_B$  is increased by the time required for driving to a charging station and charging itself.

## 5.2.4 Chance Constraints on Adapted Minimum Energy

As mentioned in [Section 5.1](#), the vertices  $\mathcal{V}_s$  of a tree stage  $s$  only represent the visits of known locations. The chance that the user, however, may also visit other, unknown locations is quantified by the estimated residual probability  $\mathbb{P}_r(s)$  as defined in (5.1). Accordingly, an additional battery reserve seems conceivable, which is derived from the expected energy consumption when visiting such an unknown location. To embed this characteristic in the charging scheduling scheme (5.2), an adapted minimum energy level

$$\tilde{\epsilon}_s = f(\mathbb{P}_r(s), \underline{\epsilon}, \Gamma_{l \notin \mathcal{L}}), \quad (5.10)$$

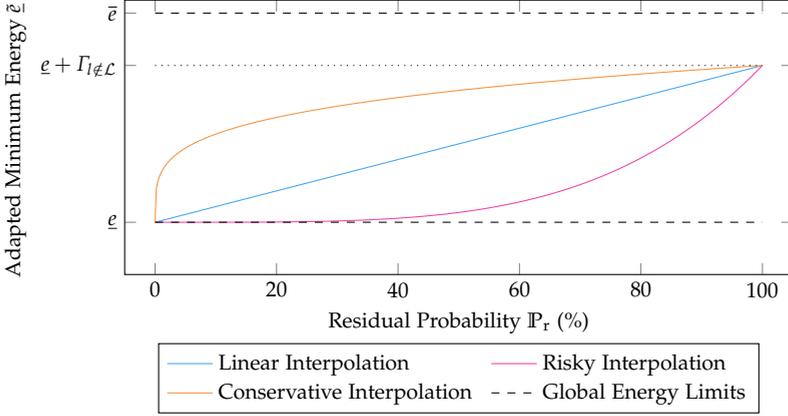
is introduced for each stage  $s$ ;  $\tilde{\epsilon}_s$  depends on the estimated residual probability  $\mathbb{P}_r(s)$  per stage  $s$ , the lower battery limit  $\underline{\epsilon}$ , and the additional energy consumption  $\Gamma_{l \notin \mathcal{L}}$  for trips to unknown locations. The last quantity  $\Gamma_{l \notin \mathcal{L}}$  can be derived from the energy consumption of historical trips to unknown locations; e.g., using the mean value (or other metrics, e.g., the 75%-quantile) of their energy consumption.<sup>7</sup> Then, the characteristic of  $\tilde{\epsilon}_s$  are designed such that  $\tilde{\epsilon}_s = \underline{\epsilon}$  for  $\mathbb{P}_r(s) = 0$ , i.e., if all possible next locations the user might visit are known. Furthermore,  $\tilde{\epsilon}_s = \underline{\epsilon} + \Gamma_{l \notin \mathcal{L}}$  for  $\mathbb{P}_r(s) = 1$ , i.e., in the case that none of the known locations is assumed to be visited. In between these two support points, a linear interpolation

$$\tilde{\epsilon}_s = \underline{\epsilon} + \Gamma_{l \notin \mathcal{L}} \cdot \mathbb{P}_r(s), \quad (5.11)$$

is used as intuitive baseline approach, see [Figure 5.3](#), drawn in blue. Alternatively, more conservative or riskier interpolations could also be used, see e.g., [Figure 5.3](#) drawn in orange or pink, respectively. This dependency, however, might also be user-specific and may therefore require an individual adaption depending on the user's risk aversion.

Depending on a user's mobility habits, the residual probability  $\mathbb{P}_r(s)$  may take high values close to 100%—especially for users with irregular mobility behavior. As a consequence, the adapted minimum energy  $\tilde{\epsilon}_s$  also takes on high values. If the lower energy bound  $\underline{\epsilon}$  in (5.2e) and (5.2f) would simply be replaced with  $\tilde{\epsilon}_s$ , the resulting charging schedule may become over-conservative. In particular, the optimization scheme would target a fully charged battery whenever possible. To provide the ability to counteract over-conservative results, instead, the adapted minimum energy requirement is formulated as chance constraints (5.2i) and (5.2j). In addition

<sup>7</sup>Note that  $\Gamma_{l \notin \mathcal{L}}$  should be chosen such that  $\underline{\epsilon} + \Gamma_{l \notin \mathcal{L}} \leq \bar{\epsilon}$ , as otherwise (5.2) may be infeasible.



**Figure 5.3:** Different interpolations for the adapted minimum energy  $\tilde{e}_s$ , i.e., the dynamic battery reserve for charging scheduling, as a function of the residual probability  $\mathbb{P}_r(s)$ .

to the characteristics of  $\tilde{e}_s$ , the user can thus also provide a personal security level  $\varepsilon_v \in [0, 1]$ ; it represents the maximum share of cases (e.g., 5%) in which the battery energy may undercut the adapted minimum energy  $\tilde{e}_s$  when arriving at or departing from a vertex.

Note that the numerical solution of (5.2) depends on computational tractability, which the chance constraints (5.2i) and (5.2j), however, impede. Hence, a suitable reformulation is required. To this end, first additional auxiliary integer variables  $z_{A,v} \in \mathbb{Z}$  and  $z_{D,v} \in \mathbb{Z}$  are introduced for each vertex  $v \in \mathcal{V}$ . These variables function as a flag to indicate whether the battery level upon arrival or departure undercuts the adapted minimum energy or not. Accordingly, the conditions

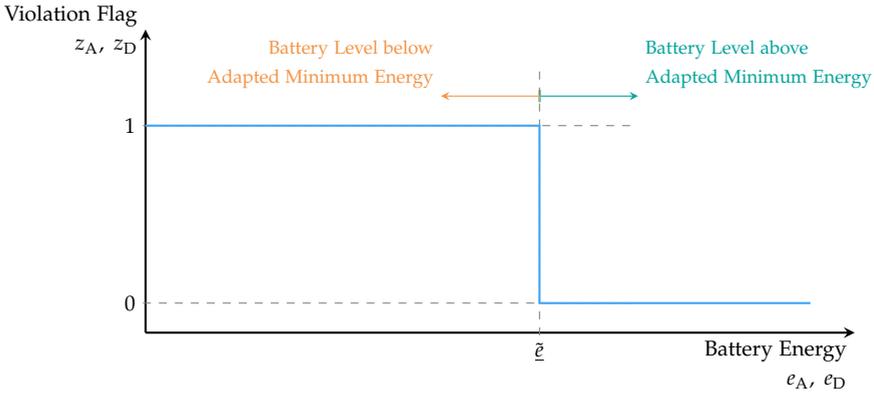
$$z_{A,v} = \begin{cases} 1, & \text{if } 0 \leq e_{A,v} < \tilde{e}_s, \\ 0, & \text{if } \tilde{e}_s \leq e_{A,v} \leq \bar{e}, \end{cases} \quad (5.12)$$

and

$$z_{D,v} = \begin{cases} 1, & \text{if } 0 \leq e_{D,v} < \tilde{e}_s, \\ 0, & \text{if } \tilde{e}_s \leq e_{D,v} \leq \bar{e}, \end{cases} \quad (5.13)$$

must hold for all vertices  $v \in \mathcal{V}$ . For the sake of comprehensibility, Figure 5.4 shows the dependency of the state variables  $e_{A,v}$  or  $e_{D,v}$  and the corresponding value for  $z_{A,v}$  or  $z_{D,v}$ , respectively. In order to establish a link between the state variables  $e_{A,v}$ ,  $e_{D,v}$  and the violation flags  $z_{A,v}$ ,  $z_{D,v}$ , the constraints

$$\tilde{e}_s \cdot (1 - z_{A,v}) \leq e_{A,v} \leq z_{A,v} \cdot (\tilde{e}_s - \bar{e}) + \bar{e}, \quad \forall v \in \mathcal{V}_s, \forall s \in \mathcal{S}, \quad (5.14)$$



**Figure 5.4:** Dependency of battery energy ( $e_A$  or  $e_D$ ) and the corresponding violation flags ( $z_A$  or  $z_D$ ) to explain integer reformulation of chance constraints (5.2i) and (5.2j).

and

$$\tilde{e}_s \cdot (1 - z_{D,v}) \leq e_{D,v} \leq z_{D,v} \cdot (\tilde{e}_s - \bar{e}) + \bar{e}, \quad \forall v \in \mathcal{V}_s, \forall s \in \mathcal{S}, \quad (5.15)$$

are formulated, which further constrain  $e_{A,v}$  and  $e_{D,v}$  beyond the global energy limits (5.2e) and (5.2f), respectively.

By introducing  $z_{A,v}$  and  $z_{D,v}$ , each minimum energy violation, i.e., each  $z_{A/D,v} = 1$ , can be weighted with the estimated occurrence probability  $\mathbb{P}(v)$  of the associated vertex  $v$ . To ensure the user's individual security level  $\varepsilon_v$  at all times, accordingly, the conditions

$$z_{A,v} \cdot \mathbb{P}(v) \leq \varepsilon_v, \quad \forall v \in \mathcal{V}, \quad (5.16)$$

and

$$z_{D,v} \cdot \mathbb{P}(v) \leq \varepsilon_v, \quad \forall v \in \mathcal{V}, \quad (5.17)$$

must hold.

## 5.2.5 Reformulated Optimization Problem

Taking into account all explanations described in Section 5.2.1 to Section 5.2.4, the original optimization problem (5.2) can be reformulated as

$$\min_{\substack{E_v \in \mathbb{R}, \\ C_v, z_{A,v}, z_{D,v} \in \mathbb{Z}, \\ \forall v \in \mathcal{V}}} \sum_{\forall v \in \mathcal{V}} \mathbb{P}(v) \cdot \left( C_v \cdot (\epsilon_{C,v} + \phi_w \cdot w_v) + \frac{\epsilon_v \cdot E_v}{\eta_{C,v}} \right) \quad (5.18a)$$

subject to

$$0 \leq C_v \leq 1, \quad C_v \in \mathbb{Z}, \forall v \in \mathcal{V}, \quad (5.18b)$$

$$\underline{E}_v \cdot C_v \leq E_v \leq \bar{E}_v \cdot C_v, \quad E_v \in \mathbb{R}, \forall v \in \mathcal{V}, \quad (5.18c)$$

$$\underline{e} \leq e_{A,v} \leq \bar{e}, \quad \forall v \in \mathcal{V}, \quad (5.18d)$$

$$\underline{e} \leq e_{D,v} \leq \bar{e}, \quad \forall v \in \mathcal{V}, \quad (5.18e)$$

$$e_{D,v} = e_{A,v} + E_v, \quad \forall v \in \mathcal{V}_d, \forall d \in \mathcal{D}, \quad (5.18f)$$

$$e_{A,v} = e_{D,0} - \sum_{\substack{\forall i, j \in \\ \{\mathcal{V}_d | i \prec j \preceq v\}}} \Gamma_{(i,j)} + \sum_{\substack{\forall k \in \\ \{\mathcal{V}_d | k \prec v\}}} E_k, \quad \forall v \in \mathcal{V}_d, \forall d \in \mathcal{D}, \quad (5.18g)$$

$$0 \leq z_{A,v} \leq 1, \quad z_{A,v} \in \mathbb{Z}, \forall v \in \mathcal{V}, \quad (5.18h)$$

$$0 \leq z_{D,v} \leq 1, \quad z_{D,v} \in \mathbb{Z}, \forall v \in \mathcal{V}, \quad (5.18i)$$

$$\underline{\tilde{e}}_s \cdot (1 - z_{A,v}) \leq e_{A,v} \leq z_{A,v} \cdot (\underline{\tilde{e}}_s - \bar{e}) + \bar{e}, \quad \forall v \in \mathcal{V}_s, \forall s \in \mathcal{S}, \quad (5.18j)$$

$$\underline{\tilde{e}}_s \cdot (1 - z_{D,v}) \leq e_{D,v} \leq z_{D,v} \cdot (\underline{\tilde{e}}_s - \bar{e}) + \bar{e}, \quad \forall v \in \mathcal{V}_s, \forall s \in \mathcal{S}, \quad (5.18k)$$

$$z_{A,v} \cdot \mathbb{P}(v) \leq \epsilon_v, \quad \forall v \in \mathcal{V}, \quad (5.18l)$$

$$z_{D,v} \cdot \mathbb{P}(v) \leq \epsilon_v, \quad \forall v \in \mathcal{V}. \quad (5.18m)$$

The major changes in (5.18) compared with (5.2) are briefly summarized here. The cost components  $J_C$  and  $J_E$  in (5.2a) are replaced with their definitions (5.3) and (5.4), respectively. Furthermore, the expected value is resolved by multiplying the costs arising for each vertex  $v$  with its occurrence probability  $\mathbb{P}(v)$  in (5.18a).

The limits (5.18b) of the integer decision variable  $C_v$  indicating a plug-in of the EV remain unchanged as in (5.2c). As the condition (5.2b) to limit the charged energy directly holds if (5.2d) is satisfied, both are combined to become (5.18c).

Both state variables  $e_{A,v}$  and  $e_{D,v}$  are still globally limited by (5.18d) and (5.18e). Alike, the transition constraints (5.18f) for charging and (5.18g) are unchanged.

To achieve computational tractability of the chance constraints (5.2i) and (5.2j), the integer variables  $z_{A,v}$  and  $z_{D,v}$  indicating a violation of the adapted minimum energy  $\tilde{\epsilon}_s$  are added; they are limited by (5.18h) and (5.18i) to have binary values only. In order to link the state variables  $e_{A,v}$  and  $e_{D,v}$  to the violation flags  $z_{A,v}$  and  $z_{D,v}$ , (5.18j) and (5.18k) are introduced. Furthermore, (5.18l) and (5.18m) ensure the user's individual security level  $\epsilon_v$ , i.e., the chance of undercutting the adapted minimum energy  $\tilde{\epsilon}_s$ , to be maintained for all vertices  $v \in \mathcal{V}$ .

## 5.3 Simulation Setup to Validate Charging Scheduling

This section describes the approach to proof functionality of the charging scheduling scheme outlined in the previous [Section 5.1](#) and [Section 5.2](#). In particular, the target is to test the system’s robustness and quantify benefits of the resulting charging schedule in comparison with alternative strategies. For this purpose, a MONTE CARLO simulation is set up in which 10 000 charging scheduling tasks are run with varying parameters and randomly choosing a scenario to realize.

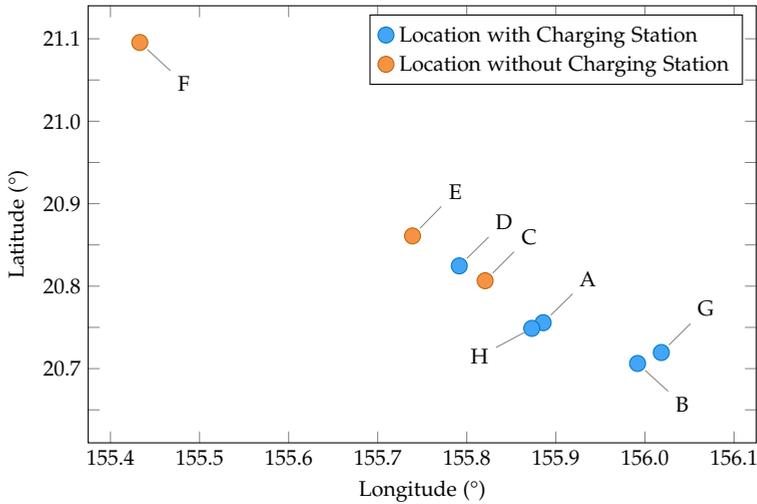
### 5.3.1 Parameter Setting for Monte Carlo Simulation

For the scenario generation (see [Section 5.1](#)), the synthetic set of known locations  $\mathcal{L}_{\text{syn}}$  as presented in [Section 4.1.2](#) is taken as a basis. [Table 5.1](#) presents location-specific properties for each known location, which are time-invariant parameters to the charging scheduling scheme. These comprise the walking distance  $w$  between known location and the closest charging station, and its mean electricity price  $\epsilon$ , both used in the objective ([5.18a](#)). Furthermore, the maximum available power  $\bar{p}$  for each charging station is provided to determine the upper energy bound  $\bar{E}$ , cf., ([5.6](#)) and ([5.18c](#)). In addition to their semantic name, all locations are given a capital

**Table 5.1:** Information on known locations  $l \in \mathcal{L}_{\text{syn}}$  (aliased with capital letters A to H) comprising the walking distance  $w$  between known location and the closest charging station, its mean electricity price  $\epsilon$ , and maximum available charging power  $\bar{p}$ ; dashes indicated unavailable charging opportunities at a location; used for validating the charging scheduling scheme (see [Section 5.1](#) and [Section 5.2](#)).

Location		Walking Distance $w$	Electricity Price $\epsilon$	Max. Charging Power $\bar{p}$
Home	A	5 m	0.23 €/kWh	3.0 kW
Work	B	500 m	0.43 €/kWh	11.0 kW
Orchestra	C	-	-	-
Parents	D	25 m	0.30 €/kWh	3.0 kW
Workout	E	-	-	-
Backoffice	F	-	-	-
Swimming	G	150 m	0.34 €/kWh	7.4 kW
Grocery	H	20 m	0.79 €/kWh	250.0 kW

letter alias to avoid ambiguous nomenclature of the resulting scenario tree components (see [Section 5.1](#)), e.g., vertices. For better orientation, [Figure 5.5](#) shows the (anonymized) geodetic positions of locations A to H.



**Figure 5.5:** Map section with position of known locations A to H based on synthetic data  $\mathcal{J}_{\text{syn}}$  (see Section 4.1.2); blue circles: locations with close-by charging station, orange circles: locations without charging opportunity.

Furthermore, parameters for trips in between all pairs of two known locations (i.e., edges of the scenario tree) are determined by means of an external routing service [290] together with a consumption model, e.g., as described in Section 3.1. Figure 5.6 shows the obtained values for the travel time  $\tau$  in between two known locations.

		End Location							
		A	B	C	D	E	F	G	H
Start Location	A		0:22	0:19	0:24	0:30	1:15	0:23	0:06
	B	0:25		0:42	0:48	0:50	1:28	0:09	0:21
	C	0:18	0:42		0:06	0:25	1:05	0:35	0:17
	D	0:25	0:47	0:07		0:14	1:00	0:39	0:22
	E	0:30	0:50	0:25	0:15		1:05	0:48	0:27
	F	1:20	1:35	1:05	1:05	1:08		1:25	1:17
	G	0:22	0:10	0:53	0:42	0:45	1:23		0:19
	H	0:05	0:22	0:19	0:24	0:26	1:15	0:18	

**Figure 5.6:** Matrix of travel times  $\tau$ , i.e., the time required to travel from a (known) start location to another (known) end location (see Table 5.1), estimated using an external routing service [290]; note the format *hours:minutes*.

Similarly, Figure 5.7 shows the energy consumption  $\Gamma$ .

		End Location							
		A	B	C	D	E	F	G	H
Start Location	A		7.0	4.2	6.0	8.8	38.6	8.4	0.8
	B	6.8		8.8	17.0	19.8	44.2	1.8	5.4
	C	5.4	10.4		2.4	5.0	30.2	18.2	4.8
	D	6.4	17.6	1.4		3.0	28.8	17.2	5.6
	E	11.2	22.6	6.4	5.2		21.6	22.2	10.6
	F	46.6	50.8	32.4	30.2	25.8		52.6	47.6
	G	7.2	1.0	16.2	16.2	19.0	43.2		6.4
	H	1.4	6.6	4.0	5.8	8.6	43.2	8.2	

**Figure 5.7:** Matrix of energy consumption  $\Gamma$ , i.e., the energy required to travel from a (known) start location to another (known) end location (see Table 5.1), estimated using an external routing service [290] together with a model as in Section 3.1; values given in kWh.

Note that both matrices are not symmetrical, as travel times and energy consumption depend on environmental factors such as road topology and traffic conditions. Taking a trip in between two locations may therefore consume more or less time and energy in one direction compared with the other. For instance, a trip from locations C to B consumes 10.4 kWh while the opposite direction only consumes 8.8 kWh due to sloping terrain; furthermore, travel times are identical for both directions in this example.

Furthermore, to estimate the transition probability  $\mathbb{P}_t$  a next place model as described in Section 4.3 is consulted. The probability  $\mathbb{P}_t$  is therefore varying for different conditions such as location and time; yet, within each simulation run, the values of  $\mathbb{P}_t$  are constant. In a similar manner, the model to estimate parking duration (see Section 4.4) is used to determine the available time window  $[t_{v,0}, t_{v,\tau_p,v}]$  of each vertex  $v \in \mathcal{V}$ .

Beyond these previously described parameters, Table 5.2 contains all remaining parameters with their corresponding values, which are constant for all locations and simulation runs. For the (maximum) expected energy consumption of trips to unknown locations, the 95%-quantile energy consumption (19.69 kWh) of all trips to unknown locations in the real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1) is used as orientation. Conservatively, the additional energy consumption is set to  $\Gamma_{l \notin \mathcal{L}} = 20$  kWh; cf., (5.18j), (5.18k) and (5.11).

**Table 5.2:** Globally constant parameters and their selected values for the MONTE CARLO simulation to validate charging scheduling.

Parameter		Selected Value
Upper Energy Bound	$\bar{e}$	80 kWh
Additional Energy Consumption	$\Gamma_{l \notin \mathcal{L}}$	20 kWh
Minimum Charging Power	$p_v$	0 kW
Weight of Walking Distance	$\phi_w$	0.01 €/m
Constant Costs for Charging	$\epsilon_{C,v}$	0€

To test the charging scheduling scheme in a variety of situations, all parameters listed in Table 5.3 are furthermore varied within a specified range of values. Here,

**Table 5.3:** Parameters that are varied in a specified range of values for the MONTE CARLO simulation to validate charging scheduling.

Parameter		Tested Values
Number of Trips		2,3,...,10
Pruning Probability		1%, 2%, ..., 10%
Lower energy bound	$\underline{e}$	8 kWh, 16 kWh, 24 kWh
User's Security Level	$\epsilon_v$	0%, 1%, ..., 10%

the maximum number of trips that can be taken within the scheduling time horizon is used as termination criterion for the scenario generation. Furthermore, the pruning probability specifies the minimum occurrence probability for a scenario  $d$  to be considered in the optimization scheme.

To additionally reproduce stochastic influences, e.g., uncertain user behavior, the following parameters are randomly chosen for each simulation run:

- **Time of Calculation:** The point in time at which the charging scheduling module is triggered, varied over the time range of one week in a resolution of one minute.
- **Start Location:** The location from which the scenario generation originates, i.e., the root vertex of the scenario tree (see Section 5.1), randomly chosen from the set of known locations  $\mathcal{L}$ , see also Table 5.1.
- **Initial Battery Energy  $e_{D,0}$ :** Congruent to the SOC at the time of calculation, randomly chosen from a range  $e_{D,0} \in [48 \text{ kWh}, 80 \text{ kWh}]$ .

- **Scenario to Realize:** The sequence of locations the user actually visits throughout the scheduling time horizon, randomly chosen considering the occurrence probabilities  $P(v)$  associated with each vertex  $v \in \mathcal{V}$ .

To finally run the simulation, all components of the charging scheduling module are implemented in *Python* [256]. For the mixed-integer programming (5.18), the optimization framework *CasADi* [291] is used together with the open-source solver *BONMIN* [292].

### 5.3.2 Alternative Strategies for Charging Scheduling

After setting up and running a MONTE CARLO simulation of the charging scheduling scheme (see Section 5.3.1), the results are benchmarked. To this end, two alternative strategies to schedule charging are defined:

- **Always Charge Strategy:** Whenever a charging station is available close to the visited location, the EV will be charged with as much energy as possible. This behavior is expected from users who just recently switched from an ICEV to an EV and suffer from range anxiety, i.e., the EV's range seems too small for the user's mobility demand, and thus, they charge as often and much energy as possible.
- **Minimum Charge Strategy:** The EV will be charged when arriving at a location, if the battery level  $e_A$  is less than the minimum energy level  $\underline{e}$ . This behavior is expected from experienced EV users who almost fully utilize their EV's driving range before recharging.

These strategies represent two opposing types of users between which the majority of actual users are assumed to be, see also Section 1.1.3. Future work could also explore a strategy of charging the EV whenever it is parked at home or if the battery level falls below the lower bound. Comparing the proposed charging scheduling scheme with alternative strategies allows to quantify the results and provide a benchmark for further research. To represent different perspectives on charging scheduling, the following four metrics are consulted:

- The number of plug-in events  $C$  per scheduling run as a measure for user comfort, i.e., the more plug-in events a strategy has, the less comfortable it is assumed to be perceived.

- The mean electricity price to quantify the average energy cost savings, i.e., the sum of all energy costs  $J_E$  divided by the sum of charged energy  $E$  throughout the time horizon per scheduling run.
- The number of adapted minimum energy violations  $z$  per scheduling run as a security measure to stay mobile, i.e., how reliably a dynamic battery reserve for unexpected trips can be maintained.
- Calendar battery aging as a measure of sustainability, i.e., the temporal decay of battery capacity within one scheduling run, calculated by means of (3.23) and summing up the increments.<sup>8</sup>

These metrics are calculated based on the charging schedules obtained from both the proposed scheduling scheme and the alternative strategies. Then, the absolute and relative difference to the optimized strategy are evaluated.

---

<sup>8</sup>Note that cyclic battery aging solely depends on the amount of charged energy, see (3.22). If more energy is charged in one scheduling run (i.e., more cyclic aging occurs), the energy does not have to be charged in a subsequent run (i.e., less cyclic aging occurs); evaluating cyclic aging is therefore unreasonable in this context.

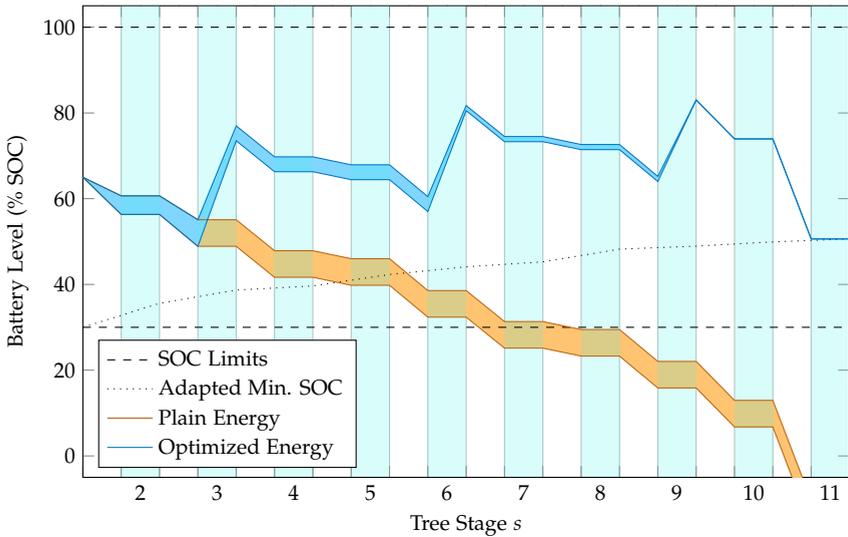
## 5.4 Results of Validating Charging Scheduling

The MONTE CARLO simulation described in Section 5.3 attempted to perform a total of 10 000 runs. However, only 6287 successful simulation runs (i.e., 62.9%) are observed. In all remaining cases, the optimization problem (5.18) could not be solved successfully. In particular, simulation runs with a large number of trips and scenarios are prone to fail. A possible reason for this observation can be the applied optimization solver, whose preprocessing heuristics might not be performant enough to cope with optimization problems of this size. In this situation, a simple backup strategy could be used, e.g., charging the EV when arriving at a location with a battery level of less than 50% SOC. However, future work should still consider testing more advanced solvers to increase the share of successfully solved problems. Alternatively, a less complex representation of the problem should be examined; e.g., locations without any charging station close by—for which the charged energy is always equal to zero—could be removed as decision variables. All evaluations in the remainder of this section are based only on the successful simulation runs.

For a real-world implementation of the proposed charging scheduling scheme, the algorithms for scenario generation (Section 5.1) and optimization (Section 5.2) need to be deployed as cloud-based microservices (see also Section 2.3). Furthermore, suitable programming interfaces are required to connect these microservices with both the EV and a user interface such as a smart phone application. Via the latter, EV users are able to track the scheduled charging stops and also manually request a recalculation.

### 5.4.1 Robustness through Adapted Minimum Energy

As described in Section 5.2.4, the obtained charging schedule incorporates an additional battery reserve, which is derived from the estimated residual probability  $\mathbb{P}_r(s)$ , i.e., the chance of the user unexpectedly visiting an unknown location. In Figure 5.8, this battery reserve, i.e., the adapted minimum energy  $\tilde{\epsilon}_s$  is shown as dotted line for one exemplary simulation run over the stages of the corresponding scenario tree (see Section 5.1). As defined in (5.18h) to (5.18m), the battery level is only allowed to undercut  $\tilde{\epsilon}_s$  in a specified portion of cases, i.e., the user's security level  $\epsilon_v$ . It can be seen that in Figure 5.8 all optimized SOC profiles (in blue) exactly meet  $\tilde{\epsilon}_s$  by the end of the scheduling time horizon; an expected result, as  $\epsilon_v = 0\%$  for the shown example. Given a suitable determination of  $\tilde{\epsilon}_s$ , e.g., as in (5.11), the optimized strategy therefore ensures to reach any destination within the maximum



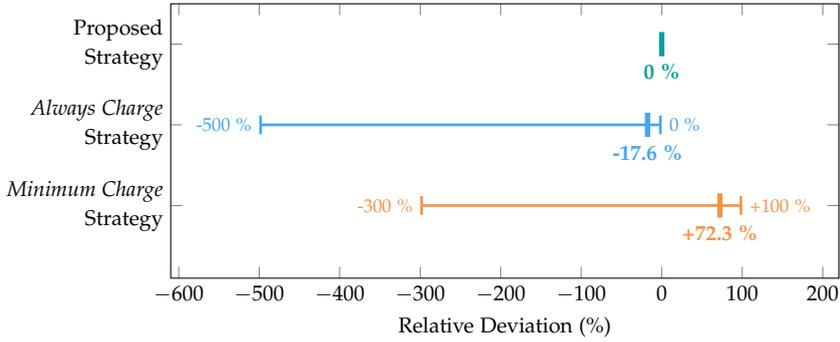
**Figure 5.8:** Exemplary charging scheduling run with sets of SOC profiles (without charging in orange, optimized in blue) over tree stages  $s$  (see Section 5.1), adapted minimum energy  $\tilde{\epsilon}_s$  drawn as dotted line ( $\epsilon_v = 0\%$ ), turquoise shaded areas represent parking phases of the EV.

range of the battery whenever desired. Note that the number of vertices grows exponentially with each scenario tree stage (see Section 5.1). As a consequence, both residual probability  $\mathbb{P}_r(s)$  and adapted minimum energy  $\tilde{\epsilon}_s$  increase throughout the scheduling time horizon. Thus,  $\tilde{\epsilon}_s$  functions as a terminal penalty to avoid large deviations of a subsequent charging schedule when recalculating prior to the end of the current time horizon, see e.g., also [293,294].

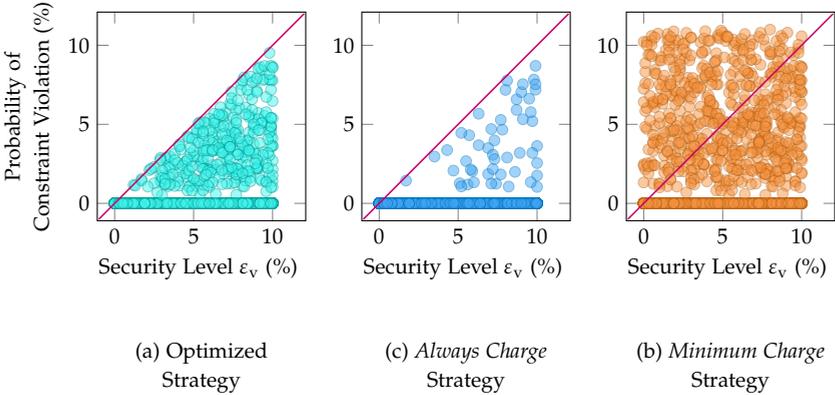
As shown in Figure 5.9, following one of the alternative strategies, however, may reveal different results in terms of robustness. Especially the *Minimum Charge* strategy yields 72.3% more cases in which the adapted minimum energy  $\tilde{\epsilon}_s$  is undercut compared with the optimized strategy. In 10.7% of all cases, even the global energy limit  $\bar{\epsilon}$  is violated if following the *Minimum Charge* strategy. As for some cases also the optimized strategy may undercut  $\tilde{\epsilon}_s$ , i.e., if  $\epsilon_v > 0\%$ , the *Always Charge* strategy on average exhibits 17.6% less violations.

Figure 5.10 examines the probability of violating the adapted lower energy limit  $\tilde{\epsilon}_s$  with respect to the specified security level  $\epsilon_v$ .

Here, the optimized strategy (a) is shown next to the alternative strategies *Always Charge* (b) and *Minimum Charge* (c); each marker represents one simulation run.



**Figure 5.9:** Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing minimum battery energy constraint violations, i.e.,  $e < \bar{e}_s$ ; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy.



**Figure 5.10:** Comparison of given security level  $\epsilon_v$  with the observed probability of violating the adapted lower energy limit  $\bar{e}_s$  for the optimized charging strategy (a) and the alternative strategies *Always Charge* (b) and *Minimum Charge* (c).

For the optimized and *Always Charge* strategy, the violation probability is always less or equal to the set security level  $\varepsilon_v$  (see the pink, diagonal line). To obtain a cost-optimal solution, however, the optimized strategy utilizes admitted constraint violations, i.e.,  $\underline{e} \leq e < \bar{e}$ , more often compared with *Always Charge*. For the *Minimum Charge* strategy, a large number of simulation runs can be observed in which the violation probability exceeds the value of  $\varepsilon_v$ . It can therefore be concluded that the optimized strategy exhibits significant benefits in terms of robust battery reserve compared with the *Minimum Charge* strategy. This is especially helpful for anxious EV users, who suffer from *range anxiety*, see also Section 1.1.3. Compared with *Always Charge*, the optimized strategy may exhibit more constraint violations; yet, the specified security level is always respected.

## 5.4.2 Benefits over Alternative Strategies for Charging Scheduling

Besides the robustness mentioned in Section 5.4.1, the scheduling scheme (5.18) holds further advantages. To demonstrate, Figure 5.11 shows the SOC profiles of both the optimized and the alternative scheduling strategies based on one exemplary simulation run. The benefits of the optimized strategy over the alternative strategies are threefold, see Table 5.4.

**Table 5.4:** Comparison of alternative charging strategies (see Section 5.3.2) with the proposed scheduling scheme (Section 5.2) in terms of reducing the number of plug-in events, battery calendar aging, and the mean electricity price; negative values indicate a benefit over the proposed strategy.

Strategy	Plug-In Events	Calendar Aging	Mean Electricity Price
<i>Always Charge</i>	88.8%	32.2%	78.4%
<i>Minimum Charge</i>	-13.1%	-2.9%	4.3%

First, the number of plug-in events is significantly reduced compared with the *Always Charge* strategy—an aspect that increases the user comfort of EV charging. In the particular example shown in Figure 5.11, the EV needs to be plugged in nine times when following *Always Charge*; the optimized result only exhibits three plug-in events. Figure 5.12 visualizes the evaluation of all simulation runs. Here, the optimized strategy on average yields 2.86 less plug-in events compared with *Always Charge* (blue), i.e., a 88.8% reduction (approx. 50% reduction is described in [271]). Compared with the *Minimum Charge* strategy (orange), on the contrary, the EV requires 0.26 more plug-in events following the optimized strategy (13.1% increase). However, note that complying with *Minimum Charge* exposes the risk of low battery (see Section 5.4.1), which in turn may cause additional, unplanned charging stops

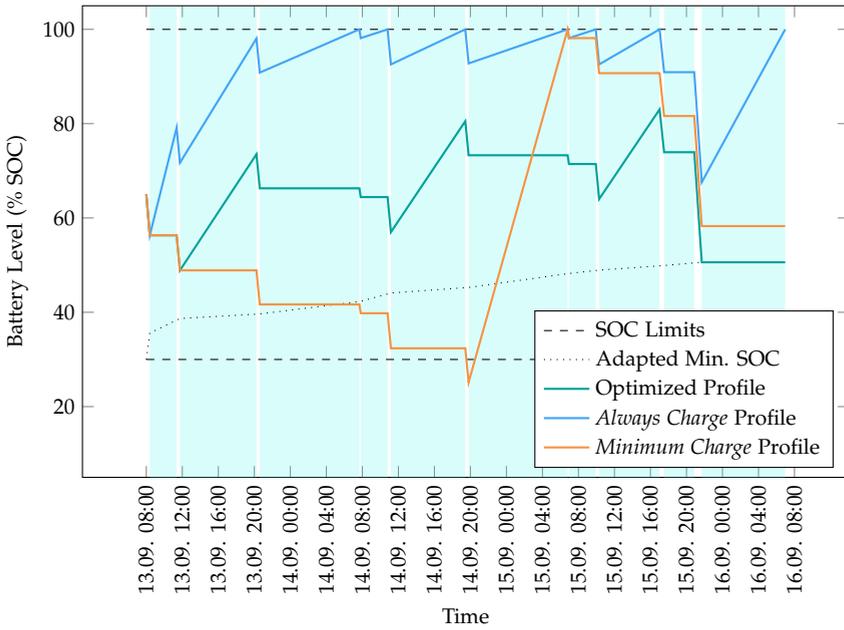


Figure 5.11: SOC profiles for proposed (green) and alternative charging scheduling strategies (blue, orange) for one simulation run over the scheduling time horizon, adapted minimum energy  $\bar{\epsilon}_s$  drawn as dotted line ( $\epsilon_v = 0\%$ ), turquoise shaded areas represent parking phases of the EV.

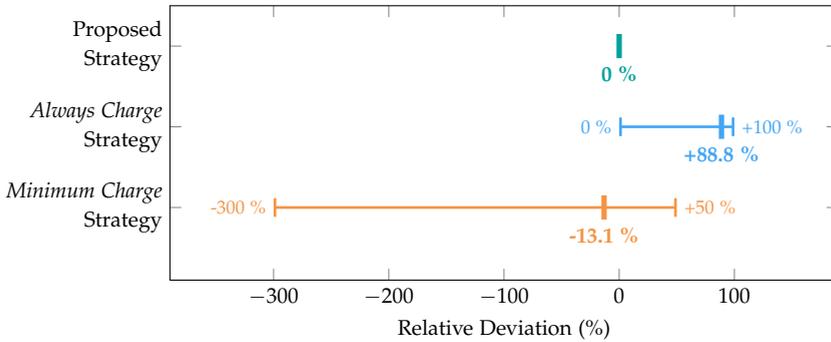
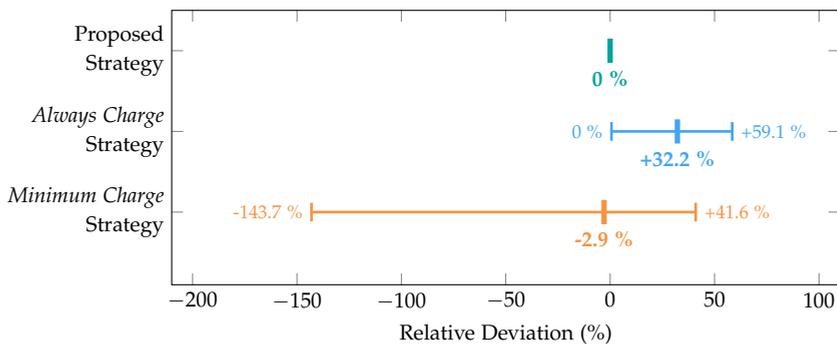


Figure 5.12: Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing the number of plug-in events  $C$ ; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy.

along the way. Hence, a trade-off between the robustness and comfort of charging scheduling exists, which may require user-individual adaptations of (5.18) in the future.

Second, the mean SOC level of the battery is lowered compared with *Always Charge*, which in turn significantly reduces calendar battery aging, according to (3.23); this aspect enhances the EV's sustainability, as the battery lifetime is prolonged. In Figure 5.11, the optimized battery SOC—both during driving and charging—is always kept in a range between approx. 40 % SOC to 80 % SOC; the SOC profile of *Always Charge*, on the contrary, is close to 100 % for the majority of time. The evaluation of all simulation runs is visualized in Figure 5.13. Compared with *Always Charge*,

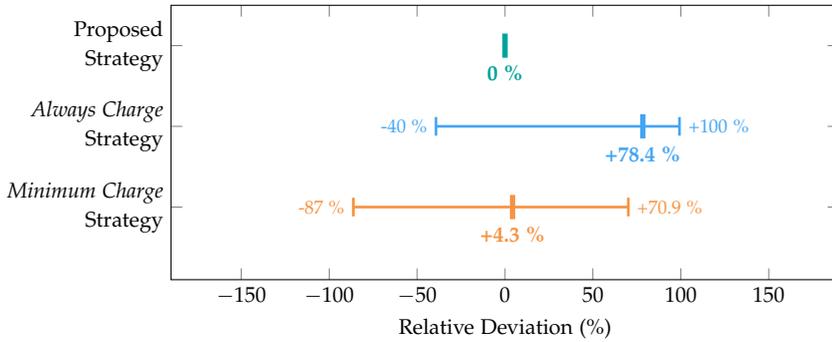


**Figure 5.13:** Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing the calendar aging; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy.

the proposed scheme on average yields 32.2 % less calendar aging. For the battery of a typically used EV, this would equal a saving of approx. 81 € per year in terms of battery aging costs.<sup>9</sup> The *Minimum Charge* strategy on average exhibits 2.9 % less calendar aging (equals approx. 7 € per year) compared with the optimized strategy. Similar to the aspect of comfort, a sufficient battery reserve is not guaranteed with *Minimum Charge*. Hence, an additional trade-off between sustainability and robustness exists, which may prospectively be parameterized according to the user's preferences.

<sup>9</sup>Assuming a yearly driving distance of 10000 km and a yearly battery depreciation of 760 € due to aging, of which approx. one third results from calendar aging, cf., Section 3.2.3, [125] and [92].

Third, the mean price for charging electricity is reduced compared with both alternative strategies *Always Charge* and *Minimum Charge*. As a consequence, the overall EV operating cost decreases. Figure 5.14 shows the relative savings over the alternative strategies considering all simulation runs. Especially in comparison with



**Figure 5.14:** Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing the mean price for electricity; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy.

*Always Charge*, a significant price reduction of 78.4% can be observed on average. For a typically used EV, this would equal electricity cost savings of approx. 737€ per year.<sup>10</sup> A traceable result considering the fact that *Always Charge* consumes energy at every charging opportunity; thus, e.g., also at the grocery location H (see Table 5.1), which offers an electricity price of 0.79€/kWh. Following *Minimum Charge*, still 4.3% less energy costs can be achieved. This price difference can be explained by the fact that the proposed optimization scheme incorporates all electricity prices at the time of calculation. In this way, charging locations with low electricity prices can be anticipated when generating a charging schedule. The *Minimum Charge* strategy, on the other hand, may in some cases force the user to charge electricity at locations with high prices, as otherwise the battery would run empty.

<sup>10</sup>Assuming a yearly driving distance of 10000km and an average energy consumption of 0.293 kWh/km, see Section 3.1.1.

## 5.5 Conclusion

Chapter 5 deals with the scheduling of charging events. In particular, the aim is to choose the most suitable charging opportunities when the EV is parked anyway. To this end, the outputs (set of known locations, predicted next place, and estimated parking duration) of the **mobility prediction** models as described in Chapter 4 are consulted.

A graph-based scenario generation (see Section 5.1) allows to quantify the uncertainty inherent in user actions via scenarios, i.e., sequences of trips and visits of known locations. Based on this, a mixed-integer programming is engineered (see Section 5.2) the solution of which yields a charging schedule providing information of where and how much energy to charge. As minimization objective, the (notional) cost for plugging in the EV at a charging station and for the charged energy is used. Furthermore, an additional battery reserve is calculated depending on the estimated probability of unexpected events to happen, and incorporated via chance constraints; to enable computational tractability, a integer reformulation of the chance constraints is conducted afterwards. To validate the proposed scheme, a MONTE CARLO simulation is run while varying input parameters (see Section 5.3). In addition, two alternative charging strategies representing corner cases of EV users are designed to benchmark the results according to four metrics evaluating robustness, comfort, sustainability, and monetary cost.

The results show that approx. 38% of all attempted simulation runs fail due to optimization solving issues. Accordingly, the applied solver might be inadequate for mixed-integer optimization problems of this size. Future work should evaluate other solvers and/or examine a simpler formulation of the decision problem.

The proposed charging scheduling strategy proves to be capable of providing sufficient battery reserve to accomplish all daily (expected and unexpected) mobility requirements. Critical situations within the scheduling time horizon in which users following an alternative strategy may end up with an empty battery, can be anticipated and successfully handled.

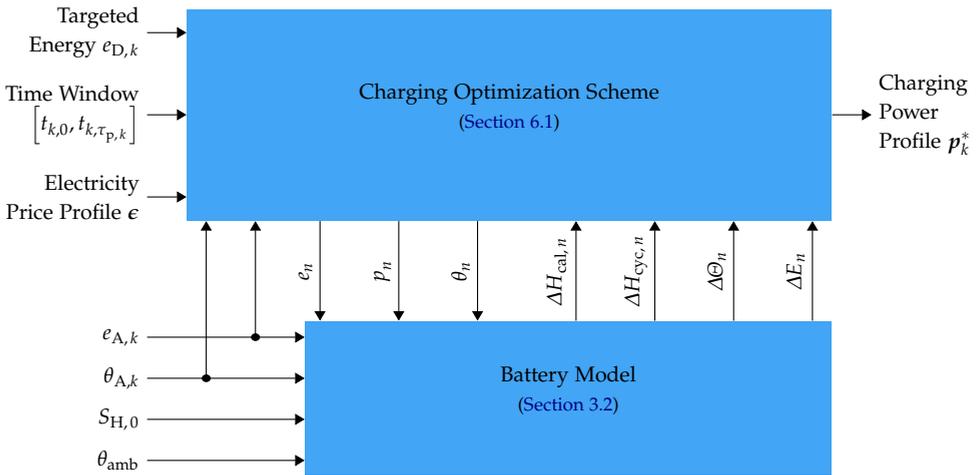
To the end of comfort, the number of plug-in events can be reduced by up to almost 88% compared with an alternative strategy anxious users would follow. Furthermore, the proposed scheme can anticipate future mobility demand, derive a corresponding energy demand, and accordingly adapt the amount of charged energy. Consequently, the mean SOC level is decreased, which in turn reduces calendar aging of the EV battery up to one third, and thus promotes a sustainable operation of the EV. Finally, anticipating charging opportunities with low electricity prices

allows to reduce the mean electricity price of charged energy compared with all alternative strategies; in some cases, energy cost savings exceed 75%.

Future work should extend the proposed scheme to include bidirectional charging. In this way, the impact and potential revenue streams from the provision of V2G services (see [Section 1.1.1.2](#)) can also be examined. Furthermore, the system needs to be embedded into a real-world smart charging environment that comprises integration of multiple users, end devices, EVs, and real-time data sources. Testing charging scheduling in such a setup can help to establish a meaningful setting of (user-individual) parameters and discover possible issues.

# 6 Charging Optimization

The time EVs are parked usually exceeds the time required to fully charge their battery, see e.g., Section 1.1.3.1. Most charging events therefore follow a standard procedure of charging the EV at maximum available power upon arrival at a charging station. Then, the fully charged EV parks until the next trip, leaving both time and energy flexibility of the charging process unused. Anxious users in particular also tend to charge their vehicles more often than necessary (see also Section 5.3.2). Figure 6.1 shows the **charging optimization** scheme presented in this chapter, which aims to leverage each charging event’s flexibility to improve efficiency, sustainability, and operating cost.



**Figure 6.1:** Schematic block diagram of the charging optimization scheme interacting with the EV battery model (see Section 3.2); both targeted battery energy  $e_{D,k}$  and available time window  $[t_{k,0}, t_{k,\tau_{p,k}}]$  are taken from the charging schedule generated with a scheme as described in Chapter 5.

Based on the energy  $e_{A,k}$ , temperature  $\theta_{A,k}$  and SOH  $S_{H,0}$  of the battery upon arrival at charging event  $k$ , a time-discrete charging power profile  $p_k^*$  is calculated for the time window  $[t_{k,0}, t_{k,\tau_{p,k}}]$ , such that the target battery energy  $e_{D,k}$  is met upon departure. To achieve this, a non-linear optimization problem is designed in Section 6.1, which takes a dynamic electricity price  $\epsilon$  into account; an approach similarly known from residential electricity consumers, see e.g., [295,296]. Using the detailed battery model as described in Section 3.2 allows to incorporate the electrical, thermal, and degradation characteristics of the EV battery.

To evaluate the optimization scheme, a two-part case study is set up in Section 6.2. First, charging optimization is simulated for the conditions of real-world charging events and compared afterwards. Second, a simplified version of the optimization scheme is implemented in a realistic environment and tested with real EVs. Section 6.3 then presents and interprets the major findings of the case study. Note that parts of this chapter are based on the evaluations SCHWENK et al. described in [250].

## 6.1 Charging Optimization Scheme

Automatically calculating an optimal charging power trajectory  $p_k^*$ , first, calls for a suitable problem formulation in a mathematical sense. For this purpose, each charging event  $k$  starting at time  $t_{k,0}$  and ending at time  $t_{k,\tau_{p,k}}$  is divided into

$$N_k = \left\lfloor \frac{t_{k,\tau_{p,k}} - t_{k,0}}{\Delta t} \right\rfloor \in \mathbb{N}, \quad (6.1)$$

equidistant time intervals possessing a duration of  $\Delta t$ . Based on this, the set of time intervals  $\mathcal{N}_k \subset \mathbb{N}$  is defined according to (3.6).<sup>1</sup> Both times  $t_{k,0}$  and  $t_{k,\tau_{p,k}}$  are taken from the previously calculated charging schedule (see Chapter 5). To account for the uncertainty especially inherent in a charging event's end time (see Section 4.4), a conservative approach may be used. Rather than a point estimation of the parking duration  $\tau_p$ , e.g., the estimated 5%-quantile could be used to establish  $t_{k,\tau_{p,k}}$ . Accordingly, only a 5% chance remains that the charging event ends before the corresponding power profile is finished.

To subsequently determine a suitable charging power  $p_n$  for each time interval  $n \in \mathcal{N}_k$ , the following non-linear optimization problem is designed such that

$$\min_{\substack{p_n \in \mathbb{R}, \\ \forall n \in \mathcal{N}_k}} \sum_{\forall n \in \mathcal{N}_k} J_{E,n}(p_n, \epsilon_n) + J_{D,n}(\theta_n, e_n, S_{H,0}) \quad (6.2a)$$

subject to

$$\underline{p}_k \leq p_n \leq \bar{p}_k, \quad p_n \in \mathbb{R}, \forall n \in \mathcal{N}_k, \quad (6.2b)$$

$$\underline{e}_k \leq e_n \leq \bar{e}_k, \quad e_n \in \mathbb{R}, \forall n \in \{0, 1, \dots, N_k\}, \quad (6.2c)$$

$$e_0 = e_{A,k}, \quad (6.2d)$$

$$e_{N_k} = e_{D,k}, \quad (6.2e)$$

$$\underline{\theta}_k \leq \theta_n \leq \bar{\theta}_k, \quad \theta_n \in \mathbb{R}, \forall n \in \{0, 1, \dots, N_k\}, \quad (6.2f)$$

$$\theta_0 = \theta_{A,k}, \quad (6.2g)$$

$$e_{n+1} = e_n + \Delta E_n(e_n, \theta_n, p_n), \quad \forall n \in \mathcal{N}_k, \quad (6.2h)$$

$$\theta_{n+1} = \theta_n + \Delta \Theta_n(e_n, \theta_{\text{amb}}, p_n), \quad \forall n \in \mathcal{N}_k. \quad (6.2i)$$

<sup>1</sup>The optimization time horizon may be shorter than the actual time window of charging event  $k$ . The end time  $t_{k,N_k}$  of the last interval  $N_k - 1$  may therefore be less than  $t_{k,\tau_{p,k}}$ .

All components of (6.2) are further explained in the following sections.

### 6.1.1 Objective of Charging Optimization

The **charging optimization** scheme (6.2) aims to minimize the sum of energy costs  $J_{E,n}$  and aging costs  $J_{D,n}$  in all time intervals  $n \in \mathcal{N}_k$ , see (6.2a). To consider the costs inherent with charging electric energy, the energy cost function

$$J_{E,n} = \begin{cases} J_{E,n}^+ & \forall p_n \geq 0, \\ J_{E,n}^- & \forall p_n < 0, \end{cases} \quad (6.3)$$

is defined. If the battery is charged in a time interval  $n$ , i.e.,  $p_n \geq 0$ , the energy expenses

$$J_{E,n}^+ = p_n \cdot \Delta t \cdot \epsilon_n, \quad (6.4)$$

apply. In a similar manner, the energy rewards

$$J_{E,n}^- = p_n \cdot \Delta t \cdot \epsilon_n. \quad (6.5)$$

account for energy sold back to the grid if  $p_n < 0$ , i.e., when discharging the EV battery. In both energy cost components,  $\epsilon_n$  corresponds to the electricity price in time interval  $n$ , which is assumed to be deterministic at the time of calculation.<sup>2</sup> If charging the EV at public charging stations, the operator of the station will usually stipulate the values for  $\epsilon_n$ , e.g., in form of an (dynamic) electricity tariff. A sole reduction of electricity costs is thus pursued. Beyond that,  $\epsilon_n$  might be defined by a notional price profile, e.g., corresponding to the share of renewable (and, if charging at home, self-produced) energy in the grid. In this way, the amount of renewable energy charged into the EV can be maximized. This, however, requires to forecast the share of renewable energy for the time window of each charging event, e.g., using an approach as presented by SCHWENK et al. in [297]. Furthermore, an additional price component is conceivable, which is derived from the load of the relevant low voltage distribution grid, as congestion issue may particularly occur in these areas, see also [23, 295, 298]

Besides charging electric energy, degradation of the battery also contributes to the total EV operating cost. For each time interval  $n \in \mathcal{N}_k$ , the increments of cyclic aging  $\Delta H_{\text{cyc},n}$  and the calendar aging  $\Delta H_{\text{cal},n}$  are thus calculated by means of (3.22)

<sup>2</sup>In a real application, the prices for buying and selling electricity might also deviate at the same time, i.e.,  $\epsilon_n$  differs in (6.4) and (6.5); for the sake of simplicity, however, both prices are assumed to be identical in this evaluation.

and (3.23), see also Section 3.2.3. Based on this, the costs for battery degradation (*aging costs*)

$$J_{D,n} = \underbrace{\Delta H_{\text{cyc},n} \cdot \frac{V_{\text{EV}}}{H_{\text{EV}}}}_{J_{D,n}^{\text{cyc}}} + \underbrace{\Delta H_{\text{cal},n} \cdot \frac{V_{\text{EV}}}{H_{\text{EV}}}}_{J_{D,n}^{\text{cal}}}, \quad (6.6)$$

are defined. For the sake of traceability,  $J_{D,n}$  is separated into the cyclic aging costs  $J_{D,n}^{\text{cyc}}$  and the calendar aging costs  $J_{D,n}^{\text{cal}}$ . In (6.6),  $V_{\text{EV}}$  denotes value loss of the EV battery due to a loss of storage capacity  $H_{\text{EV}}$  during its entire automotive application (*first life* of the battery). In particular,  $V_{\text{EV}}$  is the difference of the battery's production price and its residual value in a *second life* market.<sup>3</sup> Note that (6.6) only accounts for battery aging caused during charging events. If using a superordinate scheduling scheme, e.g., as described in Chapter 5, battery aging for trips in between charging events can also be considered. For instance, by determining an optimal target energy  $e_{D,k} < \bar{e}$ , which is adapted to the energy demand of subsequent trips.

## 6.1.2 Decision and State Variables

To obtain an optimal trajectory

$$\mathbf{p}_k^* = \left( p_0^*, p_1^*, \dots, p_{N_k-1}^* \right)^\top \in \mathbb{R}^{N_k}, \quad (6.7)$$

of charging power, a decision variable  $p_n \in \mathbb{R}$  is defined for each time interval  $n \in \mathcal{N}_k$ . Here,  $p_n$  represents the gross charging power consumed from the charging station. As given in (6.2b),  $p_n$  is constrained by the power limitations  $\underline{p}_k$  and  $\bar{p}_k$ . Those are determined e.g., by charging stations or (on-board) power electronics and represent the minimum and maximum available power at charging event  $k$ .

To compute and track the progression of the battery level, i.e., the SOC, throughout a charging event, the state variable  $e_n \in \mathbb{R}$  is defined, which represents the battery energy at time  $t_n, \forall n \in \{0, 1, \dots, N_k\}$ . As given in (6.2c), the energy limitations  $\bar{e}_k$  and  $\underline{e}_k$  apply to  $e_n$ . Their values are determined by physical restrictions, i.e., the battery's maximum storage capacity  $e_{\text{max}}$ , and/or preferences of the EV user, e.g., a minimum SOC level as mobility reserve. Furthermore, the values for  $e_{A,k}$  and  $e_{D,k}$  are taken from the previously calculated charging schedule, see Chapter 5. The

<sup>3</sup>After their automotive application, EV batteries can be used for stationary and low-dynamic applications (*second life*), such as energy buffers for the utility grid, see also [27, 92, 93].

energy  $e_{A,k}$  left in the battery upon arrival at  $k$  defines the initial battery level  $e_0$ , see (6.2d). Similarly, the battery energy  $e_{N_k}$  by the end of the charging process must equal the target energy  $e_{D,k}$ , see (6.2e). Prospectively, (6.2e) could be formulated as additional cost function to allow for some tolerance and thus enable larger energy flexibility. In case of electricity overproduction (low prices), the battery could then be charged beyond the target energy (if  $e_{D,k} < \bar{e}_k$ ). If electricity underproduction occurs (high prices), not reaching  $e_{D,k}$  could be traded off with charging expensive energy; however, only if not limiting the user's mobility.

Similar to battery energy, the state variable  $\theta_n \in \mathbb{R}$  is defined to represent the battery temperature at time  $t_n, \forall n \in \{0, 1, \dots, N_k\}$ . Here, (6.2g) ensures that the initial value  $\theta_0$  corresponds to the battery temperature  $\theta_{A,k}$  upon arrival at  $k$ . At the time of computation, all other values  $\theta_n, \forall n \in \{1, 2, \dots, N_k\}$  are unknown. As (6.2f) states, however, they are constrained within the temperature limits  $\bar{\theta}_k$  and  $\underline{\theta}_k$ , such that a safe operation of the EV battery can be guaranteed, see also Section 1.1.2.

### 6.1.3 Including Dynamics of the Electric Vehicle Battery

The values of both state variables  $e_n$  and  $\theta_n$  (see Section 6.1.2) change within each time interval of a charging process depending on the given conditions. Simulating the progression of the battery state is therefore essential to optimize a future charging process. To represent the electrical behavior of the battery throughout a single time interval  $n$ , the energy transition constraint (6.2h) is formulated. Here, the electrical battery model developed in Section 3.2.1 is consulted. Equivalently, the temperature transition constraint (6.2i) is defined by means of the thermal battery model described in Section 3.2.2.

Note that (6.2) omits the battery SOH progression throughout a charging event. Instead, the SOH  $S_{H,0}$  at the beginning of a charging event serves as a reference for all time steps  $n \in \{0, 1, \dots, N_k\}$ ; a simplification legitimated by the fact that battery aging occurs on significantly larger time scales (years) than charging (hours). The expected marginal difference in the obtained solution  $\mathbf{p}_k^*$  would therefore not justify the increased problem complexity caused by including the SOH progression.

## 6.2 Validation and Case Study

To validate the optimization scheme as described in Section 6.1, a two-part case study is set up. First, the optimization scheme is simulated and evaluated for historical charging events, see Section 6.2.1. Second, Section 6.2.2 presents a real-world implementation in which the scheme is tested with real EVs.

### 6.2.1 Simulation of Historical Charging Events

This part of the case study, in particular, aims to quantify influence factors on EV operating cost, i.e., the sum of aging costs  $J_D$  and energy costs  $J_E$ , when applying charging optimization. For simulating the proposed scheme, the data set  $\mathcal{K}$  of 279 real, unidirectional charging events  $k$  as described in Section 3.2.2.1 is consulted. To ensure a realistic test setup, 45 charging events that have sufficient duration of more than two hours are selected. For each charging event  $k$ , the parameters  $e_{A,k}$ ,  $\theta_{A,k}$ ,  $S_{H,0}$ ,  $t_{k,0}$ ,  $t_{k,\tau_p,k}$ , and  $e_{D,k}$  are set individually. In addition, Table 6.1 outlines further parameters, whose values are set constant for all  $k \in \mathcal{K}$ .

**Table 6.1:** Globally constant parameters and their selected values to simulate charging scheduling.

Parameter		Selected Value
Nominal Battery Energy	$e_{\text{nom}}$	80 kWh
Maximum Battery Energy	$\bar{e}$	80 kWh
Minimum Battery Energy	$\underline{e}$	8 kWh
Maximum Charging Power	$\bar{p}$	50 kW
Minimum Charging Power	$\underline{p}$	-50 kW
Maximum Battery Temperature	$\bar{\theta}$	60 °C
Minimum Battery Temperature	$\underline{\theta}$	-25 °C
Battery First Life Capacity Loss	$H_{\text{EV}}$	20 % SOH
Battery First Life Value Loss	$V_{\text{EV}}$	6080 €

To reproduce a dynamic electricity price  $\epsilon$ , historic hourly market prices of 2018 are used [299]. Then, typical fees and taxes are supplemented to attain a representative retail price level for private customers, see Figure 6.2. The hourly price profiles are then averaged over all workdays and weekends to level out price peaks due to electricity over- or underproduction. In this way, two characteristic price tables

for  $\epsilon$  are obtained (see Figure 6.3) to evaluate the average profitability of charging optimization.<sup>4</sup>

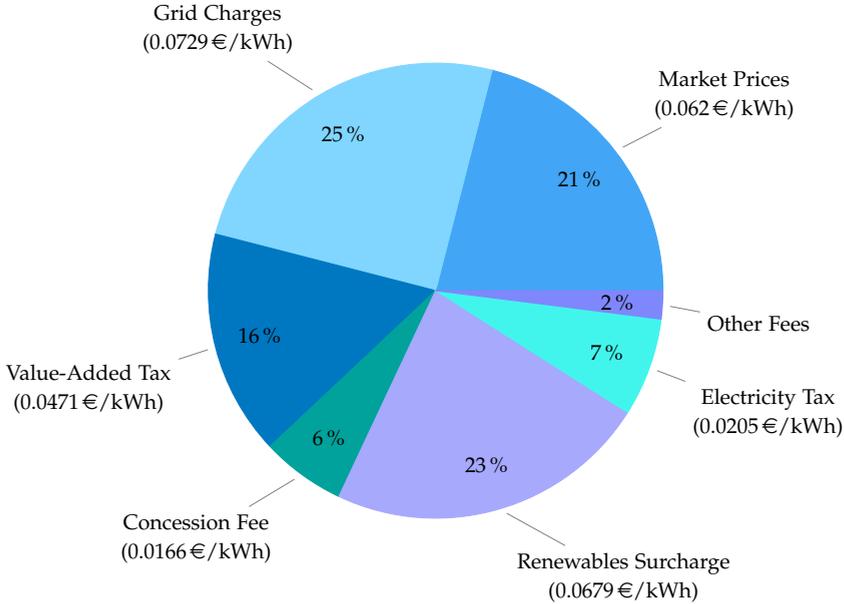


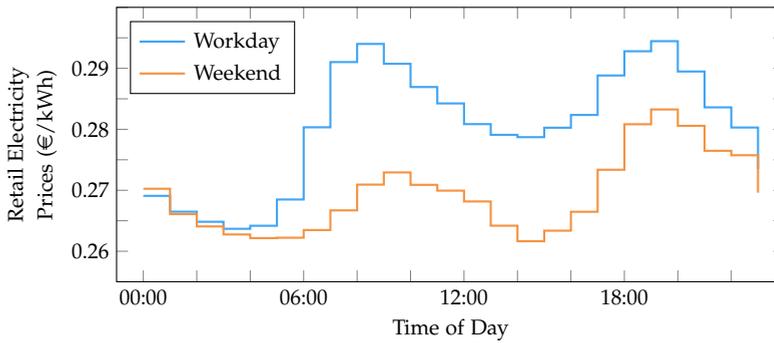
Figure 6.2: Breakdown of 2018 mean retail electricity prices for private customers in Germany [144]; absolute values given in brackets.

Finally, the optimization scheme (6.2) is implemented in *Python* [256]. Due to the use of highly non-linear models in (6.2h) and (6.2i), Discrete Dynamic Programming (DDP) is used as solving method; a detailed description and pseudo code of the algorithms can be found in Appendix B.

After generating optimal charging power trajectories  $p_k^*, \forall k \in \mathcal{K}$ , the operating cost is compared in three modes:

- **Mode I:** No optimization is performed and default charging (i.e., plug in and charge immediately at maximum available power) is applied; still, the energy and aging costs are calculated for measured battery energy and temperature profiles throughout each charging event.

<sup>4</sup>The electricity prices as given in Table 5.1 could either be constant tariffs or mean values derived from these profiles for the time window of each charging event.



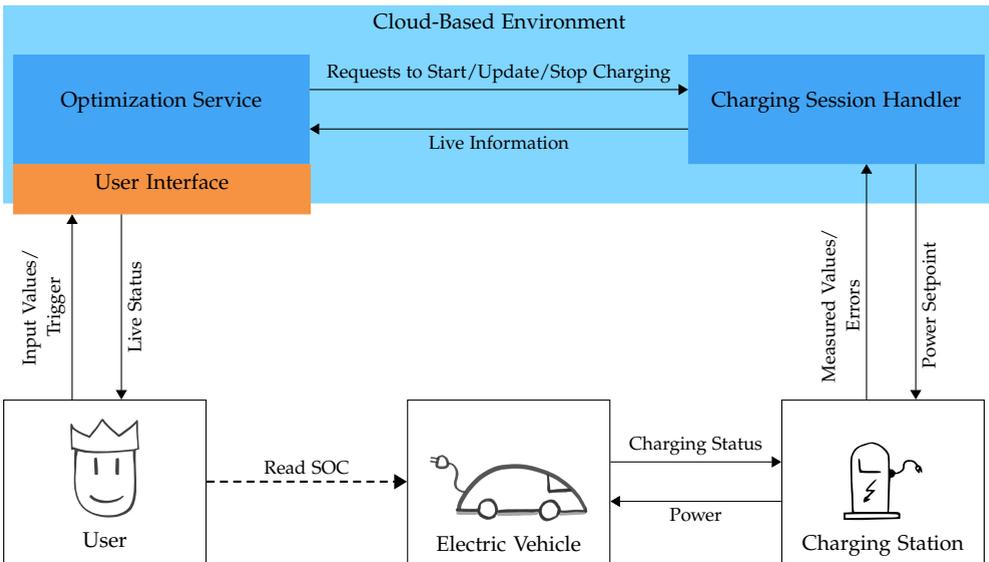
**Figure 6.3:** Characteristic price profiles of retail electricity for workdays and weekends supplemented by 0.188 €/kWh fees and 19% taxes [299].

- **Mode II:** Each charging event is optimized for energy costs only; the aging costs are calculated afterwards.
- **Mode III:** Each charging event is optimized for the sum of both energy and aging costs.

Note that  $p_k^*$  is only calculated once at the beginning of each charging event. In a real-world application, the user might wish to adapt an ongoing charging event to dynamic changes, e.g., departure time, target energy, or electricity tariff. A suitable user interface is therefore required, e.g., such as the one presented in the following Section 6.2.2.

## 6.2.2 Real-World Implementation of Charging Optimization

The second part of the case study aims to discover issues occurring while implementing charging optimization in a realistic environment. In cooperation with MEISENBACHER et al. (see [250,300–302]), a simplified charging optimization scheme is therefore developed and realized in the *Smart Energy System Simulation and Control Center of the Energy Lab 2.0*, see also HAGENMEYER et al. [303].<sup>5</sup> To attain a scalable and flexible setup, a cloud-based architecture is chosen, see Figure 6.4. All



**Figure 6.4:** Cloud-based architecture of the charging optimization scheme as realized in the *Energy Lab 2.0* [303], simplified according to [300].

web applications run as containerized microservices (by means of *Docker* [248]), deployed on a cluster of virtual machines, and orchestrated using *Kubernetes* [249]. An *MQTT*-based communication concept enables performant interaction in between components, see also Section 2.3.

The *Optimization Service* provides a runtime environment for the optimization logic (6.2), which is adapted as follows. To decrease problem complexity, a larger time step of  $\Delta t = 15$  min is chosen. In addition, the battery temperature progression is omitted and a constant battery temperature is assumed throughout the entire

<sup>5</sup>A brief overview of the project can also be found here: <https://energylabsmartcharging.github.io/Smart-Charging-Wizard/>

charging event.<sup>6</sup> Accordingly, (6.2f), (6.2g), and (6.2i) are removed from (6.2). For a faster calculation, a gradient-based solver (*Ipopt* [304]) is used instead of DDP. Furthermore, the power limits in (6.2b) are adapted to the used charging station; i.e.,  $\underline{p} = 0$  kW and  $\bar{p} = 11$  kW. All other parameters given in Section 6.2.1 remain unchanged.

Furthermore, to adapt input values (parking time window, SOC upon arrival and departure, electricity price profile), trigger the optimization scheme, and monitor the charging process, a light-weight user interface is designed by means of *Streamlit* [305]; its three-page layout is shown in Figure 6.5 and a detailed description can be found in [300,301].

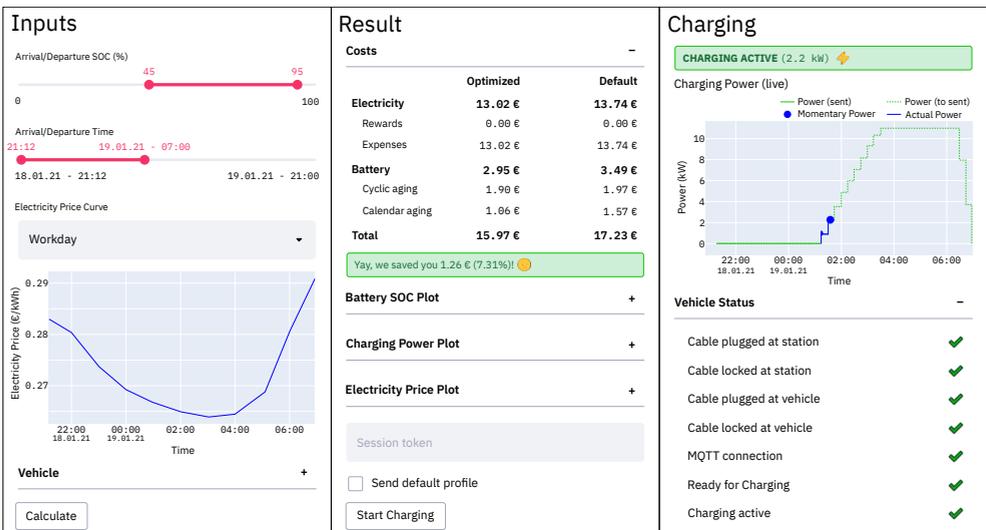


Figure 6.5: Three-page user interface of the charging optimization scheme as realized in the *Energy Lab 2.0* [303], according to [300,301].

To persist information on running charging processes, the *Charging Session Handler* is implemented as headless web application. Independently from the other services, charging events can thus be started, updated and stopped from different devices; e.g., if the *Optimization Service* crashes and needs to be restarted.

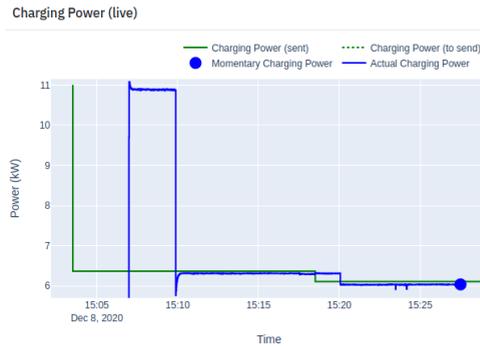
The architecture also includes a real EV charging station (KEBA [306]), whose charging power can be adapted via a programming interface. Figure 6.6a shows the

<sup>6</sup>The expected error from assuming constant battery temperature is elaborated in Section 6.3.3.

EV (Mercedes-Benz EQC) connected to this charging station during testing the real-world charging optimization setup in the *Energy Lab 2.0* [303]. For safety reasons, a programmable logic controller is used, e.g., to validate charging power setpoints received from the *Charging Session Handler*. In case of invalid values, an error message is raised to report the failed request. Furthermore, values measured at the charging station are cyclically published to provide live information for the user interface, as shown in Figure 6.6b; here, the real-time profiles of both set (green) and measured (blue) charging power can be seen.



(a) EV (Mercedes-Benz EQC) connected to controllable charging station (KEBA [306])



(b) Screenshot of user interface showing set (green) and measured (blue) charging power in realtime

Figure 6.6: Impressions of real-world charging optimization setup in the *Energy Lab 2.0* [303].

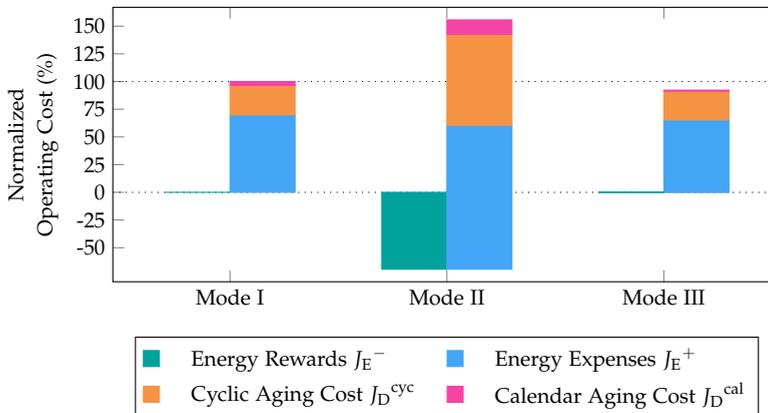
## 6.3 Results of Charging Optimization Case Study

The first part of the case study (see Section 6.2.1) compares operating cost of optimized charging with a standard procedure, see the results in Section 6.3.1. Parameter variation of the optimization scheme allows to identify crucial influence factors on the profitability of V2G, see Section 6.3.2. Although the applied DDP solving method ensures to yield globally optimal solutions, and thus, deliver precise results, its calculation time vastly exceeds acceptable levels of a user application. In some cases, its calculation time reaches up to 45 min for a single charging event. As the high computational complexity seems to arise especially from the thermal battery model (see Section 3.2.2) used in (6.2i), Section 6.3.3 examines the necessity of thermal battery models in charging optimization.

Finally, Section 6.3.4 describes the outcome of the real-world implementation of charging optimization, see Section 6.2.2.

### 6.3.1 Comparison of Electric Vehicle Operating Cost

In Figure 6.7, the operating cost components for all three modes described in Section 6.2.1 are shown; the total cost is normalized against the operating cost of Mode I.



**Figure 6.7:** Comparison of normalized operating cost and its components of 45 real charging events in three modes: standard charging without optimization (Mode I); energy cost optimization, aging costs calculated afterwards (Mode II); energy and aging cost optimization (Mode III).

On average, Mode III yields a 7.8% lower operating cost compared with Mode I; similar results can be found in the literature, e.g. 5.4% in [307] and 13.2% with simplifications in [308]. Although  $p < 0$  kW in (6.2b), i.e., discharging of the EV battery is possible, no energy rewards  $J_E^-$  can be observed in Mode III. This implies that  $J_E^-$  does not compensate for round-trip energy losses, i.e., charging and discharging, and battery aging costs.

Disregarding battery aging underestimates the total operating cost in Mode I by 30.1% on average; in [309] an underestimation of up to 52% is reported. This becomes apparent when applying Mode II: the optimization scheme utilizes fluctuations of the electricity price throughout the charging events to generate energy rewards. As a consequence, the energy costs as defined in (6.3) decrease by 13.3% compared with Mode I. Subsequently calculating the battery aging costs as defined in (6.6), however, yields a 55.8% higher total operating cost. Repeatedly charging and discharging the battery increases the battery temperature  $\theta$  and causes the calendar aging costs  $J_D^{\text{cal}}$  to rise in Mode II. TRIPPE et al. [309] report an even more drastic result for this setup: an 8% reduction of electricity costs, but a threefold increase of battery aging costs. Hence, it is concluded that especially for charging with the allocation of V2G services—in this case *energy arbitrage*—battery aging must not be neglected.

### 6.3.2 Influence Factors on Vehicle-to-Grid Profitability

Evaluating the total EV operating cost of optimized charging in Section 6.3.1 indicates that V2G services such as *energy arbitrage* might not be profitable. Two possible influence factors on V2G profitability are therefore examined in this section.

#### 6.3.2.1 Influence of Battery Prices

According to the literature [27, 92, 93], production prices of EV battery cells will likely decrease within the next decade; reasons for this include a growing EV market and improvements in battery production technology. In anticipation of charging optimization for EV fleets, the operating cost of real, historical charging events (see Section 6.3.1) is compared with the operating cost when assuming underlying future battery prices. Table 6.2 presents the setup and the corresponding results. Particularly, the value loss  $V_{EV}$  of the EV battery is varied, directly affecting the aging costs  $J_D$  in (6.6). To highlight the sole influence of decreasing battery prices, 2018

**Table 6.2:** Comparison of charging optimization (Mode III, see Section 6.2.1) for different future scenarios in terms of battery value loss (according to [92]).

Year	Battery Value Loss $V_{EV}$	Share of Aging Costs $J_D$	Decrease of Operating Cost
2020	6080 €	30.1 %	0 %
2025	4470 €	22.1 %	6.8 %
2030	2770 €	13.7 %	15.9 %

electricity prices are assumed as the cost calculation's underlying future prices. Furthermore, advances in battery technology, i.e., reduced battery aging itself, would have similar effects to decreasing battery production prices.

Compared with 2020 battery prices,  $J_D$  could on average decrease by 26.5 % in 2025 and by 54.4 % in 2030, accordingly. In terms of total operating cost, however, the decrease would only amount to 6.8 % in 2025, or 15.9 % in 2030, respectively. From an EV user's point of view, this reduction of battery aging costs is not sufficient for *energy arbitrage* to become profitable. Consequently, a suitable approach to incentivize EV owners to participate in V2G services is required. For instance, power suppliers could offer a flat compensation for battery aging costs per charging event, in case the user allows V2G services.

In addition to price-based *energy arbitrage*, power suppliers could prospectively offer further rewards for contingency reserves, i.e., feeding back energy to the grid in situations exhibiting a critical lack of power generation (*undergeneration*). Although the actual number of these situations is still unclear, they are expected to only occur several times per year, see Section 1.1.1.2. However, since the fixed cost for bidirectional charging equipment (on-board charger, charging station, communication infrastructure) is still incurred, the rewards that would need to be offered are likely to be very high. From an economic point of view, other alternatives such as stationary battery storages to provide contingency reserves may be more reasonable. Therefore, further research is needed to clarify this issue.

### 6.3.2.2 Influence of Electricity Tariff

The operating cost evaluations in Section 6.2.1 take a dynamic electricity tariff as a basis, which is similar to the ones available today, see e.g., *aWATTar* [286]. With this setup, however, V2G services such as *energy arbitrage* may be unprofitable (see Figure 6.7). A possible reason could be insufficient price variations over time to

compensate aging costs. To investigate this relation, the spread of a charging event's price profile  $\epsilon = (\epsilon_0, \epsilon_1, \dots, \epsilon_{N_k-1})^\top$  is quantified as

$$\gamma = \frac{\max\{\epsilon\} - \min\{\epsilon\}}{\epsilon_{\text{mean}}}, \quad (6.8)$$

with taking the mean workday price  $\epsilon_{\text{mean}} = 0.286 \text{ €/kWh}$  as a reference. To evaluate the sensitivity of the optimization scheme (6.2) to  $\gamma$ , an exemplary workday charging event is used. In particular, the EV arrives at 07:00 o'clock with a battery level of 60% SOC and departs at 15:00 o'clock with 100% SOC. As shown in Figure 6.8, three price profiles with different  $\gamma$  are then tested.

First, a regular workday profile with  $\gamma = 0.054$  is used (drawn in blue). Second, a real-world case is analyzed, which exhibits slight electricity underproduction in the morning and overproduction in the afternoon ( $\gamma = 0.418$ , drawn in orange). In particular, German day-ahead prices of 05.05.2021 are used to which typical fees and taxes (see Figure 6.2) are added [299]. Considering the expected rise of renewable energy sources in the near future, such cases could occur more often compared with today [310]. The power supply—particularly grid operators—may then use batteries of grid-connected EVs as power reserve to compensate drastic grid imbalances. Therefore, a third, stretched price profile derived from the real-world case with  $\gamma = 0.5$  is used (drawn in green).

Table 6.3 outlines the charging cost and its components for all three price curves. In addition, Figure 6.9 shows the corresponding battery energy profiles—here, normalized as SOC profiles. No discharging of the battery can be observed for the regular workday price profile with  $\gamma = 0.054$ ; the corresponding energy rewards  $J_E^-$  amount to 0.0 €. Instead, charging is delayed towards the end in order to reduce battery aging, as high SOC levels cause additional calendar aging, see Section 3.2.3. In contrast, the SOC profile of the real-world case (Figure 6.9, drawn in orange) with  $\gamma = 0.418$  exhibits discharging of the battery at the beginning, when  $\epsilon$  is high; accordingly, energy rewards  $J_E^-$  of  $-5.41 \text{ €}$  are gained via *energy arbitrage*, which reduce the total energy cost  $J_E$  to 5.76 €. The battery is then maintained at a level of approx. 41% SOC before charging, when  $\epsilon$  is low. Calendar aging of the battery (and the corresponding costs  $J_D^{\text{cal}} = 0.55 \text{ €}$ ) is thus reduced by decreasing the charging event's mean SOC level. However, the additional energy throughput due to initially discharging the battery causes increased cyclic aging costs  $J_D^{\text{cyc}} = 3.19 \text{ €}$ . As a consequence, the total battery aging cost  $J_D$  rise to 3.74 €—a 59.8% increase compared with the regular workday price profile. Collectively, the gained energy rewards, however, compensate for additional battery aging costs resulting in a 28.8% decreased total cost of 9.50 €.

To guide future work on grid-supporting **V2G** services, a characteristic threshold for  $\gamma$  is estimated, see also [77, 311]. In particular, a case is designed in which the battery is discharged in one time interval and charged in the subsequent time interval with equal (absolute) power  $|p| \leq 7$  kW. The characteristic threshold

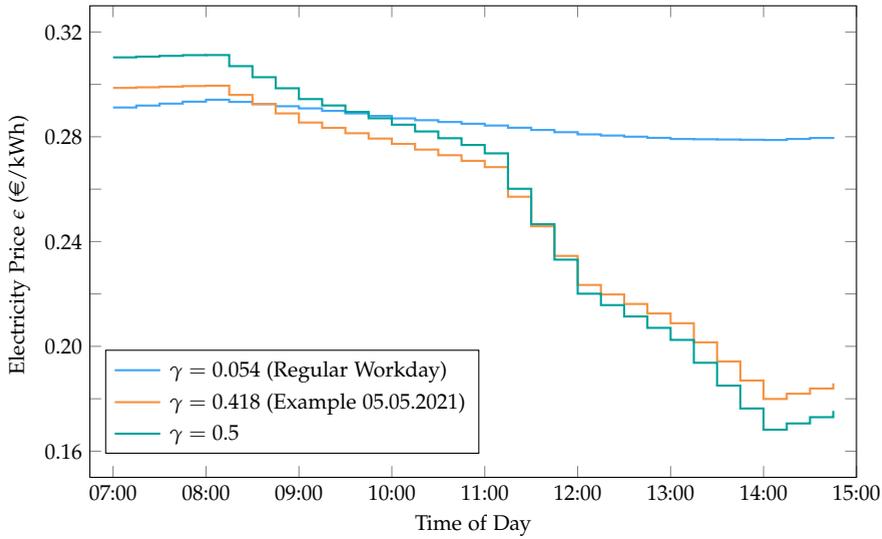
$$\gamma^* = \frac{J_E \cdot (1 - \eta_c) + 2J_D}{\epsilon_{\text{mean}}}, \quad (6.9)$$

then corresponds to the critical price spread above which **V2G** rewards fully compensate for battery aging and conversion losses. Assuming a battery temperature  $\theta = 21$  °C and a round-trip energy efficiency  $\eta_c = 0.997$ ,  $\gamma^* = 0.431$  is obtained. Thus, for this setting, **V2G** potential is fully utilized if the price spread  $\gamma$  is greater than 43.1 % relative to the mean electricity price  $\epsilon_{\text{mean}} = 0.286$  €/kWh. The green **SOC** profile shown in Figure 6.9 with  $\gamma = 0.5$  confirms this result. At the beginning, the battery is fully discharged to the lower battery limit of 10 % **SOC**. After idling for approx. three hours, the battery is charged to 100 % **SOC**. Hence, the available capacity of the **EV** battery is fully utilized to provide **V2G** services; for this setup, a corresponding (maximum) cost reduction of 36.9 % (4.92 €) is achieved via *energy arbitrage*, see Table 6.3. Conversely, **V2G** is likely to be unprofitable for price profiles featuring a smaller price spread than  $\gamma^*$ ; e.g., the ones shown in Figure 6.3, possessing price spreads of  $\gamma = 0.107$  (workday) and  $\gamma = 0.075$  (weekend), which are significantly below  $\gamma^*$ .

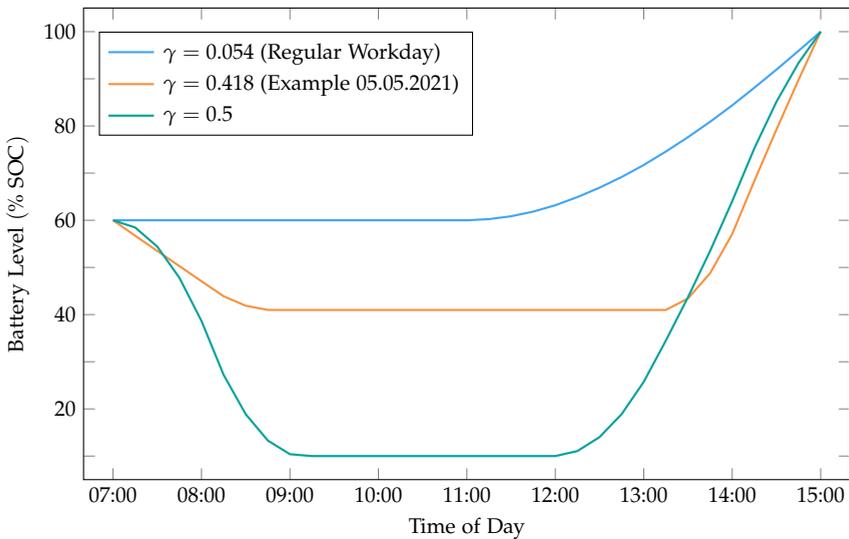
Finally, note that the threshold  $\gamma^*$  also depends on other quantities, e.g., the mean electricity price  $\epsilon_{\text{mean}}$ , and the charging event's **SOC** and time range, whose influences need to be investigated in future work. The specific value of  $\gamma^*$  may therefore not apply directly to other cases. Nevertheless, a qualitative result can be concluded. Influencing **EV** charging processes externally—e.g., as grid operator—requires an

**Table 6.3:** Cost comparison of optimized charging (Mode III, see Section 6.2.1) for three electricity price profiles (see Figure 6.8) with different spreads  $\gamma$  according to (6.8): regular workday ( $\gamma = 0.054$ ), example 05.05.2021 ( $\gamma = 0.418$ ), and stretched example 05.05.2021 ( $\gamma = 0.5$ ).

Cost Component		$\gamma = 0.054$	$\gamma = 0.418$	$\gamma = 0.5$
Energy Rewards	$J_E^-$	0.0 €	-5.41 €	-11.36 €
Energy Expenses	$J_E^+$	11.0 €	11.17 €	14.46 €
<b>Total Energy Cost</b>	$J_E$	<b>11.0 €</b>	<b>5.76 €</b>	<b>3.10 €</b>
Cyclic Aging Costs	$J_D^{\text{cyc}}$	1.63 €	3.19 €	4.82 €
Calendar Aging Costs	$J_D^{\text{cal}}$	0.71 €	0.55 €	0.50 €
<b>Total Aging Cost</b>	$J_D$	<b>2.34 €</b>	<b>3.74 €</b>	<b>5.32 €</b>
<b>Summed Cost</b>	$J_E + J_D$	<b>13.34 €</b>	<b>9.50 €</b>	<b>8.42 €</b>



**Figure 6.8:** Profiles of the electricity price  $\epsilon$  with different spreads  $\gamma$ , according to [299]; the green price profile is obtained by stretching the orange one while keeping a constant mean price  $\epsilon_{\text{mean}} = 0.24 \text{ €/kWh}$ .



**Figure 6.9:** Profiles of battery energy (normalized as SOC) over time for price profiles with different electricity spreads  $\gamma$  (see Figure 6.8).

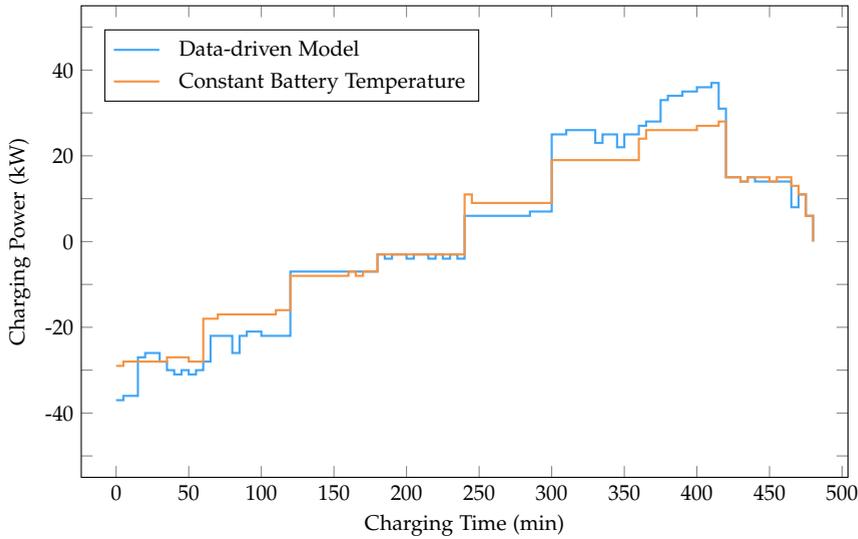
adapted price policy. Instead of offering averaged price profiles (e.g., Figure 6.8, blue), the bare fluctuations of the electricity market such as extensive price peaks (Figure 6.8, red) need to be passed to the EV customer. In this way, price-controlled charging optimization may be adequately used to realize V2G services. Yet, it is unclear whether dynamic electricity prices alone are sufficient for V2G services to support reliable grid operation. Additional control signals based on time- and location-dependent grid state (e.g., overload in low voltage distribution grids) could be required, see also [23,75,298].

### 6.3.3 Effects of Thermal Modeling

Including advanced thermal models such as ANNs (see Section 3.2.2.2) into the charging optimization problem (6.2) allows to precisely simulate the battery temperature. However, this also leads to a significant increase in problem complexity and computational effort to solve the corresponding optimization problem. For this reason, the necessity of a thermal battery model as described in Section 3.2.2 is examined here. Particularly, the assumption of constant battery temperature throughout a charging event is compared with the use of a data-driven thermal model.

In Mode III, assuming constant battery temperature (upon arrival at a charging station), i.e.,  $\Delta\Theta_n = 0\text{K}, \forall n \in \mathcal{N}_k$ , would underestimate the operating cost by 0.55% compared with a data-driven thermal model. When applying Mode II, however, the operating cost would be underestimated by 3.44%. Considering the vastly decreased computation time ( $< 15\text{s}$ , with gradient-based solver) when omitting a thermal model, the error is deemed acceptable for real-world applications. For theoretical evaluations on the profitability of charging optimization, however, thermal models should not be neglected. A compromise might be to include a heuristic thermal model that provides better temperature estimation than omitting a thermal model entirely, while maintaining reasonable computation time.

Besides the errors in estimating operating cost, the presence of a thermal model also influences the decision made by the optimization scheme, i.e., the charging power trajectory  $p^*$ . Figure 6.10 therefore shows exemplary power profiles for assuming constant battery temperature, and for using a data-driven thermal model as described in Section 3.2.2.2. For  $|p| > 7\text{kW}$ , the mean difference of charging power is 3.11 kW, when comparing the constant battery temperature assumption with the data-driven thermal model. However, for  $|p| \leq 7\text{kW}$  the mean deviation of charging power only amounts to 0.75 kW.



**Figure 6.10:** Power profiles over time of an exemplary charging event, if assuming constant battery temperature (orange), and if using a data-driven thermal model as described in [Section 3.2.2.2](#) (blue).

Finally, it can be concluded that although the operating cost only show minor deviations, the charging power profiles change significantly (see [Figure 6.10](#)). In particular, the relevance of the battery temperature rises with the (absolute) charging power. As a consequence, advanced thermal models (e.g., as in [Section 3.2.2.2](#)) should be used for charging events with  $|p| > 7$  kW. For low power levels of  $|p| \leq 7$  kW, assuming constant battery temperature is sufficient.

### 6.3.4 Insights from Real-World Implementation

To evaluate the real-world charging optimization setup, two different scenarios are defined and tested:

- **Scenario I:** Charging on a typical workday; EV arriving at 09:00 o'clock with 45 % SOC, expected departure at 17:00 o'clock with 95 % SOC. The electricity price profile forms a valley between two price peaks at 08:00 and 18:00 o'clock.
- **Scenario II:**
  - a) Charging over night in between two workdays; EV arrives at 12:00 o'clock with 20 % SOC, planned departure at 09:00 o'clock the next day with 95 % SOC; due to electricity underproduction, the price profile exhibits an extensive peak between 16:30 and 20:00 o'clock (see Figure 6.11b).
  - b) Alike to Scenario II-a, at 14:15 o'clock, however, the user decides to already depart on the same day at 18:00 o'clock with 55 % SOC; the ongoing charging process is recalculated and updated, accordingly.

The cost saved in comparison with default charging is given in Table 6.4.<sup>7</sup> Fig-

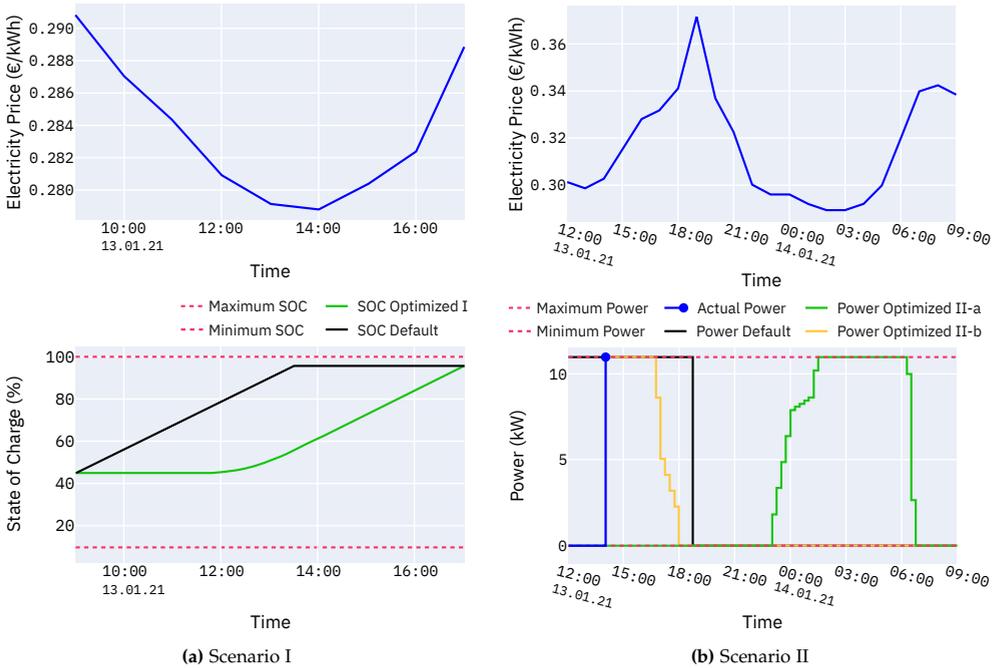
**Table 6.4:** Relative cost savings of optimized charging over default charging (plug in and charge immediately) for three different scenarios, tested with the real-world implementation of charging optimization (see Section 6.2.2), according to [300].

Scenario	Total Cost	Aging Costs	Energy Costs
Scenario I	-5 %	-12 %	-3 %
Scenario II-a	-11 %	-22 %	-8 %
Scenario II-b	-13 %	-52 %	-1 %

Figure 6.11a shows both electricity price and the calculated SOC profiles for Scenario I, as provided to the EV user on the *Result Page* of the user interface (see also Figure 6.5). For the optimized SOC profile, charging occurs in a period of low electricity prices between 12:00 and 17:00 o'clock. Furthermore, instead of symmetrically filling the price valley, considering calendar battery aging (see Section 3.2.3) shifts charging to a later time. The conditions of Scenario II-a yield the charging power profiles shown in Figure 6.11b. Due to electricity underproduction, a price peak of

<sup>7</sup>The term *default charging* refers to a standard procedure of immediately charging the EV at maximum available power upon arrival at a charging station.

up to 0.375 €/kWh occurs between 16:00 and 20:00 o'clock. While default charging (black profile) would disregard the energy shortage, optimized charging (green profile) is shifted to a time at which the electricity price peak flattens out again. In Scenario II-b, a recalculation of Scenario II-a is requested at 14:15 o'clock. As the blue power profile in Figure 6.11b shows, charging is inactive (i.e.,  $p = 0$  kW) until this point in time. Once the ongoing charging event is updated to the new conditions (lower target SOC, early departure), charging is immediately started despite high electricity prices (blue and yellow power profile), as otherwise the target SOC could not be met by the end of the specified time window.



**Figure 6.11:** Examples of real charging processes (default in black, optimized in green/yellow, actually measured charging power in blue) controlled and adapted via the user interface of the charging optimization scheme realized in the *Energy Lab 2.0* [303]; charging event of a typical workday (a); over-night charging with recalculation and early departure (b); according to [300].

Beyond these results, implementing charging optimization in a real-world environment revealed further critical challenges. To address these, some suggestions for future work are given in the following. The seamless integration of charging optimization (or generally *smart charging*) into users' mobility habits calls for uniform communication concepts, see also [312]. With introducing standards for charging

communication, e.g., *ISO 15118-20* [313], this issue is expected to improve. However, a key question still remains to be answered: Should communication with the EV be established via the charging station's power cable or via a wireless, cloud-based connection? While using the former approach ensures communication also, if the EV itself is offline (e.g., when parked in an underground garage), the latter approach provides more flexibility; of course, both approaches could also be combined. Furthermore, as already anticipated in [23], the interaction between EV and the power supply is unclear: How and on which legal and economic foundation may power suppliers externally control a charging process, e.g., by actively reducing charging power? In this context, the question arises whether a dynamic price signal alone suffices to control charging, see also [75, 298] In addition, a uniform interface for power suppliers needs to be developed that allows to control and aggregate several charging processes. Finally, a concept for user interaction needs to be elaborated in which user acceptance issues and suitable incentives for smart charging are addressed.

## 6.4 Conclusion

In Chapter 6, the optimization of EV charging events is addressed. Particularly, a non-linear optimization scheme is engineered, which aims to reduce EV operating cost by leveraging unused energy and time flexibility inherent in the charging process. To track the progression of the battery state, a detailed model of the EV battery (see Section 3.2) is incorporated. In this way, the influence of battery aging on (bidirectional) charging optimization of EVs can also be analyzed.

For validation, a two-part case study is set up. First, the conditions of historical charging events from real EVs are used to simulate optimized charging. In this way, the need for advanced thermal battery models when charging power exceeds 7kW is revealed. In terms of operating cost, taking advantage of the time and energy flexibility of unidirectional EV charging holds the potential to save an average of 7.8%. Furthermore, disregarding the costs arising from battery aging underestimates EVs' total operating cost up to 30%. Battery aging thus hinders many V2G services based on bidirectional power flow—e.g., *energy arbitrage*—from being profitable. To overcome this would require a vast decrease of battery production prices or adapted electricity tariffs that directly represent market fluctuations. Future work on this part should examine stochastic influences on charging optimization, e.g., the provision of occasional power reserves for which EV users might be rewarded—regardless of them being called or not. Furthermore, random user actions (see Chapter 4) should be included in the optimization via chance constraints and their impact on robustness should be quantified. In this context, formulating the target energy criterion as cost function—rather than a constraint—could be helpful. Particularly, a deviation of the target energy could be penalized depending on future charging opportunities; e.g., if the next location with charging opportunity can be reached, not meeting the target energy is uncritical.

In the second part of the case study, charging optimization is implemented in a real-world environment, together with a controllable EV charging station. In this context, an architecture blueprint and communication concept for optimization-based smart charging systems is developed. It is demonstrated that cost-optimized (unidirectional) charging of real EVs can be planned, started and stopped via a web-based user interface. Furthermore, the adaption of ongoing charging processes to changed conditions, e.g., an early departure time, can be successfully performed. However, further developments are required to provide a robust smart charging system that seamlessly integrates into users' daily mobility habits. In particular, issues regarding standardized charging communication, interventions of power suppliers, and user acceptance should be further examined in future work.

## 7 Conclusion and Outlook

In recent years, paradigm shifts could be observed in both energy and mobility sectors. Increasing power generation from intermittent renewable energy sources, e.g., solar and wind, is threatening an efficient and reliable utility grid operation. At the same time, the number of (Battery) Electric Vehicles (EVs) requiring significant amounts of electric energy to charge is increasing rapidly. The resulting coupling of energy and mobility sector inevitably implies that reliable electric mobility depends on a robust power supply. Furthermore, vehicle users are facing unprecedented challenges due to the fact that EVs currently provide less driving range and require more time to recharge compared with Internal Combustion Engine Vehicles (ICEVs). Consequently, EV users' perceive their individual mobility as limited.

The work summarized in this thesis intends to support users when charging their EV by means of an intelligent software application (*charging assistant*), which considers the interests of all involved stakeholders. To achieve this, design features of different possible software architectures are initially reviewed and assessed with respect to flexibility, data privacy, development effort, and complexity (Chapter 2). Furthermore, driving and charging characteristics of the EV are modeled based on historical data (Chapter 3). As core of the charging assistant, three major components are then examined, engineered and implemented as prototypes. In **mobility prediction** (Chapter 4), the user's individual mobility behavior is modeled based on recorded EV trip data. **Charging scheduling** (Chapter 5) then aims to most conveniently and cost-effectively plan charging stops at locations the EV is parked anyway. Finally, **charging optimization** (Chapter 6) identifies unused time and energy flexibility inherent in future charging processes to generate optimal charging power trajectories with respect to dynamic electricity prices and battery aging. Collectively, the major contributions of the present thesis can thus be summarized as follows:

1. A qualitative comparison of alternative architecture designs is performed in order to define a suitable structure of modules that consider all relevant sources of information, their semantic processing steps, and interconnection (Chapter 2).

2. A state-of-the-art analytical model to estimate (user-individual) energy consumption of an EV's future trips is compared with a data-driven approach (Section 3.1). It is found that although data-driven approaches may have benefits in terms of accuracy, an actual live deployment is difficult due to the variety of data sources.
3. A three-part model of the EV battery is developed and validated based on real charging data (Section 3.2). To simulate the battery's energy state, an Equivalent Circuit Model (ECM) is found to be sufficiently precise. In terms of thermal modeling, a data-driven model is developed based on historical data to diminish deficiencies of analytical models. To track the battery's State of Health (SOH), the accuracy of a battery-specific empirical aging model is examined and considered sufficient.
4. For developing and evaluating mobility prediction models, a semi-synthetic data set based on a real user is designed (Section 4.1.2). The script to generate the data set as well as the data itself is published open-source to support research on human mobility.
5. To determine a user's frequently visited locations, a novel two-stage density-based spatial clustering algorithm is engineered and tested against state-of-the-art approaches (Section 4.2). Potential alternatives of implementation and data storage with respect to data-security and privacy issues are also reviewed.
6. A random forest classification model to predict the next location a user may visit is developed and compared with a naive baseline approach using both real and synthetic data sets. In this context, the time-dependent (conditional) probability of each location to be visited is also examined (Section 4.3).
7. Different ensembles of machine learning models are developed to estimate typical parking times at users' frequent locations (Section 4.4). Validation is then conducted by means of both real and synthetic data, and reveals mostly insufficient accuracy for typical smart charging applications.
8. A graph-based scenario generation algorithm is developed and applied to quantify energy demand and occurrence probability of an EV user's mobility scenarios (Section 5.1).
9. A stochastic mixed-integer programming is set up to select EV charging stops within a defined (receding) time horizon and to determine their corresponding energy demand (Section 5.2). The optimization objective is convenient and

---

cost-effective EV charging without disturbing the user's individual mobility habits.

10. Two alternative benchmark strategies are defined to validate charging scheduling. A MONTE CARLO simulation is run to quantify the benefits of the proposed scheme over these strategies in terms of cost, comfort, sustainability, and robustness (Section 5.3).
11. A non-linear optimization scheme is engineered to identify and leverage time and energy flexibility in EV charging processes (Section 6.1). The integration of a detailed EV battery model allows accurate quantification of cost savings arising from reduced battery aging and dynamic electricity tariffs (Section 6.3.1).
12. Based on data from real EV charging events, the influence factors on the profitability of Vehicle-to-Grid (V2G) applications are elaborated, and suggestions for following research are deduced (Section 6.3.2, Section 6.3.3).
13. For a real-world implementation and end-to-end testing of the proposed scheme, an architecture blueprint and communication concept for optimization-based smart charging systems is developed, which provides a basis for future smart charging research (Section 6.3.4). In this context, further issues regarding standardized charging communication, interventions of power suppliers, and user acceptance are discovered.

Although all developed components of the proposed charging assistant are tested by themselves under realistic conditions, integrated testing of the entire system is still required. In this context, both simulation and real-world applications can be helpful to gain insights into usability, robustness, and user acceptance. For the purpose of EV modeling (Chapter 3), suitable strategies to automatically establish model parameters are necessary, e.g., when starting to use the charging assistant with new types of EVs, whose parameters are still unknown. In this case, generic models may be initially used and incrementally adapted during operation to fit the changed EV characteristics (*transfer learning*).

The models used for **mobility prediction** (Chapter 4) call for a similar *cold start* concept, which allows trustworthy predictions with a minimum of data at the beginning. To achieve this, the problem could be generalized to leverage data of several EV users in creating general predictive models. Subsequently, incremental adjustment of model parameters could help to adapt the models to individual users. Beyond that, evaluation of all **mobility prediction** models using a broad variety of

real data is inevitable to prove general applicability, and examine concepts of model surveillance and retraining.

Alike, **charging scheduling** (Chapter 5) should be tested in a real-world environment that incorporates several EV users, end devices, and real-time data sources. In this context, the impact of uncertainty inherent in user decisions can be examined, and a suitable recalculation strategy may be derived. Furthermore, bidirectional charging should be included to consider the provision of V2G services.

The **charging optimization** concept proposed in Chapter 6 disregards stochastic influences. Future work should therefore examine stochastic optimization to handle uncertain parking duration and the provision of further V2G services such as occasional power reserves. Chance constraints may then be put on available energy and power reserves throughout an EV charging process. In this way, both user's individual mobility and callable reserves for the power supply can be guarantee with a specified security level. For these aspects to work properly, however, adequate charging communication is essential, which ensures robust connectivity, e.g., also if parking in an underground garage. Furthermore, legal foundations are required to regulate external control of an EV charging process by power suppliers. In this context, uniform interfaces need to be developed, allowing power suppliers to control and aggregate several charging processes, e.g., in the case of EV fleets.

In summary, the concept of an intelligent charging assistant as proposed in the present thesis may support EV users especially in two situations: First, when switching from an ICEV to an EV by ensuring to stay mobile in the majority of cases while demanding minimal planning effort. Second,—once having overcome initial concerns such as *range anxiety*—when aiming to operate (especially charge) an EV most sustainably, conveniently, and cost-effectively by gathering and combining several sources of information. In a broader scope, however, the proposed concepts particularly help to most effectively fulfill the basic human need of *individual mobility*. Consequently extending this idea to related trends, a future charging assistant may therefore also consider other forms of mobility, e.g., public transport, (autonomous) ride hailing, and (electric) bicycles, which may be included depending on user preferences, cost, and sustainability.

# A Important Evaluation Metrics

Adequate metrics constitute a prerequisite to evaluate both feature and model performance when designing data-driven models. This chapter provides definitions of all metrics used within this thesis, which are oriented at [246].

## A.1 Metrics to Evaluate Feature Performance

To assess the (isolated) performance of single input features of a model—especially data-driven models such as Linear Regression (LR), ANNs and random forests—the following two metrics are used.

### A.1.1 Spearman Rank Correlation Coefficient

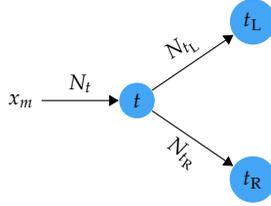
Given a set of inputs  $\mathcal{X} = \{x_1, x_2, \dots, x_N\}$  (e.g., multiple values of one feature) and a corresponding set of outputs  $\mathcal{Y} = \{y_1, y_2, \dots, y_N\}$  (e.g., values of a label), univariate feature relevance (independent of model type) can be determined using the SPEARMAN [314] rank correlation coefficient

$$q_{\text{Sp}}(\mathcal{X}, \mathcal{Y}) = \frac{\sum_{i=1}^N (R(x_i) - \bar{R}_{\mathcal{X}}) (R(y_i) - \bar{R}_{\mathcal{Y}})}{\sqrt{\sum_{i=1}^N (R(x_i) - \bar{R}_{\mathcal{X}})^2 \cdot \sum_{i=1}^N (R(y_i) - \bar{R}_{\mathcal{Y}})^2}}. \quad (\text{A.1})$$

Here,  $R(x_i)$  indicates the rank of  $x_i$  in  $\mathcal{X}$  if sorted in ascending order (1 for the smallest,  $N$  for the largest element). To obtain offset compensation, the mean rank  $\bar{R}_{\mathcal{X}} = \frac{1}{N} \sum_{i=1}^N R(x_i)$  is used. Thus,  $q_{\text{Sp}}$  is a non-parametric rank correlation measure that quantifies the fitness of a monotonic function to describe the relation between  $\mathcal{X}$  and  $\mathcal{Y}$ . It takes values in  $[-1, 1]$  expressing the association of ranks, i.e., 1 for perfect positive association, 0 for no association, and  $-1$  for perfect negative association.

### A.1.2 Gini Importance in Random Forest Models

A random forest  $\mathcal{R}$  (see [260, 276]) consists of a number  $N_T$  of decision trees  $\mathcal{T}$ . Each tree, in turn, consists of nodes  $t$  representing a split of data, see Figure A.1. In particular, each node splits a set of  $N_t$  data samples into two subsets of samples



**Figure A.1:** Node  $t$  of decision tree  $\mathcal{T} \in \mathcal{R}$  performing a binary split of  $N_t$  data samples into subsets with  $N_{t_L}$  and  $N_{t_R}$  samples.

that are passed to successor nodes  $t_L$  and  $t_R$ . The number of data samples  $N_{t_L}$  and  $N_{t_R}$  passed to each successor is determined in the model training procedure for which data with the features  $\mathbf{x} = (x_0, x_1, \dots)^\top$  is used. To evaluate how a feature  $x_m$  influences the performance of a random forest model in terms of accuracy, the corresponding GINI importance

$$q_{\text{Gini}}(x_m) = \frac{1}{N_T} \sum_{\forall \mathcal{T} \in \mathcal{R}} \sum_{\forall t \in \mathcal{T}_{x_m}} \frac{N_t}{N} \cdot \Delta i(t), \quad (\text{A.2})$$

can be consulted, which represents the mean decrease of node *impurity* that feature  $x_m$  causes [277]. To achieve this, the impurity decrease

$$\Delta i(t) = i(t) - \frac{N_{t_L}}{N_t} \cdot i(t_L) - \frac{N_{t_R}}{N_t} \cdot i(t_R), \quad (\text{A.3})$$

is calculated for all nodes in set  $\mathcal{T}_{x_m} \subseteq \mathcal{T}$ , which only contains nodes whose split ratio was influenced by feature  $x_m$  in the training process. The decrease of impurity is then weighted by the portion  $\frac{N_t}{N}$  of data samples reaching node  $t$ , and summed up for all nodes  $t \in \mathcal{T}_{x_m}$ . Finally, the sum over all decision trees  $\mathcal{T} \in \mathcal{R}$  is calculated and divided by the total number of trees  $N_T$ . Here, the GINI index

$$i(t) = 1 - \left( \frac{N_{t_L}}{N_t} \right)^2 - \left( \frac{N_{t_R}}{N_t} \right)^2, \quad (\text{A.4})$$

is used as impurity metric [260, 276, 278]; other adequate metrics may, however, also be used [277]. Note that using GINI importance does not allow to distinguish similarly performing, yet redundant features.

## A.2 Metrics to Evaluate Model Performance

To assess the average model performance, a variety of distance metrics can be applied. For this purpose, true values  $\mathcal{Y} = \{y_1, y_2, \dots, y_N\}$  (i.e., observed values in a data set, also called *ground truth*), are compared with the corresponding values  $\hat{\mathcal{Y}} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N\}$  estimated by a model.<sup>1</sup> The elements of  $\mathcal{Y}$  and  $\hat{\mathcal{Y}}$  can be both scalars and vectors.

### A.2.1 Accuracy

An intuitive metric to quantify (especially classification) model performance is the accuracy

$$q_{\text{ACC}}(\mathcal{Y}, \hat{\mathcal{Y}}) = \frac{1}{N} \sum_{i=1}^N \begin{cases} 1, & \text{if } y_i = \hat{y}_i, \\ 0, & \text{if } y_i \neq \hat{y}_i, \end{cases} \quad (\text{A.5})$$

based on *trivial distance*, which represents the portion of correct estimations in relation to the total number of estimations  $N$ . Note that  $q_{\text{ACC}}$  can be misleading for imbalanced data sets (e.g., 95% of the data samples in one class). For these cases, advanced metrics such as *balanced accuracy* may be more appropriate, see e.g., [315].

### A.2.2 Coefficient of Determination

To quantify the benefit of a model over a naive baseline approach estimating the average  $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$  regardless of the model input, the coefficient of determination (also called  $R^2$ -score)

$$q_{R^2}(\mathcal{Y}, \hat{\mathcal{Y}}) = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (\text{A.6})$$

is defined. Reasonable values are within  $[0, 1]$ , where 0 indicates the absence of any relation and 1 indicates a deterministic relation. Similar to (A.1), an offset compensation is achieved by considering  $\bar{y}$ . Note that by replacing  $\hat{y}_i$  with values  $x_i$  of an input feature,  $q_{R^2}$  can also be used to evaluate (univariate) feature relevance.

---

<sup>1</sup>To quantify unbiased model performance, an independent subset (*test set*) of the entire data must be used, which has previously not been consulted for model creation or parameter tuning.

### A.2.3 Absolute Error Metrics

The Mean Absolute Error (**MAE**)

$$q_{\text{MAE}}(\mathcal{Y}, \hat{\mathcal{Y}}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (\text{A.7})$$

represents the average deviation of a model's estimates  $\hat{y}_i$  compared with true values  $y_i$ . It possesses the same unit as the estimates and is therefore not generally dimensionless. Furthermore, only the absolute deviation is quantified without direction. A systematic over- or underestimation of a model can thus not be determined by means of  $q_{\text{MAE}}$ .

Similar to (A.7), the Relative Mean Absolute Error (**RMAE**) (also *mean absolute percentage error*)

$$q_{\text{RMAE}}(\mathcal{Y}, \hat{\mathcal{Y}}) = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad \forall y_i \neq 0, \quad (\text{A.8})$$

quantifies the average deviation of a model's estimates compared with true values. In addition, each deviation is divided by the true value  $y_i$  to obtain a dimensionless quantity. Accordingly,  $q_{\text{RMAE}}$  is however not defined, if  $y_i = 0$ .

### A.2.4 Squared Error Metrics

Extending definition (A.7) by squaring each deviation between a model's estimates  $\hat{y}_i$  and the true values  $y_i$  yields the Mean Squared Error (**MSE**)

$$q_{\text{MSE}}(\mathcal{Y}, \hat{\mathcal{Y}}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (\text{A.9})$$

which possesses the squared unit of the estimates. Similar to the **MAE**, only the absolute deviation is quantified without direction.

To allow a direct comparison of error and estimates, the Root Mean Squared Error (**RMSE**)

$$q_{\text{RMSE}}(\mathcal{Y}, \hat{\mathcal{Y}}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (\text{A.10})$$

is additionally defined by taking the square root of  $q_{\text{MSE}}$ .

# B Discrete Dynamic Programming Solving Algorithms

The use of non-linear models in (6.2h) and (6.2i) increase the complexity of (6.2) and potentially impede computational tractability. For this reason, Discrete Dynamic Programming (DDP) [316] is used to solve (6.2), see also [250,317]. The remainder of this chapter describes the solving procedure and provides detailed pseudo code of the algorithms.<sup>1</sup>

First, Algorithm B.1 initializes the cost grid  $\mathfrak{J}$ , which is spanned by the energy state variables  $e_n$ , the temperature state variables  $\theta_n$  and all time intervals  $n \in \mathcal{N}_k$ . Furthermore, the grid of all optimal actions  $\mathfrak{P}$  is initially filled with zeros.

---

**Algorithm B.1:** Initialization of backward induction algorithm (see Algorithm B.2), according to [316].

---

**Input:**  $N_k, e_0, e_{N_k}, \theta_0, \underline{e}, \bar{e}, \underline{\theta}, \bar{\theta}, \underline{p}, \bar{p}, \lambda$

*# discretize state and action:*

- 1:  $e_d \leftarrow \text{range}(\text{start: } \underline{e}, \text{stop: } \bar{e}, \text{step: } 0.8 \text{ kWh})$
- 2:  $\theta_d \leftarrow \text{range}(\text{start: } \underline{\theta}, \text{stop: } \bar{\theta}, \text{step: } 1 \text{ K})$
- 3:  $p_d \leftarrow \text{range}(\text{start: } \underline{p}, \text{stop: } \bar{p}, \text{step: } 1 \text{ kW})$

*# initialize cost grid and action grid (penalty value  $\lambda = 1000 \text{ €}$ ):*

- 4:  $\mathfrak{J} \leftarrow \text{zeros}(N_k, \text{length}(e_d), \text{length}(\theta_d))$
- 5:  $\mathfrak{J}[0, :, :], \mathfrak{J}[N_k, :, :] \leftarrow \lambda, \lambda$
- 6:  $\mathfrak{J}[0, \text{argmin}(|e_d - e_0|), \text{argmin}(|\theta_d - \theta_0|)] \leftarrow 0$
- 7:  $\mathfrak{J}[N_k, \text{argmin}(|e_d - e_{N_k}|), :] \leftarrow 0$
- 8:  $\mathfrak{P} \leftarrow \text{zeros}(N_k, \text{length}(e_d), \text{length}(\theta_d))$

**Output:**  $\mathfrak{J}, \mathfrak{P}, e_d, \theta_d, p_d$

---

<sup>1</sup>In Algorithm B.1, Algorithm B.2, and Algorithm B.3, the notation  $x[n]$  indicates accessing the  $n$ -th element of vector  $x$ .

Second, [Algorithm B.2](#) performs a backward induction in which the cost for all possible transitions in between two stages is calculated. In particular, [Algorithm B.2](#) runs backwards ( $n \in \{N_k - 1, N_k - 2, \dots, 0\}$ ) to update  $\mathfrak{J}$  for all possible  $e_i \in e_d$  and  $\theta_j \in \theta_d$ . Similarly, all corresponding optimal actions  $\mathfrak{P}$  are determined (*backward induction*). To avoid infeasible trajectories, a penalty value  $\lambda$  is assigned to the according value in  $\mathfrak{J}$ , if a constraint (6.2b) to (6.2i) is violated. Here, the value of  $\lambda$  should be chosen significantly higher than the sum of usual costs.

---

**Algorithm B.2:** Backward induction algorithm to create cost grid  $\mathfrak{J}$  and corresponding optimal actions  $\mathfrak{P}$ , according to [316].

---

**Input:**  $N_k, \mathfrak{J}, \mathfrak{P}, e_d, \theta_d, p_d, \underline{e}, \bar{e}, \underline{\theta}, \bar{\theta}, \underline{p}, \bar{p}, \epsilon, \lambda$

```

1: for  $n \leftarrow N_k - 1$  to 0 :
2:   for all  $e_i \in e_d$  :
3:     for all  $\theta_j \in \theta_d$  :
4:       # initialize cached total cost:
5:        $J \leftarrow \text{ones}(\text{length}(p_d)) \cdot \lambda$ 
6:       for all  $p_m \in p_d$  :
7:         # validate charging power constraints (6.2b):
8:         if  $\underline{p}(e_i, \theta_j) \leq p_m \leq \bar{p}(e_i, \theta_j)$  :
9:           # calculate state transitions:
10:           $e_{n+1} \leftarrow e_i + \Delta E(e_i, \theta_j, p_m)$ 
11:           $\theta_{n+1} \leftarrow \theta_j + \Delta \Theta(p_m, \dot{Q}_{\text{loss}}(e_i, \theta_j, p_m), \Delta E(e_i, \theta_j, p_m), \theta_j)$ 
12:          # validate state constraints (6.2c) and (6.2f):
13:          if  $\underline{e} \leq e_{n+1} \leq \bar{e}$  and  $\underline{\theta} \leq \theta_{n+1} \leq \bar{\theta}$  :
14:            # calculate transition costs (6.3) and (6.6):
15:             $J_E \leftarrow J_E(p_m, \epsilon_n)$ 
16:             $J_D \leftarrow \frac{V_{EV}}{H_{EV}} (\Delta H_{\text{cal}}(e_i, \theta_j, H_0) + \Delta H_{\text{cyc}}(|\Delta E(e_i, \theta_j, p_m)|))$ 
17:            # calculate and cache total cost:
18:             $J[m] \leftarrow$ 
19:             $J_E + J_D + \mathfrak{J}[\text{argmin}(|e_d - e_{n+1}|), \text{argmin}(|\theta_d - \theta_{n+1}|)]$ 
20:          # assign minimum cached cost and corresponding action:
21:           $\mathfrak{J}[n, i, j] \leftarrow \min(J)$ 
22:           $\mathfrak{P}[n, i, j] \leftarrow p_d[\text{argmin}(J)]$ 

```

**Output:**  $\mathfrak{J}, \mathfrak{P}$

---

Once the cost grid  $\mathfrak{J}$  is filled, [Algorithm B.3](#) conducts a forward integration in which a path of minimum cost through  $\mathfrak{J}$  is chosen, starting from the initial state. Particularly, [Algorithm B.3](#) integrates forward in time ( $n \in \{0, 1, \dots, N_k\}$ ), starting from the initial battery energy  $e_0$  and battery temperature  $\theta_0$ . For each  $n \in \mathcal{N}_k$ , the cost-optimal action is taken from  $\mathfrak{P}$  based on the current state  $e_n$  and  $\theta_n$ . This procedure yields the globally optimal charging power trajectory  $\mathbf{p}_k^*$  (*forward integration*) [316].

---

**Algorithm B.3:** Forward integration algorithm to find the optimal charging power trajectory  $\mathbf{p}_k^*$ , according to [316].

---

**Input:**  $N_k, \mathfrak{J}, \mathfrak{P}, e_0, \theta_0, e_d, \theta_d, \epsilon$

# find starting point in the cost grid:

1:  $i, j \leftarrow \text{argmin}(\mathfrak{J}[0, :, :])$

# initialize output and assign corresponding action:

2:  $\mathbf{p}_k^* \leftarrow \text{zeros}(N_k)$

3:  $\mathbf{p}_k^*[0] \leftarrow \mathfrak{P}[0, i, j]$

# initialize costs:

4:  $J_E, J_D \leftarrow 0, 0$

# start forward integration loop:

5: **for**  $n \leftarrow 0$  **to**  $N_k - 1$  :

# calculate state transitions:

6:  $e_{n+1} \leftarrow e_n + \Delta E(e_n, \theta_n, \mathbf{p}_k^*[n])$

7:  $\theta_{n+1} \leftarrow \theta_n + \Delta \Theta(\mathbf{p}_k^*[n], \dot{Q}_{\text{loss}}(e_n, \theta_n, \mathbf{p}_k^*[n]), \Delta E(e_n, \theta_n, \mathbf{p}_k^*[n]), \theta_n)$

# calculate costs with (6.3) and (6.6):

8:  $J_E \leftarrow J_E + J_E(p_n, \epsilon_n)$

9:  $J_D \leftarrow J_D + \frac{V_{\text{EV}}}{H_{\text{EV}}} (\Delta H_{\text{cal}}(e_n, \theta_n, H_0)) + \Delta H_{\text{cyc}}(|\Delta E(e_n, \theta_n, \mathbf{p}_k^*[n])|)$

# find nearest discrete state and assign corresponding action:

10:  $\mathbf{p}_k^*[n+1] \leftarrow \mathfrak{P}[n+1, \text{argmin}(|e_d - e_{n+1}|), \text{argmin}(|\theta_d - \theta_{n+1}|)]$

**Output:**  $\mathbf{p}_k^*, J_E, J_D$

---



# List of Abbreviations

<b>ANN</b>	Artificial Neural Network
<b>API</b>	Application Programming Interface
<b>BMS</b>	Battery Management System
<b>CAN</b>	Controller Area Network
<b>CO<sub>2</sub></b>	Carbon Dioxide
<b>DDP</b>	Discrete Dynamic Programming
<b>ECM</b>	Equivalent Circuit Model
<b>EV</b>	(Battery) Electric Vehicle
<b>FCEV</b>	Fuel Cell Electric Vehicle
<b>GPS</b>	Global Positioning System
<b>ICEV</b>	Internal Combustion Engine Vehicle
<b>JSON</b>	JavaScript Object Notation
<b>LR</b>	Linear Regression
<b>LiCoO<sub>2</sub></b>	Lithium-Cobalt-Oxide
<b>MAE</b>	Mean Absolute Error
<b>MLP</b>	Multi Layer Perceptron
<b>MQTT</b>	Message Queuing Telemetry Transport
<b>MSE</b>	Mean Squared Error
<b>RMAE</b>	Relative Mean Absolute Error
<b>RES</b>	Renewable Energy Source
<b>RMSE</b>	Root Mean Squared Error
<b>SBC</b>	Single-Board Computer
<b>SOC</b>	State of Charge
<b>SOH</b>	State of Health
<b>TRL</b>	Technology Readiness Level
<b>V2G</b>	Vehicle-to-Grid



# List of Symbols

This section summarizes the notation used throughout the thesis. The taxonomy is partly based on the formulation given in [246]. It follows these general rules:

- Scalars, parameters, indexing variables and functions are denominated by lower case letters, e.g.,  $a$  for parameters, scalars or variables,  $f(\bullet)$  for functions or functionals,  $i$  for indices
- Upper case bold face letters represent matrices, e.g.,  $\mathbf{M}$ , lower case bold face letters represent vectors, e.g.,  $\mathbf{x}$  for column vectors and  $\mathbf{x}^\top$  for row vectors
- Upper case script letters represent sets, e.g.,  $\mathcal{A} = \{a_1, a_2, \dots, a_A\}$ , with the entities  $a_{1..A}$ , and the cardinality  $A$  denominated by an upper case letter.
- $x_n$  indicates the value of  $x$  at time  $t_n$
- The symbol  $\hat{x}$  indicates an estimate of  $x$
- The symbol  $\dot{x}$  indicates the derivative of  $x$  with respect to time  $t$
- The symbol  $\bar{x}$  represents the upper limit of  $x$ ,  $\underline{x}$  represents the lower limit of  $x$
- Optimal solutions are indicated by an asterisk, e.g.,  $x^*$  would denote the optimal solution to  $\min\{f(x)\}$

## Latin Symbols

$\mathcal{B}$	Set of Historical Battery Data Samples	[-]
$c_h$	Heat Capacity of the Battery	[J/K]
$C$	Plug-In Event	[-]
$d$	Scenario of Scenario Tree	[-]
$\mathcal{D}$	Set of all Scenarios of Scenario Tree	[-]
$e$	Battery Energy	[J]
$e_A$	Battery Energy upon Arrival	[J]

LIST OF SYMBOLS

---

$e_D$	Battery Energy upon Departure	[J]
$e_{\max}$	Momentary Maximum Available Battery Capacity	[J]
$e_{\text{nom}}$	Nominal Available Battery Capacity	[J]
$\underline{e}$	Lower Bound of Battery Energy	[J]
$\tilde{e}$	Adapted Lower Bound of Battery Energy	[J]
$\bar{e}$	Upper Bound of Battery Energy	[J]
$\Delta E$	Energy Throughput	[J]
$\Delta \hat{E}$	Estimated Energy Throughput	[J]
$E$	Charged Energy	[J]
$\underline{E}$	Lower Bound of Charged Energy	[J]
$\bar{E}$	Upper Bound of Charged Energy	[J]
$\mathbb{E}$	Expected Value	[-]
$\Delta H_{\text{cyc}}$	Cyclic Battery Capacity Fade	[%]
$\Delta H_{\text{cal}}$	Calendar Battery Capacity Fade	[%]
$H_{\text{EV}}$	Total Battery Capacity Fade, EV Application	[%]
$I_{\text{bat}}$	Battery Current	[A]
$\mathcal{J}$	Set of Historical Trips	[-]
$J_E$	Energy Cost Function	[€]
$J_D$	Battery Degradation Cost Function	[€]
$J_C$	Plug-In Cost Function	[€]
$\mathfrak{J}$	Discrete Dynamic Programming Cost Grid	[-]
$k$	Charging Event	[-]
$\mathcal{K}$	Set of Historical Charging Events	[-]
$l$	Geodetic Location	[-]
$\hat{l}_{j+1,\tau}$	Predicted Parking Location of Next Trip	[-]
$l_{\text{lat}}$	Geodetic Location, Latitude	[°]
$l_{\text{lon}}$	Geodetic Location, Longitude	[°]
$l_{\text{alt}}$	Geodetic Location, Altitude	[m]
$\mathcal{L}$	Set of Known, Distinct Locations	[-]
$N$	Number of Time Intervals	[-]
$\mathcal{N}$	Set of Time Intervals	[-]
$o$	Mileage of Vehicle	[m]
$p$	Gross Charging Power	[W]
$\underline{p}$	Lower Bound of Charging Power	[W]
$\bar{p}$	Upper Bound of Charging Power	[W]
$\mathfrak{P}$	Discrete Dynamic Programming Optimal Action Grid	[-]
$\mathbb{P}$	Probability	[-]
$\mathbb{P}_r$	Residual Probability	[-]
$\mathbb{P}_t$	Transition Probability	[-]

$\dot{Q}_{amb}$	Heat Flow between Battery and Environment	[W]
$\dot{Q}_{BMS}$	Heat Flow Withdrawn or Injected from BMS	[W]
$\dot{Q}_{loss}$	Heat Flow from Internal Battery Losses	[W]
$q_{ACC}$	Accuracy, see <a href="#">Section A.2.1</a>	[-]
$q_{Gini}$	Gini Importance, see <a href="#">Section A.1.2</a>	[-]
$q_{MAE}$	Mean Absolute Error, see <a href="#">Section A.2.3</a>	[-]
$q_{MSE}$	Mean Squared Error, see <a href="#">Section A.2.4</a>	[-]
$q_{RMAE}$	Relative Mean Absolute Error, see <a href="#">Section A.2.3</a>	[-]
$q_{RMSE}$	Root Mean Squared Error, see <a href="#">Section A.2.4</a>	[-]
$q_{R^2}$	R <sup>2</sup> -Score, Coefficient of Determination, see <a href="#">Section A.2.2</a>	[-]
$q_{Sp}$	Spearman Rank Correlation Coefficient, see <a href="#">Section A.1.1</a>	[-]
$R_i$	Battery Internal Resistance	[ $\Omega$ ]
$s$	Stage of Scenario Tree	[-]
$\mathcal{S}$	Set of all Stages of Scenario Tree	[-]
$S_C$	Battery State of Charge	[%]
$S_H$	Battery State of Health	[%]
$\Delta t$	Time Step Duration	[min]
$U_{OCV}$	Battery Open-Circuit Voltage	[V]
$U_{bat}$	Battery Terminal Voltage	[V]
$v$	Tree Vertex	[-]
$V_{EV}$	Total Battery Value Loss, EV Application	[€]
$\mathcal{V}$	Set of Tree Vertices	[-]
$w$	Walking Distance to Charging Station	[m]
$\mathbf{x}_\Gamma$	Feature Vector of Consumption Model	[-]
$\mathbf{x}_\Theta$	Feature Vector of Thermal Battery Model	[-]
$\mathbf{x}_{\tau_p}$	Feature Vector of Parking Duration Model	[-]
$\mathbf{x}_l$	Feature Vector of Next Place Model	[-]
$z$	Constraint Violation Flag	[-]

## Greek Symbols

$\alpha$	Heat Transition Coefficient between Battery and Environment	[W/K]
$\beta$	Battery Degradation Parameter	[-]
$\gamma$	Electricity Price Selling-Buying-Ratio	[%]
$\Gamma$	Electric Energy Consumption	[kWh/km]
$\hat{I}_{BL}$	Baseline Estimation of Electric Energy Consumption	[kWh/km]

## LIST OF SYMBOLS

---

$\hat{\Gamma}_{DD}$	Data-Driven Estimation of Electric Energy Consumption	[kWh/km]
$\Gamma_{I \neq \mathcal{L}}$	Electric Energy Consumption for Visiting Unknown Locations	[kWh/km]
$\epsilon$	Electricity Price	[\\$/kWh]
$\epsilon_C$	Constant Cost for Plug-In Event	[\\$]
$\epsilon_v$	Security Level for Arrival Energy	[-]
$\epsilon_{\tau_p}$	Tolerance of Parking Duration Estimation	[min]
$\zeta$	Battery Calendar Aging Time Equivalent	[s]
$\eta_c$	Efficiency of Charging Process	[%]
$\theta_{amb}$	Ambient Temperature	[°C]
$\theta$	Battery Temperature	[°C]
$\Delta\Theta$	Battery Temperature Difference	[K]
$\Delta\hat{\Theta}_{RO}$	Estimated Battery Temperature Difference, Reduced-Order Model	[K]
$\Delta\hat{\Theta}_{DD}$	Estimated Battery Temperature Difference, Data-Driven Model	[K]
$\underline{\theta}$	Lower Bound of Battery Temperature	[°C]
$\bar{\theta}$	Upper Bound of Battery Temperature	[°C]
$v$	Vehicle Speed	[m/s]
$\tau$	Travel Time	[s]
$\tau_p$	Parking Duration	[s]
$\hat{\tau}_p$	Estimated Parking Duration	[s]
$\phi_{\Gamma}$	Weight of Baseline Energy Consumption Estimation	[-]
$\phi_w$	Cost Factor of Walking Distance to Charging Station	[\\$/m]

# List of Figures

1.1	Comparison of life-cycle CO <sub>2</sub> emissions of ICEV and EV for varying RES-share [6]. . . . .	1
1.2	Relative share of hydro power, solar power, and wind power in total electricity generation in Germany, 1990 to 2020 [14]. . . . .	2
1.3	Electricity consumption in the European Union grouped by sectors; illustration of the transport sector’s reluctant consumption, 1990 to 2020 [14]. . . . .	3
1.4	Breakdown of EV manufacturing cost focusing on powertrain-related components [28]. . . . .	4
1.5	Schematic grid topology with transmission and distribution grids in different voltage levels, as well as energy producers and consumers, see also [31]. . . . .	6
1.6	Road map of EV charging modes also considering a connection to the electric utility grid, approx. 2015 to 2025, according to [49]. . . . .	8
1.7	Schematic section view of lithium ion battery. Positive electrode (LiCoO <sub>2</sub> ) releases electrons when charging; negative electrode (graphite) releases electrons when discharging; lithium ions diffuse through the separator accordingly [89]. . . . .	12
1.8	Breakdown of battery manufacturing cost comprising raw material values, labor and overhead costs [27]. . . . .	13
1.9	Equivalent circuit network of EV battery for low-dynamic or stationary operation [104]. . . . .	16
1.10	Number of public charging stations and disproportionately growing amount of registered EVs in Germany, 2011-2021 [144,145]. . . . .	20
1.11	Histogram illustrating the frequency of daily driven distance [171]. . . . .	23
1.12	Characterization of charging habits according to frequency of simultaneously charging EVs in residential areas [155]. . . . .	23
1.13	Characterization of charging habits according to SOC at the beginning and the end of charging events [184]. . . . .	24

2.1 Layout of all stakeholders (white boxes), the EV (light blue box) represented by the vehicle-specific models (orange boxes), all developed modules (dark blue boxes), and their respective connection among each other. . . . . 31

2.2 Scheme of cloud-based environment for data acquisition via a MQTT-broker, storage in several databases, and application of data-driven models or microservices [243,244]. . . . . 42

3.1 Schematic block diagram of the consumption model and its components. . . . . 48

3.2 Comparison of energy consumption estimation  $\hat{\Gamma}_{BL}$  of the Baseline Model (see Figure 3.1) with the true energy consumption  $\Gamma$  of recorded EV trips; the orange line indicates ideal model behavior. . . 49

3.3 Histogram of training data classified by the specific energy consumption  $\Gamma$  in bins of 0.05 kWh/km. . . . . 51

3.4 Comparison of energy consumption estimation  $\hat{\Gamma}_{DD,LR}$  of the linear regression model (see Section 3.1.2.1) with the true energy consumption  $\Gamma$  of recorded EV trips in independent test set  $\mathcal{J}_{test}$ ; the orange line indicates ideal model behavior. . . . . 58

3.5 Comparison of energy consumption estimation  $\hat{\Gamma}_{DD,MLP}$  of the ANN regression model (see Section 3.1.2.2) with the true energy consumption  $\Gamma$  of recorded EV trips in independent test set  $\mathcal{J}_{test}$ ; the orange line indicates ideal model behavior; the dashed pink line indicates a systematic misrepresentation. . . . . 59

3.6 Comparison of energy consumption estimation  $\hat{\Gamma}_{DD,RF}$  of the random forest regression model (see Section 3.1.2.3) with the true energy consumption  $\Gamma$  of recorded EV trips in independent test set  $\mathcal{J}_{test}$ ; the orange line indicates ideal model behavior; the dashed pink line indicates a systematic misrepresentation. . . . . 60

3.7 Schematic block diagram of EV battery model and its components. . . 61

3.8 Equivalent circuit model of an EV battery for low-dynamic operation with internal resistance  $R_i$  and voltage source  $U_{OCV}$ . . . . . 62

3.9 Characteristic curve of open circuit voltage  $U_{OCV}$  with respect to normalized battery energy  $e$ . . . . . 63

3.10 Characteristic curve of internal battery resistance  $R_i$  with respect to battery SOC  $S_C$  and battery temperature  $\theta$ . . . . . 64

3.11 Histogram of battery temperature change  $\Delta\theta_b$  for all data samples  $b$  in battery data set  $\mathcal{B}$ . . . . . 67

3.12	Exemplary profiles of a quantity's true values $y$ and estimated values $\hat{y}$ over the time window $[t_{k,0}, t_{k,N_k}]$ of charging event $k$ to explain the error measures of the battery model (see Section 3.2).	72
3.13	True and estimated energy throughput for individual samples $b \in \mathcal{B}_k$ for all $k \in \mathcal{K}$ (local error, normalized as SOC) using the electrical battery model (Section 3.2.1), the orange line indicates ideal model behavior.	73
3.14	Deviation $\hat{e}_{N_k} - e_{N_k}$ of true and estimated SOC at the end of all charging events $k \in \mathcal{K}$ (global error) using the electrical battery model (Section 3.2.1).	74
3.15	True and estimated change of battery temperature for individual samples $b \in \mathcal{B}_k, \forall k \in \mathcal{K}$ (local error) using the reduced-order thermal battery model (Section 3.2.2), the orange line indicates ideal model behavior.	76
3.16	Deviation $\hat{\theta}_{N_k} - \theta_{N_k}$ of true and estimated battery temperature $\theta$ at the end of all charging events $k \in \mathcal{K}$ (global error) using the reduced-order thermal battery model (Section 3.2.2).	77
3.17	True and estimated change in battery temperature for individual samples $b \in \mathcal{B}_{\text{test}}$ (local error) using the LR thermal battery model (Section 3.2.2.2), the orange line indicates ideal model behavior.	78
3.18	Deviation $\hat{\theta}_{N_k} - \theta_{N_k}$ of true and estimated battery temperature $\theta$ at the end of all charging events $k \in \mathcal{K}_{\text{test}}$ (global error) using the LR thermal battery model (Section 3.2.2.2).	78
3.19	True and estimated change in battery temperature for individual samples $b \in \mathcal{B}_{\text{test}}$ (local error) using the ANN thermal battery model (2 hidden layers, 10 neurons each, Section 3.2.2.2), the orange line indicates ideal model behavior.	79
3.20	Deviation $\hat{\theta}_{N_k} - \theta_{N_k}$ of true and estimated battery temperature $\theta$ at the end of all charging events $k \in \mathcal{K}_{\text{test}}$ (global error) using the ANN thermal battery model (2 hidden layers, 10 neurons each, Section 3.2.2.2).	80
3.21	True and estimated battery SOH values using the battery aging model as describes in Section 3.2.3, the orange line indicates ideal model behavior.	81
4.1	Schematic block diagram of the mobility prediction module and its components; inputs of the EV user, e.g., via a calendar interface, may provide information on planned trips.	85

4.2 Map section of southwestern Germany with inferred known locations  $\mathcal{L}_{\text{real}}$  (blue circles) and noise points (pink crosses) using the proposed two-stage clustering approach, based on real data  $\mathcal{J}_{\text{real}}$ , orange double lines indicate major highways. . . . . 94

4.3 Inferred known locations  $\mathcal{L}_{\text{real}}$  and their relative relevance to compare default *DBSCAN* with the proposed two-stage clustering approach (see Section 4.2.2) on the real data set  $\mathcal{J}_{\text{real}}$ ; in addition, the share of data points labeled as “Noise” is shown. . . . . 95

4.4 Comparison of the default *DBSCAN* (a) and the two-stage spatial clustering approach (including data cleaning) (b). The two-stage approach combines clusters that are not clearly separated to ensure well-defined single locations for further assessment. . . . . 96

4.5 Schematic sequence of trips (light blue) and parking periods (dark blue) to indicate the point in time (red arrow) of predicting the next location to visit. . . . . 97

4.6 Transition matrix containing transition probabilities  $\mathbb{P}_t$  for all combinations of current and next location on a Monday morning at 05:00 o’clock; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1). . . . . 105

4.7 Stacked transition probabilities  $\mathbb{P}_t$  for a trip starting from Location 1 (*home*) to all possible end locations on an arbitrary Monday; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1). . . . . 106

4.8 Transition matrix containing transition probabilities  $\mathbb{P}_t$  for all combinations of current and next location on a Monday afternoon at 16:00 o’clock; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1). . . . . 107

4.9 Stacked transition probabilities  $\mathbb{P}_t$  for a trip starting from Location 2 (*work*) to all possible end locations on an arbitrary Monday; random forest model (see Section 4.3.2.2) trained on real data set  $\mathcal{J}_{\text{real}}$  (see Section 4.1.1). . . . . 108

4.10 Schematic sequence of trips (light blue) and parking periods (dark blue) to indicate the point in time (red arrow) of estimating the duration  $\tau_p$  of the next parking period. . . . . 110

4.11 Accuracy ( $q_{p,60}$ ) of different models to estimate parking duration (Label A) and next departure time (Label B) on the synthetic data set  $\mathcal{J}_{\text{syn}}$  with five-fold cross-validation, horizontal line indicates mean accuracy. . . . . 116

4.12	Accuracy ( $q_{p,60}$ ) of different models to estimate parking duration (Label A) and next departure time (Label B) on the real data set $\mathcal{J}_{\text{real}}$ with five-fold cross-validation, horizontal line indicates mean accuracy. . . . .	117
5.1	Block diagram of the charging scheduling scheme; mobility prediction (Chapter 4) determines predicted next location $\hat{l}_{j+1,\tau}$ and estimated parking duration $\hat{\tau}_p$ , EV driving and charging characteristics are derived from the corresponding EV models (Chapter 3). . . . .	125
5.2	Exemplary generation of scenarios based on a set $\mathcal{L} = \{A, B, C\}$ of three known locations (a) and the constructed out-tree $\mathcal{T}$ (b) to demonstrate the scenario generation procedure and nomenclature. . . . .	128
5.3	Different interpolations for the adapted minimum energy $\tilde{e}_s$ , i.e., the dynamic battery reserve for charging scheduling, as a function of the residual probability $\mathbb{P}_r(s)$ . . . . .	136
5.4	Dependency of battery energy ( $e_A$ or $e_D$ ) and the corresponding violation flags ( $z_A$ or $z_D$ ) to explain integer reformulation of chance constraints (5.2i) and (5.2j). . . . .	137
5.5	Map section with position of known locations A to H based on synthetic data $\mathcal{J}_{\text{syn}}$ (see Section 4.1.2); blue circles: locations with close-by charging station, orange circles: locations without charging opportunity. . . . .	141
5.6	Matrix of travel times $\tau$ , i.e., the time required to travel from a (known) start location to another (known) end location (see Table 5.1), estimated using an external routing service [290]; note the format <i>hours:minutes</i> . . . . .	141
5.7	Matrix of energy consumption $\Gamma$ , i.e., the energy required to travel from a (known) start location to another (known) end location (see Table 5.1), estimated using an external routing service [290] together with a model as in Section 3.1; values given in kWh. . . . .	142
5.8	Exemplary charging scheduling run with sets of SOC profiles (without charging in orange, optimized in blue) over tree stages $s$ (see Section 5.1), adapted minimum energy $\tilde{e}_s$ drawn as dotted line ( $\varepsilon_v = 0\%$ ), turquoise shaded areas represent parking phases of the EV. . . . .	147

5.9 Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing minimum battery energy constraint violations, i.e.,  $e < \tilde{e}_s$ ; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy. . . . . 148

5.10 Comparison of given security level  $\varepsilon_v$  with the observed probability of violating the adapted lower energy limit  $\tilde{e}_s$  for the optimized charging strategy (a) and the alternative strategies *Always Charge* (b) and *Minimum Charge* (c). . . . . 148

5.11 SOC profiles for proposed (green) and alternative charging scheduling strategies (blue, orange) for one simulation run over the scheduling time horizon, adapted minimum energy  $\tilde{e}_s$  drawn as dotted line ( $\varepsilon_v = 0\%$ ), turquoise shaded areas represent parking phases of the EV. 150

5.12 Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing the number of plug-in events C; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy. . . . . 150

5.13 Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing the calendar aging; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy. . . . . 151

5.14 Comparison of alternative charging scheduling strategies (see Section 5.3.2) with the proposed charging scheduling scheme (Section 5.2) in terms of reducing the mean price for electricity; minimum, maximum and mean relative difference is given; negative values indicate a benefit over the proposed strategy. . . . . 152

6.1 Schematic block diagram of the charging optimization scheme interacting with the EV battery model (see Section 3.2); both targeted battery energy  $e_{D,k}$  and available time window  $[t_{k,0}, t_{k,\tau_{p,k}}]$  are taken from the charging schedule generated with a scheme as described in Chapter 5. . . . . 155

6.2 Breakdown of 2018 mean retail electricity prices for private customers in Germany [144]; absolute values given in brackets. . . . . 162

6.3 Characteristic price profiles of retail electricity for workdays and weekends supplemented by 0.188 €/kWh fees and 19 % taxes [299]. . 163

6.4	Cloud-based architecture of the charging optimization scheme as realized in the <i>Energy Lab 2.0</i> [303], simplified according to [300]. . . . .	164
6.5	Three-page user interface of the charging optimization scheme as realized in the <i>Energy Lab 2.0</i> [303], according to [300,301]. . . . .	165
6.6	Impressions of real-world charging optimization setup in the <i>Energy Lab 2.0</i> [303]. . . . .	166
6.7	Comparison of normalized operating cost and its components of 45 real charging events in three modes: standard charging without optimization (Mode I); energy cost optimization, aging costs calculated afterwards (Mode II); energy and aging cost optimization (Mode III). . . . .	167
6.8	Profiles of the electricity price $\epsilon$ with different spreads $\gamma$ , according to [299]; the green price profile is obtained by stretching the orange one while keeping a constant mean price $\epsilon_{\text{mean}} = 0.24 \text{ €/kWh}$ . . . . .	172
6.9	Profiles of battery energy (normalized as SOC) over time for price profiles with different electricity spreads $\gamma$ (see Figure 6.8). . . . .	172
6.10	Power profiles over time of an exemplary charging event, if assuming constant battery temperature (orange), and if using a data-driven thermal model as described in Section 3.2.2.2 (blue). . . . .	174
6.11	Examples of real charging processes (default in black, optimized in green/yellow, actually measured charging power in blue) controlled and adapted via the user interface of the charging optimization scheme realized in the <i>Energy Lab 2.0</i> [303]; charging event of a typical workday (a); over-night charging with recalculation and early departure (b); according to [300]. . . . .	176
A.1	Node $t$ of decision tree $\mathcal{T} \in \mathcal{R}$ performing a binary split of $N_t$ data samples into subsets with $N_{t_L}$ and $N_{t_R}$ samples. . . . .	184



# List of Tables

1.1	Literature review on <b>V2G</b> services grouped by categorical characteristics of the services [53] . . . . .	9
1.2	Literature review on electric battery models grouped by model categories. . . . .	15
1.3	Literature review on thermal battery models grouped by model categories. . . . .	16
1.4	Literature review on battery degradation grouped by model categories. . . . .	18
1.5	Concerns and future expectations towards <b>EV</b> usage according to recent customer surveys [48,147,148]. . . . .	20
1.6	Literature review on <b>EV</b> usage, i.e., driving and charging habits, mobility prediction, and user acceptance. . . . .	22
1.7	Literature review on <b>EV</b> charging assistants and strategies grouped by stakeholder perspectives. . . . .	27
2.1	Possible alternatives and their respective properties for design features of the charging assistant shown in <b>Figure 2.1</b> . . . . .	32
3.1	Selection of consumption model features according to univariate feature relevance based on the <b>SPEARMAN</b> -correlation coefficient (see <b>Section A.1.1</b> ) in relation to the specific energy consumption $\Gamma$ and manual screening. . . . .	52
3.2	Tested and selected hyperparameters for <b>ANN</b> regression model to calculate $\hat{\Gamma}_{DD,MLP}$ . . . . .	55
3.3	Result of preliminary tests to examine different architectures of the <b>ANN</b> model to calculate $\hat{\Gamma}_{DD,MLP}$ . . . . .	56
3.4	Selected hyperparameters for random forest regression model to calculate $\hat{\Gamma}_{DD,RF}$ . . . . .	57
3.5	Comparison of baseline model and data-driven models according to the error metrics <b>MSE</b> $q_{MSE}$ (in $(\text{kWh}/\text{km})^2$ ), <b>RMSE</b> $q_{RMSE}$ (in $\text{kWh}/\text{km}$ ), <b>MAE</b> $q_{MAE}$ (in $\text{kWh}/\text{km}$ ), and <b>RMAE</b> $q_{RMAE}$ , on the independent test data set $\mathcal{J}_{\text{test}}$ . . . . .	57

3.6 Selection of thermal battery model features according to univariate feature relevance based on the SPEARMAN-correlation coefficient (see Section A.1.1) in relation to the change of battery temperature  $\Delta\Theta$  and manual screening. . . . . 67

3.7 Tested and selected hyperparameters for ANN regression model to calculate  $\Delta\hat{\Theta}_{DD,MLP}$ . . . . . 69

3.8 Result of three preliminary test runs to examine different architectures of the ANN model to calculate  $\Delta\hat{\Theta}_{DD,MLP}$ ; for evaluation, the  $R^2$ -Score  $q_{R^2}$  of estimated and true values (see (A.6)) is used. . . . . 70

3.9 Local and global error of thermal battery models (Section 3.2). . . . . 75

4.1 Parameters and selected values for two-stage density-based spatial clustering algorithm (see Section 4.2.2) to determine EV users' known locations. . . . . 93

4.2 GINI importance  $q_{Gini}$  using a random forest model for all features when predicting the next trip destination on the real data  $\mathcal{J}_{real}$ . Higher numbers indicate that the feature causes a larger mean decrease in impurity and is therefore more relevant to the random forest [276,277]. . . . . 99

4.3 Comparison of baseline and random forest model (see Section 4.3.2) accuracy  $q_{ACC}$  for different scopes (start locations) when predicting the next location to visit on the real data set  $\mathcal{J}_{real}$  using a five-fold cross-validation. . . . . 103

4.4 Comparison of baseline and random forest model (see Section 4.3.2) accuracy  $q_{ACC}$  for different scopes (start locations) when predicting the next location to visit on the synthetic data set  $\mathcal{J}_{syn}$  using a five-fold cross-validation. . . . . 104

4.5 GINI importance  $q_{Gini}$  using a default random forest for all features when considering both Label A and Label B on the real data  $\mathcal{J}_{real}$ . Higher numbers indicate that the feature causes a larger mean decrease in impurity and is therefore more relevant to the random forest [276,277]. . . . . 112

4.6 Per-location accuracy ( $q_{p,60}$ , mean of five-fold cross-validation) of estimated parking duration (Label A) and next departure time (Label B), on the real data set  $\mathcal{J}_{real}$  using the default random forest model. . . . . 118

4.7 Comparison of mean error metrics for predictive models to estimate parking duration (Label A) based on the real data set  $\mathcal{J}_{real}$  with a five-fold cross-validation. . . . . 119

4.8	Comparison of mean error metrics for predictive models to estimate the next departure time (Label B) based on the real data set $\mathcal{J}_{\text{real}}$ with a five-fold cross-validation. . . . .	119
4.9	Comparison of time ensemble model accuracy for different scopes (known locations with corresponding number of data samples) when estimating parking duration (Label A) on the real data set $\mathcal{J}_{\text{real}}$ with a five-fold cross-validation. . . . .	121
5.1	Information on known locations $l \in \mathcal{L}_{\text{syn}}$ (aliased with capital letters A to H) comprising the walking distance $w$ between known location and the closest charging station, its mean electricity price $\epsilon$ , and maximum available charging power $\bar{p}$ ; dashes indicated unavailable charging opportunities at a location; used for validating the charging scheduling scheme (see Section 5.1 and Section 5.2). . . . .	140
5.2	Globally constant parameters and their selected values for the MONTE CARLO simulation to validate charging scheduling. . . . .	143
5.3	Parameters that are varied in a specified range of values for the MONTE CARLO simulation to validate charging scheduling. . . . .	143
5.4	Comparison of alternative charging strategies (see Section 5.3.2) with the proposed scheduling scheme (Section 5.2) in terms of reducing the number of plug-in events, battery calendar aging, and the mean electricity price; negative values indicate a benefit over the proposed strategy. . . . .	149
6.1	Globally constant parameters and their selected values to simulate charging scheduling. . . . .	161
6.2	Comparison of charging optimization (Mode III, see Section 6.2.1) for different future scenarios in terms of battery value loss (according to [92]). . . . .	169
6.3	Cost comparison of optimized charging (Mode III, see Section 6.2.1) for three electricity price profiles (see Figure 6.8) with different spreads $\gamma$ according to (6.8): regular workday ( $\gamma = 0.054$ ), example 05.05.2021 ( $\gamma = 0.418$ ), and stretched example 05.05.2021 ( $\gamma = 0.5$ ). . . . .	171
6.4	Relative cost savings of optimized charging over default charging (plug in and charge immediately) for three different scenarios, tested with the real-world implementation of charging optimization (see Section 6.2.2), according to [300]. . . . .	175



# List of Own Publications

The following, chronologically ordered list presents all publications to which the author of this thesis contributed:

- [271] **K. Schwenk**, M. Faix, R. Mikut, V. Hagenmeyer, and R. R. Appino, “On calendar-based scheduling for user-friendly charging of plug-in electric vehicles,” in *2019 IEEE 2nd Connected and Automated Vehicles Symposium (CAVS)*. Honolulu, HI, USA, IEEE, 2019, pp. 1–5.
- [243] **K. Schwenk**, T. Harr, R. Großmann, R. R. Appino, V. Hagenmeyer, and R. Mikut, “Data-driven charging strategies for grid-beneficial, customer-oriented and battery-preserving electric mobility,” in *Proceedings 29th Workshop Computational Intelligence*. Dortmund, Germany, KIT Scientific Publishing, 2019, pp. 73–93.
- [275] **K. Schwenk**, F. Hepperle, and F. Herzog, “Verfahren zur Planung von Ladevorgängen für ein Elektrofahrzeug,” International Patent Application WO 2021259677 A1, 2019.
- [300] S. Meisenbacher, **K. Schwenk**, J. Galenzowski, S. Waczowicz, R. Mikut, and V. Hagenmeyer, “Poster: Smart Charging of Electric Vehicles with Cloud-Based Optimization and a Lightweight User Interface: A Real-World Application in the Energy Lab 2.0,” in *e-Energy '21: Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. Virtual Event, Italy, Association for Computing Machinery, 2021, pp. 284–285.
- [250] **K. Schwenk**, S. Meisenbacher, B. Briegel, T. Harr, V. Hagenmeyer, and R. Mikut, “Integrating Battery Aging in the Optimization for Bidirectional Charging of Electric Vehicles,” *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5135–5145, 2021.
- [242] **K. Schwenk** and J. Frick, “Verfahren zum inkrementellen Kondensieren von Positionsdaten elektrischer Fahrzeuge,” German Patent Application 2021ID01 931, 2021.

- [297] **K. Schwenk**, F. Hepperle, and T. Harr, "Verfahren zur Schätzung des zukünftigen Anteils an erneuerbar erzeugter Energie im Stromnetz zum Kohlendioxid-reduzierten Laden von batterie- und hybridelektrischen Fahrzeugen," German Patent Request 2021ID01 226, 2021.
- [301] S. Meisenbacher, **K. Schwenk**, J. Galenzowski, S. Waczowicz, R. Mikut, and V. Hagenmeyer, "A Lightweight User Interface for Smart Charging of Electric Vehicles: A Real-World Application," in *2021 9th International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*. Sarawak, Malaysia, IEEE, 2021, pp. 57–61.
- [270] **K. Schwenk**, K. Phipps, B. Briegel, V. Hagenmeyer, and R. Mikut, "A Benchmark for Parking Duration Prediction of Electric Vehicles for Smart Charging Applications," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. Orlando, FL, USA, IEEE, 2021, pp. 1–8.
- [282] **K. Schwenk**, V. Hagenmeyer, and R. Mikut, "Multi-Day Stochastic Scheduling of Electric Vehicle Charging for Reliability and Convenience," 2022, submitted.
- [283] **K. Schwenk**, German Patent Request 2022ID00 158, 2022, subject to nondisclosure.
- [281] K. Phipps, **K. Schwenk**, B. Briegel, R. Mikut and V. Hagenmeyer, "Uncertainty in Parking Duration Prediction of Electric Vehicles for Smart Charging Applications," 2022, submitted.

# Bibliography

- [1] S. Shafiee and E. Topal, "When will fossil fuel reserves be diminished?" *Energy Policy*, vol. 37, no. 1, pp. 181–189, 2009.
- [2] S. J. Davis, K. Caldeira, and H. D. Matthews, "Future CO<sub>2</sub> Emissions and Climate Change from Existing Energy Infrastructure," *Science*, vol. 329, no. 5997, pp. 1330–1333, 2010.
- [3] Statista, "Distribution of oil demand in the OECD in 2017 by sector," 2021, accessed on 04.12.2021. [Online]. Available: <https://www.statista.com/statistics/307194/top-oil-consuming-sectors-worldwide/>
- [4] S. Gottwalt, A. Schuller, C. Flath, H. Schmeck, and C. Weinhardt, "Assessing load flexibility in smart grids: Electric vehicles for renewable energy integration," in *2013 IEEE Power & Energy Society General Meeting*. Vancouver, BC, Canada, IEEE, 2013, pp. 1–5.
- [5] S. Steuer, J. Gärttner, A. Schuller, H. Schmeck, and C. Weinhardt, "Economic Evaluation of Local Photovoltaic Generation in Electric Vehicle Car Parks," in *Tagungsband des VDE-Kongress 2014*. Frankfurt, Germany, VDE-Verlag, 2014.
- [6] R. Cornell, "The Climate Change Mitigation Potential of Electric Vehicles as a Function of Renewable Energy," *The International Journal of Climate Change: Impacts and Responses*, vol. 11, no. 1, pp. 15–24, 2019.
- [7] P. A. Owusu and S. Asumadu-Sarkodie, "A review of renewable energy sources, sustainability issues and climate change mitigation," *Cogent Engineering*, vol. 3, no. 1, p. 1167990, 2016.
- [8] A. Hoekstra, "The Underestimated Potential of Battery Electric Vehicles to Reduce Emissions," *Joule*, vol. 3, no. 6, pp. 1412–1414, 2019.
- [9] C. Wulff, A. Merkel, and N. Röttgen, "Dreizehntes Gesetz zur Änderung des Atomgesetzes," 2011, accessed on 04.05.2020. [Online]. Available: <https://dip.bundestag.de/vorgang/dreizehntes-gesetz-zur-C3%A4nderung-des-atomgesetzes/36232>

- [10] Bundesministerium für Wirtschaft und Energie (BMWi), “Kommission Wachstum, Strukturwandel und Beschäftigung,” 2020, accessed on 28.01.2020. [Online]. Available: <https://www.bmwi.de/>
- [11] T. Brown, D. Schlachtberger, A. Kies, S. Schramm, and M. Greiner, “Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system,” *Energy*, vol. 160, pp. 720–739, 2018.
- [12] S. Naumann and P. Bretschneider, “Planning Strategy for a Zero Carbon Cross Sectoral Energy System,” in *ETG Congress 2021*. Virtual Event, VDE-Verlag, 2021, pp. 1–6.
- [13] D. Keles, D. Möst, and W. Fichtner, “The development of the German energy market until 2030 - A critical survey of selected scenarios,” *Energy Policy*, vol. 39, no. 2, pp. 812–825, 2011.
- [14] International Energy Agency, “Data and Statistics,” 2021, accessed on 03.12.2021. [Online]. Available: <https://www.iea.org/data-and-statistics>
- [15] L. Mehigan, J. Deane, B. Ó. Gallachóir, and V. Bertsch, “A review of the role of distributed generation (DG) in future electricity systems,” *Energy*, vol. 163, pp. 822–836, 2018.
- [16] R. R. Appino, M. Munoz-Ortiz, J. A. G. Ordiano, R. Mikut, V. Hagenmeyer, and T. Faulwasser, “Reliable Dispatch of Renewable Generation via Charging of Time-Varying PEV Populations,” *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1558–1568, 2019.
- [17] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, “Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks,” *Applied Energy*, vol. 162, pp. 218–230, 2016.
- [18] S. Klaiber, S. Waczowicz, I. Konotop, D. Westermann, R. Mikut, and P. Bretschneider, “Prognose für preisbeeinflusstes Verbrauchsverhalten,” *at - Automatisierungstechnik*, vol. 65, no. 3, pp. 179–188, 2017.
- [19] S. Weitemeyer, D. Kleinhans, T. Vogt, and C. Agert, “Integration of Renewable Energy Sources in future power systems: The role of storage,” *Renewable Energy*, vol. 75, pp. 14–20, 2015.
- [20] S. Weitemeyer, D. Kleinhans, L. Wienholt, T. Vogt, and C. Agert, “A European Perspective: Potential of Grid and Storage for Balancing Renewable Power Systems,” *Energy Technology*, vol. 4, no. 1, pp. 114–122, 2016.

- [21] R. Jarvis and P. Moses, "Smart Grid Congestion Caused by Plug-in Electric Vehicle Charging," in *2019 IEEE Texas Power and Energy Conference (TPEC)*. College Station, TX, USA, IEEE, 2019, pp. 1–5.
- [22] D. Beyer, S. Ruhe, P. Bretschneider, and J. Ehrhardt, "Results of a Monte Carlo based risk analysis on the impact of PEVs to the distribution grid," in *2016 IEEE International Energy Conference (ENERGYCON)*. Leuven, Belgium, IEEE, 2016, pp. 1–7.
- [23] O. Warweg, F. Schaller, S. Ritter, and P. Bretschneider, "Technical and regulatory challenges for the integration of electric vehicles into the energy system," in *2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies*. Manchester, UK, IEEE, 2011, pp. 1–8.
- [24] L. Held, A. März, D. Krohn, J. Wirth, M. Zimmerlin, M. R. Suriyah, T. Leibfried, P. Jochem, and W. Fichtner, "The Influence of Electric Vehicle Charging on Low Voltage Grids with Characteristics Typical for Germany," *World Electric Vehicle Journal*, vol. 10, no. 4, p. 88, 2019.
- [25] Y. Noorollahi, A. Aligholian, and A. Golshanfard, "Stochastic energy modeling with consideration of electrical vehicles and renewable energy resources- A review," *Journal of Energy Management and Technology*, vol. 4, no. 1, pp. 13–26, 2020.
- [26] W. Funk and B. Strigel, "From Vehicle Manufacturer to Mobility Service Provider – Business Challenges of Electromobility for OEM and Supplier," in *CTI SYMPOSIUM 2018*. Berlin, Germany, Springer, 2019, pp. 143–152.
- [27] G. Berckmans, M. Messagie, J. S. N. O. L. Vanhaverbeke, and J. V. Mierlo, "Cost Projection of State of the Art Lithium-Ion Batteries for Electric Vehicles up to 2030," *Energies*, vol. 10, no. 9, p. 1314, 2017.
- [28] E. D. GmbH, "CTI - car training institute," 2019, accessed on 03.12.2019. [Online]. Available: <https://drivetrain-symposium.world/typical-manufacturing-costs-of-a-battery-electric-vehicle-did-you-know-3/>
- [29] C. Will and A. Schuller, "Understanding user acceptance factors of electric vehicle smart charging," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 198–214, 2016.
- [30] R. Gottumukkala, R. Merchant, A. Tauzin, K. Leon, A. Roche, and P. Darby, "Cyber-physical System Security of Vehicle Charging Stations," in *2019 IEEE*

- Green Technologies Conference(GreenTech)*. Lafayette, LA, USA, IEEE, 2019, pp. 1–5.
- [31] A. J. Schwab, *Elektroenergiesysteme*. Berlin/Heidelberg: Springer, 2012.
- [32] International Energy Agency, “Global EV outlook 2017: Two million and counting,” 2017, accessed on 18.04.2019. [Online]. Available: <https://www.iea.org/reports/global-ev-outlook-2017>
- [33] —, “Global EV Outlook 2018: Towards cross-modal electrification,” 2018, accessed on 18.04.2019. [Online]. Available: <https://www.iea.org/reports/global-ev-outlook-2018>
- [34] B. Khaki, Y.-W. Chung, C. Chu, and R. Gadh, “Probabilistic Electric Vehicle Load Management in Distribution Grids,” in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*. Detroit, MI, USA, IEEE, 2019, pp. 1–6.
- [35] D. Danner and H. de Meer, “State Estimation in the Power Distribution System,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 46, no. 3, pp. 86–88, 2019.
- [36] M. Eider, M. Stolba, D. Sellner, A. Berl, R. Basmadjian, H. de Meer, S. Klingert, T. Schulze, F. Kutzner, and C. Kacperski, “Seamless Electromobility,” in *e-Energy '17: Proceedings of the Eighth International Conference on Future Energy Systems*. Shatin, Hong Kong, Association for Computing Machinery, 2017, pp. 316–321.
- [37] A. G. Boulanger, A. C. Chu, S. Maxx, and D. L. Waltz, “Vehicle Electrification: Status and Issues,” *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1116–1138, 2011.
- [38] M. Lützenberger, N. Masuch, T. Küster, D. Freund, M. Voß, C.-E. Hrabia, D. Pozo, J. Fährndrich, F. Trollmann, J. Keiser, and S. Albayrak, “A common approach to intelligent energy and mobility services in a smart city environment,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 6, no. 3, pp. 337–350, 2015.
- [39] C. Crozier, T. Morstyn, M. Deakin, and M. McCulloch, “The case for Bi-directional charging of electric vehicles in low voltage distribution networks,” *Applied Energy*, vol. 259, p. 114214, 2020.
- [40] S. Nazari, F. Borrelli, and A. Stefanopoulou, “Electric Vehicles for Smart Buildings: A Survey on Applications, Energy Management Methods, and Battery Degradation,” *Proceedings of the IEEE*, vol. 109, no. 6, pp. 1128–1144, 2020.

- [41] G. Saldaña, J. I. S. Martin, I. Zamora, F. J. Asensio, and O. Oñederra, "Electric Vehicle into the Grid: Charging Methodologies Aimed at Providing Ancillary Services Considering Battery Degradation," *Energies*, vol. 12, no. 12, p. 2443, 2019.
- [42] O. Sundström and C. Binding, "Flexible Charging Optimization for Electric Vehicles Considering Distribution Grid Constraints," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 26–37, 2012.
- [43] K. Seddig, P. Jochem, and W. Fichtner, "Two-stage stochastic optimization for cost-minimal charging of electric vehicles at public charging stations with photovoltaics," *Applied Energy*, vol. 242, pp. 769–781, 2019.
- [44] C. Will, P. Jochem, S. Pfahl, and W. Fichtner, "Economic and sustainability-potential of carbon-neutral charging services for electric vehicle customers," in *Proceedings of the 30th International Electric Vehicle Symposium and Exhibition, EVS30*. Stuttgart, Germany, 2017.
- [45] R. Das, K. Thirugnanam, P. Kumar, R. Lavudiya, and M. Singh, "Mathematical Modeling for Economic Evaluation of Electric Vehicle to Smart Grid Interaction," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 712–721, 2014.
- [46] M. Hayn, A. Zander, W. Fichtner, S. Nickel, and V. Bertsch, "The impact of electricity tariffs on residential demand side flexibility: results of bottom-up load profile modeling," *Energy Systems*, vol. 9, no. 3, pp. 759–792, 2018.
- [47] A. W. Thompson, "Economic implications of lithium ion battery degradation for Vehicle-to-Grid (V2X) services," *Journal of Power Sources*, vol. 396, pp. 691–709, 2018.
- [48] Daimler AG, "Evaluation E-Mobility Concepts: Full Report," Sindelfingen, Germany, 2019, internal document.
- [49] VISPIRON SYSTEMS GmbH, "V2G Roadmap," 2020, accessed on 02.02.2020. [Online]. Available: <https://systems.vispiron.com/>
- [50] S. Arens, K. Derendorf, F. Schuldt, K. V. Maydell, and C. Agert, "Effect of EV Movement Schedule and Machine Learning-Based Load Forecasting on Electricity Cost of a Single Household," *Energies*, vol. 11, no. 2913, pp. 1–19, 2018.

- [51] T. U. Solanke, V. K. Ramachandaramurthy, J. Y. Yong, J. Pasupuleti, P. Kasiathan, and A. Rajagopalan, "A review of strategic charging–discharging control of grid-connected electric vehicles," *Journal of Energy Storage*, vol. 28, p. 101193, 2020.
- [52] E. Sortomme and M. A. El-Sharkawi, "Optimal Scheduling of Vehicle-to-Grid Energy and Ancillary Services," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 351–359, 2012.
- [53] N. S. Pearre and H. Ribberink, "Review of research on V2X technologies, strategies, and operations," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 61–70, 2019.
- [54] F. Wu and R. Sioshansi, "A stochastic operational model for controlling electric vehicle charging to provide frequency regulation," *Transportation Research Part D: Transport and Environment*, vol. 67, pp. 475–490, 2019.
- [55] J. Donadee and M. Ilić, "Stochastic co-optimization of charging and frequency regulation by electric vehicles," in *Proc. North American Power Symp. (NAPS)*. Champaign, IL, USA, IEEE, 2012, pp. 1–6.
- [56] M. Rahmani-Andebili, "Spinning Reserve Capacity Provision by the Optimal Fleet Management of Plug-In Electric Vehicles Considering the Technical and Social Aspects," in *Planning and Operation of Plug-In Electric Vehicles*. Springer International Publishing, 2019, pp. 49–74.
- [57] P. Aliasghari, B. Mohammadi-Ivatloo, M. Alipour, M. Abapour, and K. Zare, "Optimal scheduling of plug-in electric vehicles and renewable micro-grid in energy and reserve markets considering demand response program," *Journal of Cleaner Production*, vol. 186, pp. 293–303, 2018.
- [58] O. Beade, Y. He, and M. Hennebel, "Introducing decentralized EV charging coordination for the voltage regulation," in *IEEE PES ISGT Europe 2013*. Lyngby, Denmark, IEEE, 2013, pp. 1–5.
- [59] K. Knezović, M. Marinelli, R. J. Møller, P. B. Andersen, C. Træholt, and F. Sossan, "Analysis of voltage support by electric vehicles and photovoltaic in a real Danish low voltage network," in *2014 49th International Universities Power Engineering Conference (UIPEC)*. Cluj-Napoca, Romania, IEEE, 2014, pp. 1–6.

- [60] A. Rajaei, M. Jooshaki, M. Fotuhi-Firuzabad, and M. Moeini-Aghaie, "Enhancing Power Distribution System Flexibility Using Electric Vehicle Charging Management," in *2019 27th Iranian Conference on Electrical Engineering (ICEE)*. Yazd, Iran, IEEE, 2019, pp. 782–786.
- [61] P. Denholm and W. Short, "An evaluation of utility system impacts and benefits of optimally dispatched plug-in hybrid electric vehicles," National Renewable Energy Laboratory, Golden, CO, USA, 2006.
- [62] D. F. R. Melo, A. Trippe, H. B. Gooi, and T. Massier, "Robust Electric Vehicle Aggregation for Ancillary Service Provision Considering Battery Aging," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1728–1738, 2018.
- [63] H. Gan and C. Zheng, "An electric vehicle operation optimization method based on demand-side management," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 23, pp. 1–12, 2019.
- [64] F. Erden, M. C. Kisacikoglu, and N. Erdogan, "Adaptive V2G Peak Shaving and Smart Charging Control for Grid Integration of PEVs," *Electric Power Components and Systems*, vol. 46, no. 13, pp. 1494–1508, 2018.
- [65] S. Khemakhem, M. Rekik, and L. Krichen, "Double layer home energy supervision strategies based on demand response and plug-in electric vehicle control for flattening power load curves in a smart grid," *Energy*, vol. 167, pp. 312–324, 2019.
- [66] R. H. Ashique, Z. Salam, M. J. B. A. Aziz, and A. R. Bhatti, "Integrated photovoltaic-grid dc fast charging system for electric vehicle: A review of the architecture and control," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 1243–1257, 2017.
- [67] H. Lund and W. Kempton, "Integration of renewable energy into the transport and electricity sectors through V2G," *Energy policy*, vol. 36, no. 9, pp. 3578–3587, 2008.
- [68] M. Ghofrani, E. Detert, N. N. Hosseini, A. Arabali, N. Myers, and P. Ngin, "V2G Services for Renewable Integration," in *Modeling and Simulation for Electric Vehicle Applications*. Rijeka, Croatia: InTech Open, 2016.
- [69] H. N. Nguyen, C. Zhang, and M. A. Mahmud, "Optimal coordination of G2V and V2G to support power grids with high penetration of renewable energy," *IEEE Transactions on Transportation Electrification*, vol. 1, no. 2, pp. 188–195, 2015.

- [70] D. Dallinger and M. Wietschel, "Grid integration of intermittent renewable energy sources using price-responsive plug-in electric vehicles," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3370–3382, 2012.
- [71] M. R. Mozafar, M. H. Moradi, and M. H. Amini, "A simultaneous approach for optimal allocation of renewable energy sources and electric vehicle charging stations in smart grids based on improved GA-PSO algorithm," *Sustainable Cities and Society*, vol. 32, pp. 627–637, 2017.
- [72] F. Salah, A. Schuller, M. Maurer, and C. Weinhardt, "Pricing of demand flexibility: Exploring the impact of Electric Vehicle customer diversity," in *2016 13th International Conference on the European Energy Market (EEM)*. Porto, Portugal, IEEE, 2016, pp. 1–5.
- [73] C. Zhou, K. Qian, M. Allan, and W. Zhou, "Modeling of the cost of EV battery wear due to V2G application in power systems," *IEEE Transactions on Energy Conversion*, vol. 26, no. 4, pp. 1041–1050, 2011.
- [74] G. Graber, V. Calderaro, P. Mancarella, and V. Galdi, "Two-stage stochastic sizing and packetized energy scheduling of BEV charging stations with quality of service constraints," *Applied Energy*, vol. 260, p. 114262, 2020.
- [75] C. M. Flath, J. P. Ilg, S. Gottwalt, H. Schmeck, and C. Weinhardt, "Improving Electric Vehicle Charging Coordination Through Area Pricing," *Transportation Science*, vol. 48, no. 4, pp. 619–634, 2014.
- [76] H. J. Monfared and A. Ghasemi, "Retail electricity pricing based on the value of electricity for consumers," *Sustainable Energy, Grids and Networks*, vol. 18, p. 100205, 2019.
- [77] M. Carrion, "Determination of the Selling Price Offered by Electricity Suppliers to Electric Vehicle Users," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6655–6666, 2019.
- [78] J. Leadbetter and L. Swan, "Battery storage system for residential electricity peak demand shaving," *Energy and Buildings*, vol. 55, pp. 685–692, 2012.
- [79] S. Karrari, N. Ludwig, V. Hagenmeyer, and M. Noe, "A Method for Sizing Centralised Energy Storage Systems Using Standard Patterns," in *IEEE Milan PowerTech*. Milan, Italy, IEEE, 2019, pp. 1–6.
- [80] W. Kempton and J. Tomić, "Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy," *Journal of Power Sources*, vol. 144, no. 1, pp. 280–294, 2005.

- [81] W. Kempton, V. Udo, K. Huber, K. Komara, S. Letendre, S. Baker, D. Brunner, and N. Pearre, "A test of vehicle-to-grid (V2G) for energy storage and frequency regulation in the PJM system," Results from an Industry-University Research Partnership, University of Delaware, Newark, DE, USA, 2008.
- [82] J. Mullan, D. Harries, T. Bräunl, and S. Whitely, "The technical, economic and commercial viability of the vehicle-to-grid concept," *Energy Policy*, vol. 48, pp. 394–406, 2012.
- [83] K. M. Tan, V. K. Ramachandaramurthy, and J. Y. Yong, "Integration of electric vehicles in smart grid: A review on vehicle to grid technologies and optimization techniques," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 720–732, 2016.
- [84] F. Mwasilu, J. J. Justo, E.-K. Kim, T. D. Do, and J.-W. Jung, "Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 501–516, 2014.
- [85] L. Liu, F. Kong, X. Liu, Y. Peng, and Q. Wang, "A review on electric vehicles interacting with renewable energy in smart grid," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 648–661, 2015.
- [86] Y. Benomar, M. El Baghdadi, O. Hegazy, Y. Yang, M. Messagie, and J. Van Mierlo, "Design and modeling of V2G inductive charging system for light-duty Electric Vehicles," in *2017 Twelfth International Conference on Ecological Vehicles and Renewable Energies (EVER)*. Monte Carlo, Monaco, IEEE, 2017, pp. 1–7.
- [87] J. C. Mankins, "Technology Readiness Level: A White Paper," Advanced Concepts Office, Office of Space Access and Technology, NASA, Washington, DC, USA, 1995.
- [88] Y. Ligen, H. Vrabel, and H. Girault, "Mobility from Renewable Electricity: Infrastructure Comparison for Battery and Hydrogen Fuel Cell Vehicles," *World Electric Vehicle Journal*, vol. 9, no. 1, p. 3, 2018.
- [89] T. F. Fuller, M. Doyle, and J. Newman, "Simulation and optimization of the dual lithium ion insertion cell," *Journal of the Electrochemical Society*, vol. 141, no. 1, pp. 1–10, 1994.

- [90] B. Huang, Z. Pan, X. Su, and L. An, "Recycling of lithium-ion batteries: Recent advances and perspectives," *Journal of Power Sources*, vol. 399, pp. 274–286, 2018.
- [91] C. Pillot, "The rechargeable battery market past and future," in *Batteries 2011*. Cannes, France, 2011.
- [92] S. I. Sun, A. J. Chipperfield, M. Kiaee, and R. G. A. Wills, "Effects of market dynamics on the time-evolving price of second-life electric vehicle batteries," *Journal of Energy Storage*, vol. 19, pp. 41–51, 2018.
- [93] I.-Y. L. Hsieh, M. S. Pan, Y.-M. Chiang, and W. H. Green, "Learning only buys you so much: Practical limits on battery price reduction," *Applied Energy*, vol. 239, pp. 218–224, 2019.
- [94] M. Brill, "Entwicklung und Implementierung einer neuen Onboard-Diagnosemethode für Lithium-Ionen-Fahrzeugbatterien," Dissertation, Universität Ulm, Ulm, Germany, 2012.
- [95] K. Liu, K. Li, Q. Peng, and C. Zhang, "A brief review on key technologies in the battery management system of electric vehicles," *Frontiers of Mechanical Engineering*, vol. 14, no. 1, pp. 47–64, 2018.
- [96] M. Hannan, M. Lipu, A. Hussain, and A. Mohamed, "A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations," *Renewable and Sustainable Energy Reviews*, vol. 78, pp. 834–854, 2017.
- [97] H. He, X. Zhang, R. Xiong, Y. Xu, and H. Guo, "Online model-based estimation of state-of-charge and open-circuit voltage of lithium-ion batteries in electric vehicles," *Energy*, vol. 39, no. 1, pp. 310–318, 2012.
- [98] Y. Xing, W. He, M. Pecht, and K. L. Tsui, "State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures," *Applied Energy*, vol. 113, pp. 106–115, 2014.
- [99] A. Kustiman, B. R. Dewangga, O. Wahyunggoro, and A. I. Cahyadi, "Battery State of Charge Estimation Based on Coulomb Counting Combined with Recursive Least Square and PI Controller," in *2019 International Conference on Information and Communications Technology (ICOIACT)*. Yogyakarta, Indonesia, IEEE, 2019, pp. 663–668.

- [100] L. W. Juang, P. J. Kollmeyer, R. Zhao, T. M. Jahns, and R. D. Lorenz, "Coulomb counting state-of-charge algorithm for electric vehicles with a physics-based temperature dependent battery model," in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*. Montreal, QC, Canada, IEEE, 2015, pp. 5052–5059.
- [101] L. He and D. Guo, "An Improved Coulomb Counting Approach Based on Numerical Iteration for SOC Estimation With Real-Time Error Correction Ability," *IEEE Access*, vol. 7, pp. 74 274–74 282, 2019.
- [102] M. A. Rahman, S. Anwar, and A. Izadian, "Electrochemical model parameter identification of a lithium-ion battery using particle swarm optimization method," *Journal of Power Sources*, vol. 307, pp. 86–97, 2016.
- [103] Y. Zheng, W. Gao, X. Han, M. Ouyang, L. Lu, and D. Guo, "An accurate parameters extraction method for a novel on-board battery model considering electrochemical properties," *Journal of Energy Storage*, vol. 24, p. 100745, 2019.
- [104] X. Zhang, J. Lu, S. Yuan, J. Yang, and X. Zhou, "A novel method for identification of lithium-ion battery equivalent circuit model parameters considering electrochemical properties," *Journal of Power Sources*, vol. 345, pp. 21–29, 2017.
- [105] Y. Li, M. Vilathgamuwa, T. Farrell, S. S. Choi, N. T. Tran, and J. Teague, "A physics-based distributed-parameter equivalent circuit model for lithium-ion batteries," *Electrochimica Acta*, vol. 299, pp. 451–469, 2019.
- [106] Z. Deng, L. Yang, Y. Cai, H. Deng, and L. Sun, "Online available capacity prediction and state of charge estimation based on advanced data-driven algorithms for lithium iron phosphate battery," *Energy*, vol. 112, pp. 469–480, 2016.
- [107] X. Gong, R. Xiong, and C. Mi, "A data-driven bias correction method based lithiumion battery modeling approach for electric vehicle applications," *IEEE Transactions on Industry Applications*, vol. 52, no. 2, pp. 1759–1765, 2015.
- [108] Q.-K. Wang, Y.-J. He, J.-N. Shen, Z.-F. Ma, and G.-B. Zhong, "A unified modeling framework for lithium-ion batteries: An artificial neural network based thermal coupled equivalent circuit model approach," *Energy*, vol. 138, pp. 118–132, 2017.
- [109] C.-H. Lee and C.-H. Wu, "A Novel Big Data Modeling Method for Improving Driving Range Estimation of EVs," *IEEE Access*, vol. 3, pp. 1980–1993, 2015.

- [110] E. Chemali, "Intelligent State-of-Charge and State-of-Health Estimation Framework for Li-ion Batteries in Electrified Vehicles using Deep Learning Techniques," Dissertation, McMaster University, Hamilton, Canada, 2018.
- [111] R. Xiong, J. Cao, Q. Yu, H. He, and F. Sun, "Critical Review on the Battery State of Charge Estimation Methods for Electric Vehicles," *IEEE Access*, vol. 6, pp. 1832–1843, 2017.
- [112] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," *Journal of Power Sources*, vol. 226, pp. 272–288, 2013.
- [113] K. S. Ng, C.-S. Moo, Y.-P. Chen, and Y.-C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Applied Energy*, vol. 86, no. 9, pp. 1506–1511, 2009.
- [114] S. Pelletier, O. Jabali, G. Laporte, and M. Veneroni, "Battery degradation and behaviour for electric vehicles: Review and numerical analyses of several models," *Transportation Research Part B: Methodological*, vol. 103, pp. 158–187, 2017.
- [115] S. Sun, J. Zhang, J. Bi, and Y. Wang, "A Machine Learning Method for Predicting Driving Range of Battery Electric Vehicles," *Journal of Advanced Transportation*, vol. 2019, pp. 1–14, 2019.
- [116] F. Richter, S. Kjelstrup, P. J. Vie, and O. S. Burheim, "Thermal conductivity and internal temperature profiles of Li-ion secondary batteries," *Journal of Power Sources*, vol. 359, pp. 592–600, 2017.
- [117] H. Dai, L. Zhu, J. Zhu, X. Wei, and Z. Sun, "Adaptive Kalman filtering based internal temperature estimation with an equivalent electrical network thermal model for hard-cased batteries," *Journal of Power Sources*, vol. 293, pp. 351–365, 2015.
- [118] K. Yoo and J. Kim, "Thermal behavior of full-scale battery pack based on comprehensive heat-generation model," *Journal of Power Sources*, vol. 433, p. 226715, 2019.
- [119] M. Guo, G.-H. Kim, and R. E. White, "A three-dimensional multi-physics model for a Li-ion battery," *Journal of Power Sources*, vol. 240, pp. 80–94, 2013.
- [120] S. Panchal, I. Dincer, M. Agelin-Chaab, R. Fraser, and M. Fowler, "Transient electrochemical heat transfer modeling and experimental validation of a large

- sized LiFePO<sub>4</sub>/graphite battery," *International Journal of Heat and Mass Transfer*, vol. 109, pp. 1239–1251, 2017.
- [121] M. Muratori, M. Canova, Y. Guezennec, and G. Rizzoni, "A Reduced-Order Model for the Thermal Dynamics of Li-Ion Battery Cells," *IFAC Proceedings Volumes*, vol. 43, no. 7, pp. 192–197, 2010.
- [122] P. Gambhire, N. Ganesan, S. Basu, K. S. Hariharan, S. M. Kolake, T. Song, D. Oh, T. Yeo, and S. Doo, "A reduced order electrochemical thermal model for lithium ion cells," *Journal of Power Sources*, vol. 290, pp. 87–101, 2015.
- [123] N. Nieto, L. Díaz, J. Gastellurrutia, I. Alava, F. Blanco, J. C. Ramos, and A. Rivas, "Thermal Modeling of Large Format Lithium-Ion Cells," *Journal of the Electrochemical Society*, vol. 160, no. 2, pp. 212–217, 2012.
- [124] M. Dietze, "Entwicklung optimierter Betriebs- und Ladestrategien für Fahrzeuge mit Vehicle-2-Grid-Funktionalität," Dissertation, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, 2015.
- [125] M. Woody, M. Arbabzadeh, G. M. Lewis, G. A. Keoleian, and A. Stefanopoulou, "Strategies to limit degradation and maximize Li-ion battery service lifetime - Critical review and guidance for stakeholders," *Journal of Energy Storage*, vol. 28, p. 101231, 2020.
- [126] C. Lyu, Y. Zhao, W. Luo, and L. Wang, "Aging mechanism analysis and its impact on capacity loss of lithium ion batteries," in *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. Xi'an, China, IEEE, 2019, pp. 2148–2153.
- [127] M. Bercibar, I. Gandiaga, I. Villarreal, N. Omar, J. V. Mierlo, and P. V. den Bossche, "Critical review of state of health estimation methods of Li-ion batteries for real applications," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 572–587, 2016.
- [128] M. A. Roscher, J. Assfalg, and O. S. Bohlen, "Detection of Utilizable Capacity Deterioration in Battery Systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 98–103, 2011.
- [129] M. Coleman, W. Hurley, and C. K. Lee, "An Improved Battery Characterization Method Using a Two-Pulse Load Test," *IEEE Transactions on Energy Conversion*, vol. 23, no. 2, pp. 708–713, 2008.

- [130] J. Jiang, Z. Lin, Q. Ju, Z. Ma, C. Zheng, and Z. Wang, "Electrochemical Impedance Spectra for Lithium-ion Battery Ageing Considering the Rate of Discharge Ability," *Energy Procedia*, vol. 105, pp. 844–849, 2017.
- [131] J. Bi, T. Zhang, H. Yu, and Y. Kang, "State-of-health estimation of lithium-ion battery packs in electric vehicles based on genetic resampling particle filter," *Applied Energy*, vol. 182, pp. 558–568, 2016.
- [132] M. Gholizadeh and F. R. Salmasi, "Estimation of State of Charge, Unknown Nonlinearities, and State of Health of a Lithium-Ion Battery Based on a Comprehensive Unobservable Model," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 3, pp. 1335–1344, 2014.
- [133] Y. Zheng, J. Wang, C. Qin, L. Lu, X. Han, and M. Ouyang, "A novel capacity estimation method based on charging curve sections for lithium-ion batteries in electric vehicles," *Energy*, vol. 185, pp. 361–371, 2019.
- [134] M. Petit, E. Prada, and V. Sauvant-Moynot, "Development of an empirical aging model for Li-ion batteries and application to assess the impact of Vehicle-to-Grid strategies on battery lifetime," *Applied Energy*, vol. 172, pp. 398–407, 2016.
- [135] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, "Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods," *Journal of Power Sources*, vol. 239, pp. 680–688, 2013.
- [136] C. Hu, G. Jain, P. Zhang, C. Schmidt, P. Gomadam, and T. Gorka, "Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery," *Applied Energy*, vol. 129, pp. 49–55, 2014.
- [137] G.-W. You, S. Park, and D. Oh, "Real-time state-of-health estimation for electric vehicle batteries: A data-driven approach," *Applied Energy*, vol. 176, pp. 92–103, 2016.
- [138] S. S. Ng, Y. Xing, and K. L. Tsui, "A naive Bayes model for robust remaining useful life prediction of lithium-ion battery," *Applied Energy*, vol. 118, pp. 114–123, 2014.
- [139] C. Hu, G. Jain, C. Schmidt, C. Strief, and M. Sullivan, "Online estimation of lithium-ion battery capacity using sparse Bayesian learning," *Journal of Power Sources*, vol. 289, pp. 105–113, 2015.

- [140] V. Klass, M. Behm, and G. Lindbergh, "A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation," *Journal of Power Sources*, vol. 270, pp. 262–272, 2014.
- [141] A. Eddahech, O. Briat, N. Bertrand, J.-Y. Delétage, and J.-M. Vinassa, "Behavior and state-of-health monitoring of Li-ion batteries using impedance spectroscopy and recurrent neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 487–494, 2012.
- [142] D. D. Susilo, A. Widodo, T. Prahasto, and M. Nizam, "State of Health Estimation of Lithium-Ion Batteries Based on Combination of Gaussian Distribution Data and Least Squares Support Vector Machines Regression," *Materials Science Forum*, vol. 929, pp. 93–102, 2018.
- [143] Y. Li, S. Zhong, Q. Zhong, and K. Shi, "Lithium-Ion Battery State of Health Monitoring Based on Ensemble Learning," *IEEE Access*, vol. 7, pp. 8754–8762, 2019.
- [144] BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., "Elektromobilität Ausbau der Ladeinfrastruktur," 2021, accessed on 14.12.2021. [Online]. Available: <https://www.bdew.de>
- [145] Energie Codes und Services GmbH, "Ladesäulenregister," 2020, accessed on 21.12.2021. [Online]. Available: <https://ladesaeulenregister.de>
- [146] M. Cuchý, M. Štolba, and M. Jakob, "Whole Day Mobility Planning with Electric Vehicles," in *Proceedings of the 10th International Conference on Agents and Artificial Intelligence*. Funchal, Madeira, Portugal, SCITEPRESS - Science and Technology Publications, 2018, pp. 154–164.
- [147] Daimler AG, "Longitudinal Study e-Mobility 2018," Sindelfingen, Germany, 2018, internal document.
- [148] N. Kühl, M. Goutier, A. Ensslen, and P. Jochem, "Literature vs. Twitter: Empirical insights on customer needs in e-mobility," *Journal of Cleaner Production*, vol. 213, pp. 508–520, 2019.
- [149] M.-H. Abu-Bakar, R. Razuwan, and S. Kamal, "Teenage Driving Behavior Modeling Using Deep Learning for Driver Behavior Classification," in *Proceedings of the Third International Conference on Computing, Mathematics and Statistics (iCMS2017)*. Singapore, Springer, 2019, pp. 449–455.
- [150] W. Dong, J. Li, R. Yao, C. Li, T. Yuan, and L. Wang, "Characterizing Driving Styles with Deep Learning," IBM Research China, Shanghai, China, 2016.

- [151] X. Li, Q. Zhang, Z. Peng, A. Wang, and W. Wang, "A data-driven two-level clustering model for driving pattern analysis of electric vehicles and a case study," *Journal of Cleaner Production*, vol. 206, pp. 827–837, 2019.
- [152] M. Knowles, H. Scott, and D. Baglee, "The effect of driving style on electric vehicle performance, economy and perception," *International Journal of Electric and Hybrid Vehicles*, vol. 4, no. 3, pp. 228–247, 2012.
- [153] M. Faria, G. Duarte, and P. Baptista, "Assessing electric mobility feasibility based on naturalistic driving data," *Journal of cleaner production*, vol. 206, pp. 646–660, 2019.
- [154] E. Azadfar, V. Sreeram, and D. Harries, "The investigation of the major factors influencing plug-in electric vehicle driving patterns and charging behaviour," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 1065–1076, 2015.
- [155] Netze BW GmbH, "DIE E-MOBILITYALLEE - Das Stromnetz-Reallabor zur Erforschung des zukünftigen E-Mobility-Alltags," Stuttgart, Germany, 2019.
- [156] R. R. Desai, R. B. Chen, and W. Armington, "A Pattern Analysis of Daily Electric Vehicle Charging Profiles: Operational Efficiency and Environmental Impacts," *Journal of Advanced Transportation*, vol. 2018, pp. 1–15, 2018.
- [157] M. Schmidt, P. Staudt, and C. Weinhardt, "Evaluating the importance and impact of user behavior on public destination charging of electric vehicles," *Applied Energy*, vol. 258, p. 114061, 2020.
- [158] C. Goebel and M. Voß, "Forecasting driving behavior to enable efficient grid integration of plug-in electric vehicles," in *2012 IEEE Online Conference on Green Communications (GreenCom)*. Piscataway, NJ, USA, IEEE, 2012, pp. 74–79.
- [159] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility markov chains," in *MPM '12: Proceedings of the First Workshop on Measurement, Privacy, and Mobility*. Bern, Switzerland, Association for Computing Machinery, 2012, pp. 1–6.
- [160] X. Zhang, Z. Zhao, Y. Zheng, and J. Li, "Prediction of Taxi Destinations Using a Novel Data Embedding Method and Ensemble Learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 68–78, 2020.
- [161] P. Rathore, D. Kumar, S. Rajasegarar, M. Palaniswami, and J. C. Bezdek, "A Scalable Framework for Trajectory Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3860–3874, 2019.

- [162] V. Koolwal and K. K. Mohbey, "A comprehensive survey on trajectory-based location prediction," *Iran Journal of Computer Science*, vol. 3, pp. 65–91, 2020.
- [163] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts," in *Thirtieth AAAI Conference on Artificial Intelligence*. Phoenix, Arizona USA, AAAI, 2016, pp. 194–200.
- [164] Z. Zhang, C. Li, Z. Wu, A. Sun, D. Ye, and X. Luo, "NEXT: a neural network framework for next POI recommendation," *Frontiers of Computer Science*, vol. 14, no. 2, pp. 314–333, 2019.
- [165] O. Frendo, N. Gaertner, and H. Stuckenschmidt, "Improving Smart Charging Prioritization by Predicting Electric Vehicle Departure Time," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6646–6653, 2021.
- [166] F. Zong, Y. Tian, Y. He, J. Tang, and J. Lv, "Trip destination prediction based on multi-day GPS data," *Physica A: Statistical Mechanics and its Applications*, vol. 515, pp. 258–269, 2019.
- [167] Z. Zhao, H. N. Koutsopoulos, and J. Zhao, "Individual mobility prediction using transit smart card data," *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 19–34, 2018.
- [168] A. Rossi, G. Barlacchi, M. Bianchini, and B. Lepri, "Modelling Taxi Drivers' Behaviour for the Next Destination Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 2980–2989, 2020.
- [169] G. Broadbent, G. Metternicht, and D. Drozdowski, "An Analysis of Consumer Incentives in Support of Electric Vehicle Uptake: An Australian Case Study," *World Electric Vehicle Journal*, vol. 10, no. 1, p. 11, 2019.
- [170] J. Huber, E. Schaule, D. Jung, and C. Weinhardt, "Quo Vadis Smart Charging? A Literature Review and Expert Survey on Technical Potentials and User Acceptance of Smart Charging Systems," *World Electric Vehicle Journal*, vol. 10, no. 4, p. 85, 2019.
- [171] J. Smart, W. Powell, and S. Schey, "Extended Range Electric Vehicle Driving and Charging Behavior Observed Early in the EV Project," *SAE Technical Paper Series*, vol. 2013, pp. 1–10, 2013.
- [172] D. Andre, A. Nuhic, T. Soczka-Guth, and D. U. Sauer, "Comparative study of a structured neural network and an extended Kalman filter for state of health determination of lithium-ion batteries in hybrid electric vehicles," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 3, pp. 951–961, 2013.

- [173] Daimler AG, "Longitudinal Study e-Mobility 2015-2018," Sindelfingen, Germany, 2018, internal document.
- [174] M. Allison, E. Akakabota, and G. Pillai, "Future load profiles under scenarios of increasing renewable generation and electric transport," in *2018 5th International Conference on Renewable Energy: Generation and Applications (ICREGA)*. Al Ain, United Arab Emirates, IEEE, 2018, pp. 296–300.
- [175] J. Huber, D. Dann, and C. Weinhardt, "Probabilistic forecasts of time and energy flexibility in battery electric vehicle charging," *Applied Energy*, vol. 262, p. 114525, 2020.
- [176] E. S. Xydas, C. E. Marmaras, L. M. Cipcigan, A. S. Hassan, and N. Jenkins, "Forecasting Electric Vehicle charging demand using Support Vector Machines," in *2013 48th International Universities' Power Engineering Conference (UPEC)*. Dublin, Ireland, IEEE, 2013, pp. 1–6.
- [177] C. Bikcora, L. Verheijen, and S. Weiland, "Density forecasting of daily electricity demand with ARMA-GARCH, CAViaR, and CARE econometric models," *Sustainable Energy, Grids and Networks*, vol. 13, pp. 148–156, 2018.
- [178] M. Amini, O. Karabasoglu, M. D. Ilić, K. G. Boroojeni, and S. S. Iyengar, "ARIMA-based demand forecasting method considering probabilistic model of electric vehicles' parking lots," in *2015 IEEE Power Energy Society General Meeting*. Denver, CO, USA, IEEE, 2015, pp. 1–5.
- [179] M. B. Arias and S. Bae, "Electric vehicle charging demand forecasting model based on big data technologies," *Applied Energy*, vol. 183, pp. 327–339, 2016.
- [180] D. Panahi, S. Deilami, M. A. S. Masoum, and S. M. Islam, "Forecasting plug-in electric vehicles load profile using artificial neural networks," in *2015 Australasian Universities Power Engineering Conference (AUPEC)*. Wollongong, NSW, Australia, IEEE, 2015, pp. 1–6.
- [181] 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.
- [182] 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.

- [183] F. Ruelens, S. Vandael, W. Leterme, B. J. Claessens, M. Hommelberg, T. Holvoet, and R. Belmans, "Demand side management of electric vehicles with uncertainty on arrival and departure times," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*. Berlin, Germany, IEEE, 2012, pp. 1–8.
- [184] Daimler AG, "Evaluations of Electric Development Vehicles," Sindelfingen, Germany, 2019, internal report.
- [185] P. Staudt, M. Schmidt, J. Gärttner, and C. Weinhardt, "A decentralized approach towards resolving transmission grid congestion in Germany using vehicle-to-grid technology," *Applied Energy*, vol. 230, pp. 1435–1446, 2018.
- [186] L. Piras, E. Paja, P. Giorgini, J. Mylopoulos, R. Cuel, and D. Ponte, "Gamification solutions for software acceptance: A comparative study of Requirements Engineering and Organizational Behavior techniques," in *2017 11th International Conference on Research Challenges in Information Science (RCIS)*. Brighton, UK, IEEE, 2017, pp. 255–265.
- [187] Daimler AG, "Report Consumer Study e-Charging," Sindelfingen, Germany, 2018, internal document.
- [188] A. Schuller, B. Dietz, C. M. Flath, and C. Weinhardt, "Charging Strategies for Battery Electric Vehicles: Economic Benchmark and V2G Potential," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2014–2022, 2014.
- [189] J. García-Villalobos, I. Zamora, J. S. Martín, F. Asensio, and V. Aperribay, "Plug-in electric vehicles in electric distribution networks: A review of smart charging approaches," *Renewable and Sustainable Energy Reviews*, vol. 38, pp. 717–731, 2014.
- [190] A. S. Al-Ogaili, T. J. T. Hashim, N. A. Rahmat, A. K. Ramasamy, M. B. Marsadek, M. Faisal, and M. A. Hannan, "Review on Scheduling, Clustering, and Forecasting Strategies for Controlling Electric Vehicle Charging: Challenges and Recommendations," *IEEE Access*, vol. 7, pp. 128 353–128 371, 2019.
- [191] Q. Wang, X. Liu, J. Du, and F. Kong, "Smart Charging for Electric Vehicles: A Survey From the Algorithmic Perspective," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1500–1517, 2016.
- [192] O. Sundström and C. Binding, "Optimization methods to plan the charging of electric vehicle fleets," in *International Conference on Control, Communication,*

*Power Engineering*. Chennai, India, The Institute of Doctors Engineers and Scientists (IDES), 2010, pp. 323–328.

- [193] N. DeForest, J. S. MacDonald, and D. R. Black, “Day ahead optimization of an electric vehicle fleet providing ancillary services in the Los Angeles Air Force Base vehicle-to-grid demonstration,” *Applied Energy*, vol. 210, pp. 987–1001, 2018.
- [194] O. Frendo, J. Graf, N. Gaertner, and H. Stuckenschmidt, “Data-driven smart charging for heterogeneous electric vehicle fleets,” *Energy and AI*, vol. 1, p. 100007, 2020.
- [195] O. Frendo, N. Gaertner, and H. Stuckenschmidt, “Open Source Algorithm for Smart Charging of Electric Vehicle Fleets,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 6014–6022, 2021.
- [196] M. Schmidt, P. Staudt, and C. Weinhardt, “Decision support and strategies for the electrification of commercial fleets,” *Transportation Research Part D: Transport and Environment*, vol. 97, p. 102894, 2021.
- [197] H. Chen, Z. Hu, H. Luo, J. Qin, R. Rajagopal, and H. Zhang, “Design and planning of a multiple-charger multiple-port charging system for PEV charging station,” *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 173–183, 2017.
- [198] C. G. Cassandras and Y. Geng, “Optimal dynamic allocation and space reservation for electric vehicles at charging stations,” *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 4056–4061, 2014.
- [199] A. Gusrialdi, Z. Qu, and M. A. Simaan, “Distributed scheduling and cooperative control for charging of electric vehicles at highway service stations,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2713–2727, 2017.
- [200] Y. Kim, J. Kwak, and S. Chong, “Dynamic pricing, scheduling, and energy management for profit maximization in PHEV charging stations,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1011–1026, 2016.
- [201] L. Hou, C. Wang, and J. Yan, “Bidding for Preferred Timing: An Auction Design for Electric Vehicle Charging Station Scheduling,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3332–3343, 2019.

- [202] J. González, R. Alvaro, C. Gamallo, M. Fuentes, J. Fraile-Ardanuy, L. Knapen, and D. Janssens, "Determining Electric Vehicle Charging Point Locations Considering Drivers' Daily Activities," *Procedia Computer Science*, vol. 32, pp. 647–654, 2014.
- [203] J. Li, X. Sun, Q. Liu, W. Zheng, H. Liu, and J. A. Stankovic, "Planning Electric Vehicle Charging Stations Based on User Charging Behavior," in *2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI)*. Orlando, FL, USA, IEEE, 2018, pp. 225–236.
- [204] Y. Zhou, R. Kumar, and S. Tang, "Incentive-based distributed scheduling of electric vehicle charging under uncertainty," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 3–11, 2018.
- [205] Z. Liu, Q. Wu, S. Huang, L. Wang, M. Shahidehpour, and Y. Xue, "Optimal day-ahead charging scheduling of electric vehicles through an aggregative game model," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5173–5184, 2017.
- [206] V. Subramanian and T. K. Das, "A two-layer model for dynamic pricing of electricity and optimal charging of electric vehicles under price spikes," *Energy*, vol. 167, pp. 1266–1277, 2019.
- [207] F. Yucel, K. Akkaya, and E. Bulut, "Efficient and privacy preserving supplier matching for electric vehicle charging," *Ad Hoc Networks*, vol. 90, p. 101730, 2019.
- [208] I. T. Vadium, R. Das, Y. Wang, G. Putrus, and R. Kotter, "Electric vehicle Carbon footprint reduction via intelligent charging strategies," in *2019 8th International Conference on Modern Power Systems (MPS)*. Cluj-Napoca, Cluj, Romania, IEEE, 2019, pp. 1–6.
- [209] A. Alsabbagh, H. Yin, and C. Ma, "Distributed electric vehicles charging management with social contribution concept," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3483–3492, 2020.
- [210] B. Alinia, M. H. Hajiesmaili, and N. Crespi, "Online ev charging scheduling with on-arrival commitment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4524–4537, 2019.
- [211] G. Alface, J. C. Ferreira, and R. Pereira, "Electric Vehicle Charging Process and Parking Guidance App," *Energies*, vol. 12, no. 11, p. 2123, 2019.

- [212] N. Hoch, K. Zemmer, B. Werther, and R. Y. Siegwart, "Electric vehicle travel optimization-customer satisfaction despite resource constraints," in *2012 IEEE Intelligent Vehicles Symposium*. Madrid, Spain, IEEE, 2012, pp. 172–177.
- [213] T. Steffen, A. Fly, and W. Mitchell, "Optimal Electric Vehicle Charging Considering the Effects of a Financial Incentive on Battery Ageing," *Energies*, vol. 13, no. 18, p. 4742, 2020.
- [214] M. Ebrahimi, M. Rastegar, M. Mohammadi, A. Palomino, and M. Parvania, "Stochastic Charging Optimization of V2G-Capable PEVs: A Comprehensive Model for Battery Aging and Customer Service Quality," *IEEE Transactions on Transportation Electrification*, vol. 6, no. 3, pp. 1026–1034, 2020.
- [215] K. Zhou, L. Cheng, X. Lu, and L. Wen, "Scheduling model of electric vehicles charging considering inconvenience and dynamic electricity prices," *Applied Energy*, vol. 276, p. 115455, 2020.
- [216] L. K. Maia, L. Drünert, F. L. Mantia, and E. Zondervan, "Expanding the lifetime of Li-ion batteries through optimization of charging profiles," *Journal of Cleaner Production*, vol. 225, pp. 928–938, 2019.
- [217] S.-G. Yoon, Y.-J. Choi, J.-K. Park, and S. Bahk, "Stackelberg-game-based demand response for at-home electric vehicle charging," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 4172–4184, 2015.
- [218] J. de Hoog, T. Alpcan, M. Brazil, D. A. Thomas, and I. Mareels, "A market mechanism for electric vehicle charging under network constraints," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 827–836, 2015.
- [219] G. Zhang, S. T. Tan, and G. G. Wang, "Real-time smart charging of electric vehicles for demand charge reduction at non-residential sites," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4027–4037, 2017.
- [220] M. Alipour, B. Mohammadi-Ivatloo, M. Moradi-Dalvand, and K. Zare, "Stochastic scheduling of aggregators of plug-in electric vehicles for participation in energy and ancillary service markets," *Energy*, vol. 118, pp. 1168–1179, 2017.
- [221] F. Hertrampf, S. Späthe, and S. Apel, "Comparison of Charging Strategies of Electric Vehicles Using Local Power Production to Minimize Carbon Emissions," in *2019 IEEE International Conference on Internet of Things and Intelligence System (IoT&IS)*. Bali, Indonesia, IEEE, 2019, pp. 177–183.

- [222] A. D. Giorgio, F. Liberati, and S. Canale, "Electric vehicles charging control in a smart grid: A model predictive control approach," *Control Engineering Practice*, vol. 22, pp. 147–162, 2014.
- [223] K. Valogianni, W. Ketter, J. Collins, and D. Zhdanov, "Sustainable Electric Vehicle Charging using Adaptive Pricing," *Production and Operations Management*, vol. 29, no. 6, pp. 1550–1572, 2020.
- [224] O. Sundström and C. Binding, "Planning electric-drive vehicle charging under constrained grid conditions," in *2010 International Conference on Power System Technology*. Zhejiang, China, IEEE, 2010, pp. 1–6.
- [225] S. Li, J. Li, C. Su, and Q. Yang, "Optimization of Bi-Directional V2G Behavior With Active Battery Anti-Aging Scheduling," *IEEE Access*, vol. 8, pp. 11 186–11 196, 2020.
- [226] Q. Yang, J. Li, W. Cao, S. Li, J. Lin, D. Huo, and H. He, "An improved vehicle to the grid method with battery longevity management in a microgrid application," *Energy*, vol. 198, p. 117374, 2020.
- [227] H. Hesse, V. Kumteveli, M. Schimpe, J. Reniers, D. Howey, A. Tripathi, Y. Wang, and A. Jossen, "Ageing and Efficiency Aware Battery Dispatch for Arbitrage Markets Using Mixed Integer Linear Programming," *Energies*, vol. 12, no. 6, p. 999, 2019.
- [228] R. Das, Y. Wang, G. Putrus, R. Kotter, M. Marzband, B. Herteleer, and J. Warmerdam, "Multi-objective techno-economic-environmental optimisation of electric vehicle for energy services," *Applied Energy*, vol. 257, p. 113965, 2020.
- [229] Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-Free Real-Time EV Charging Scheduling Based on Deep Reinforcement Learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5246–5257, 2019.
- [230] B. Lunz, H. Walz, and D. U. Sauer, "Optimizing vehicle-to-grid charging strategies using genetic algorithms under the consideration of battery aging," in *2011 IEEE Vehicle Power and Propulsion Conference*. Chicago, IL, USA, IEEE, 2011, pp. 1–7.
- [231] W. Vermeer, G. R. C. Mouli, and P. Bauer, "Real-Time Building Smart Charging System Based on PV Forecast and Li-Ion Battery Degradation," *Energies*, vol. 13, no. 13, p. 3415, 2020.

- [232] R. A. Sousa, V. Monteiro, J. C. Ferreira, A. A. N. Melendez, J. L. Afonso, and J. A. Afonso, "Development of an IoT System with Smart Charging Current Control for Electric Vehicles," in *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*. Washington, DC, USA, IEEE, 2018, pp. 4662–4667.
- [233] A. Anil, S. Azeem, B. Panicker, and V. K. Saifudeen, "Electric vehicle charging assistive system using internet of electric vehicles," in *Proceedings of the International Conference on Microelectronics, Signals and Systems 2019*. Kollam, India, AIP Publishing, 2020.
- [234] F. Grée, V. Laznikova, B. Kim, G. Garcia, T. Kigezi, and B. Gao, "Cloud-Based Big Data Platform for Vehicle-to-Grid (V2G)," *World Electric Vehicle Journal*, vol. 11, no. 2, p. 30, 2020.
- [235] H. Khalloof, W. Jakob, J. Liu, E. Braun, S. Shahoud, C. Duepmeier, and V. Hagenmeyer, "A generic distributed microservices and container based framework for metaheuristic optimization," in *GECCO '18: Proceedings of the Genetic and Evolutionary Computation Conference Companion*. Kyoto, Japan, Association for Computing Machinery, 2018, pp. 1363–1370.
- [236] H. Khalloof, P. Ostheimer, W. Jakob, S. Shahoud, C. Duepmeier, and V. Hagenmeyer, "Superlinear Speedup of Parallel Population-Based Metaheuristics: A Microservices and Container Virtualization Approach," in *Intelligent Data Engineering and Automated Learning – IDEAL 2019*. Springer International Publishing, 2019, pp. 386–393.
- [237] J. Sidler, E. Braun, T. Schlachter, C. Döpmeier, and V. Hagenmeyer, "Design of a Web-Service for Formal Descriptions of Domain-Specific Data," in *IFIP Advances in Information and Communication Technology*. Springer International Publishing, 2020, pp. 201–215.
- [238] T. Kiss, P. Kacsuk, J. Kovacs, B. Rakoczi, A. Hajnal, A. Farkas, G. Gesmier, and G. Terstyanszky, "MiCADO—microservice-based cloud application-level dynamic orchestrator," *Future Generation Computer Systems*, vol. 94, pp. 937–946, 2019.
- [239] S. Shahoud, S. Gunnarsdottir, H. Khalloof, C. Duepmeier, and V. Hagenmeyer, "Facilitating and Managing Machine Learning and Data Analysis Tasks in Big Data Environments using Web and Microservice Technologies," in *MEDES '19: Proceedings of the 11th International Conference on Management of Digital*

- EcoSystems*. Limassol, Cyprus, Association for Computing Machinery, 2019, pp. 80–87.
- [240] M. Böhlend, W. Doneit, L. Gröll, R. Mikut, and M. Reischl, “Automated design process for hybrid regression modeling with a one-class SVM,” *at - Automatisierungstechnik*, vol. 67, no. 10, pp. 843–852, 2019.
- [241] R. Großmann, T. Gothan, and B. Bäker, “Adaption des Fahrerverhaltens in Abhängigkeit der Fahrsituation durch intelligente Fahrstrategien,” in *VDI Autorex 2017: Automatisiertes Fahren und vernetzte Mobilität*. Berlin, Germany, VDI Verlag, 2017, pp. 143–156.
- [242] K. Schwenk and J. Frick, “Verfahren zum inkrementellen Kondensieren von Positionsdaten elektrischer Fahrzeuge,” German Patent Application 2021ID01 931, 2021.
- [243] K. Schwenk, T. Harr, R. Großmann, R. R. Appino, V. Hagenmeyer, and R. Mikut, “Data-driven charging strategies for grid-beneficial, customer-oriented and battery-preserving electric mobility,” in *Proceedings 29th Workshop Computational Intelligence*. Dortmund, Germany, KIT Scientific Publishing, 2019, pp. 73–93.
- [244] R. Großmann, T. Gothan, S. Langhammer, and B. Bäker, “Cloudbasiertes Framework für die Entwicklung digitaler Dienste und vernetzter Funktionen,” *ATZextra*, vol. 24, no. S8, pp. 42–46, 2019.
- [245] S. Reinheimer, *Cloud Computing*. Wiesbaden: Springer Fachmedien, 2018.
- [246] R. Mikut, *Data Mining in der Medizin und Medizintechnik*, ser. Schriftenreihe des Instituts für Angewandte Informatik - Automatisierungstechnik, Universität Karlsruhe (TH). Karlsruhe: Universitätsverlag Karlsruhe, 2008, vol. 22.
- [247] D. Chen and H. Zhao, “Data Security and Privacy Protection Issues in Cloud Computing,” in *2012 International Conference on Computer Science and Electronics Engineering*. Hangzhou, China, IEEE, 2012, pp. 647–651.
- [248] D. Merkel, “Docker: lightweight linux containers for consistent development and deployment,” *Linux Journal*, vol. 2014, no. 239, p. 2, 2014.
- [249] The Linux Foundation, “Kubernetes,” 2018, accessed on 12.12.2020. [Online]. Available: <https://kubernetes.io/>

- [250] K. Schwenk, S. Meisenbacher, B. Briegel, T. Harr, V. Hagenmeyer, and R. Mikut, "Integrating Battery Aging in the Optimization for Bidirectional Charging of Electric Vehicles," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5135–5145, 2021.
- [251] C. Fiori, K. Ahn, and H. A. Rakha, "Power-based electric vehicle energy consumption model: Model development and validation," *Applied Energy*, vol. 168, pp. 257–268, 2016.
- [252] J. Hong, S. Park, and N. Chang, "Accurate remaining range estimation for Electric vehicles," in *2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*. Macao, China, IEEE, 2016, pp. 781–786.
- [253] A. Braun, *Effiziente Elektrofahrzeuge*. Wiesbaden: Springer Fachmedien, 2019.
- [254] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [255] S. Weisberg, *Applied linear regression*, 3rd ed. Hoboken, NJ, USA: Wiley-Interscience, 2005.
- [256] Python Software Foundation, "Python," 2019, accessed on 23.12.2021. [Online]. Available: <https://www.python.org/>
- [257] A. Alanis, *Artificial neural networks for engineering applications*. St. Louis, MO, USA: Elsevier, 2019.
- [258] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2014, accessed on 19.08.2019. [Online]. Available: <http://arxiv.org/abs/1412.6980v8>
- [259] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, accessed on 12.07.2021. [Online]. Available: <http://tensorflow.org/>
- [260] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

- [261] J. Brittain, "Thevenin's theorem," *IEEE Spectrum*, vol. 27, no. 3, p. 42, 1990.
- [262] N. Rotering and M. Ilic, "Optimal Charge Control of Plug-In Hybrid Electric Vehicles in Deregulated Electricity Markets," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1021–1029, 2011.
- [263] G. Plett, "High-Performance Battery-Pack Power Estimation Using a Dynamic Cell Model," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 5, pp. 1586–1593, 2004.
- [264] R. Millikan, E. Bishop, and A. T. Society, *Elements of Electricity: A Practical Discussion of the Fundamental Laws and Phenomena of Electricity and Their Practical Applications in the Business and Industrial World*. Chicago, IL, USA: American Technical Society, 1917.
- [265] W. Dahmen and A. Reusken, *Numerik für Ingenieure und Naturwissenschaftler*. Berlin/Heidelberg: Springer, 2008.
- [266] V. Gold, Ed., *The IUPAC Compendium of Chemical Terminology*. Zürich: International Union of Pure and Applied Chemistry (IUPAC), 2019.
- [267] B. Xu, A. Oudalov, A. Ulbig, G. Andersson, and D. S. Kirschen, "Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1131–1140, 2018.
- [268] M. Bercibar, "Machine-learning techniques used to accurately predict battery life," *Nature*, vol. 568, pp. 325–326, 2019.
- [269] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the Support of a High-Dimensional Distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
- [270] K. Schwenk, K. Phipps, B. Briegel, V. Hagenmeyer, and R. Mikut, "A Benchmark for Parking Duration Prediction of Electric Vehicles for Smart Charging Applications," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. Orlando, FL, USA, IEEE, 2021, pp. 1–8.
- [271] K. Schwenk, M. Faix, R. Mikut, V. Hagenmeyer, and R. R. Appino, "On Calendar-Based Scheduling for User-Friendly Charging of Plug-In Electric Vehicles," in *2019 IEEE 2nd Connected and Automated Vehicles Symposium (CAVS)*. Honolulu, HI, USA, IEEE, 2019, pp. 1–5.

- [272] E. Mynatt and J. Tullio, "Inferring calendar event attendance," in *IUI '01: Proceedings of the 6th international conference on Intelligent user interfaces*. Santa Fe, New Mexico, USA, Association for Computing Machinery, 2001, pp. 121–128.
- [273] Europäisches Parlament und Rat der europäischen Union, "Datenschutz-Grundverordnung, VERORDNUNG (EU) 2016/679," 2018, accessed on 04.07.2019. [Online]. Available: <https://dsgvo-gesetz.de/>
- [274] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, vol. 96, no. 34. Portland, Oregon, USA, AAAI, 1996, pp. 226–231.
- [275] K. Schwenk, F. Hepperle, and F. Herzog, "Verfahren zur Planung von Ladevorgängen für ein Elektrofahrzeug," International Patent Application WO 2021 259 677 A1, 2019.
- [276] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. London: Taylor & Francis Ltd., 2017.
- [277] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," *Advances in Neural Information Processing Systems*, vol. 26, pp. 431–439, 2013.
- [278] F. A. Farris, "The Gini Index and Measures of Inequality," *The American Mathematical Monthly*, vol. 117, no. 10, p. 851, 2010.
- [279] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of Predictability in Human Mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
- [280] E. Frank and M. Hall, "A simple approach to ordinal classification," in *European conference on machine learning*. Freiburg, Germany, Springer, 2001, pp. 145–156.
- [281] K. Phipps, K. Schwenk, B. Briegel, R. Mikut, and V. Hagenmeyer, "Uncertainty in Parking Duration Prediction of Electric Vehicles for Smart Charging Applications," 2022, submitted.
- [282] K. Schwenk, V. Hagenmeyer, and R. Mikut, "Multi-Day Stochastic Scheduling of Electric Vehicle Charging for Reliability and Convenience," 2022, submitted.

- [283] K. Schwenk, German Patent Request 2022ID00 158, 2022, subject to nondisclosure.
- [284] J. Dupacova, G. Consigli, and S. W. Wallace, "Scenarios for Multistage Stochastic Programs," *Annals of Operations Research*, vol. 100, pp. 25–53, 2000.
- [285] S. O. Krumke and H. Noltemeier, *Graphentheoretische Konzepte und Algorithmen*. Kaiserslautern/Würzburg: Springer Vieweg, 2012.
- [286] aWATTar Deutschland GmbH, "aWATTar HOURLY," 2021, accessed on 28.11.2021. [Online]. Available: <https://www.awattar.de/>
- [287] J. Dupacova, N. Groewe-Kuska, and W. Roemisch, "Scenario reduction in stochastic programming - An approach using probability metrics," *Mathematical Programming*, vol. 100, pp. 493–511, 2003.
- [288] D. Michie, "'Memo' Functions and Machine Learning," *Nature*, vol. 218, no. 5136, pp. 19–22, 1968.
- [289] F. Morlock, B. Rolle, M. Bauer, and O. Sawodny, "Time Optimal Routing of Electric Vehicles Under Consideration of Available Charging Infrastructure and a Detailed Consumption Model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 12, pp. 5123–5135, 2020.
- [290] E. Overbeek, "HERE.com," 2019, accessed on 04.11.2021. [Online]. Available: <https://www.here.com/>
- [291] J. Andersson, J. Gillis, G. Horn, and M. Diehl, "CasADi," 2018, accessed on 06.10.2021. [Online]. Available: <https://web.casadi.org/docs/>
- [292] P. Bonami, P. Belotti, J. J. Forrest, L. Ladanyi, C. Laird, J. Lee, F. Margot, and A. Waechter, "Basic Open-source Nonlinear Mixed INteger programming," 2019, accessed on 14.10.2021. [Online]. Available: <https://www.coin-or.org/Bonmin/>
- [293] R. Appino, J. G. Ordiano, R. Mikut, T. Faulwasser, and V. Hagenmeyer, "On the use of probabilistic forecasts in scheduling of renewable energy sources coupled to storages," *Applied Energy*, vol. 210, pp. 1207–1218, 2018.
- [294] W. H. Kwon and S. Han, *Receding Horizon Control*. London: Springer, 2005.
- [295] S. Waczowicz, M. Reischl, V. Hagenmeyer, R. Mikut, S. Klaiber, P. Bretschneider, I. Konotop, and D. Westermann, "Demand response clustering - How do dynamic prices affect household electricity consumption?" in *2015 IEEE Eindhoven PowerTech*. Eindhoven, Netherlands, IEEE, 2015, pp. 1–6.

- [296] S. Waczowicz, M. Reischl, S. Klaiber, P. Bretschneider, I. Konotop, D. Westermann, V. Hagenmeyer, and R. Mikut, "Virtual Storages as Theoretically Motivated Demand Response Models for Enhanced Smart Grid Operations," *Energy Technology*, vol. 4, no. 1, pp. 163–176, 2015.
- [297] K. Schwenk, F. Hepperle, and T. Harr, "Verfahren zur Schätzung des zukünftigen Anteils an erneuerbar erzeugter Energie im Stromnetz zum Kohlendioxid-reduzierten Laden von batterie- und hybridelektrischen Fahrzeugen," German Patent Request 2021ID01 226, 2021.
- [298] I. Konotop, S. Waczowicz, S. Klaiber, P. Bretschneider, R. Mikut, and D. Westermann, "Investigation and mathematical modelling of the impact of incentive signals to consumers on their consumption, load forecast and network operation," *IET Generation, Transmission & Distribution*, vol. 10, no. 9, pp. 2138–2146, 2016.
- [299] Fraunhofer Institute for Solar Energy Systems ISE, "Energy-Charts," 2021, accessed on 27.11.2021. [Online]. Available: <https://energy-charts.info/>
- [300] S. Meisenbacher, K. Schwenk, J. Galenzowski, S. Waczowicz, R. Mikut, and V. Hagenmeyer, "Poster: Smart Charging of Electric Vehicles with Cloud-Based Optimization and a Lightweight User Interface: A Real-World Application in the Energy Lab 2.0," in *e-Energy '21: Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. Virtual Event, Italy, Association for Computing Machinery, 2021, pp. 284–285.
- [301] —, "A Lightweight User Interface for Smart Charging of Electric Vehicles: A Real-World Application," in *2021 9th International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*. Sarawak, Malaysia, IEEE, 2021, pp. 57–61.
- [302] J. Galenzowski, "Anforderungsanalyse und Implementierung einer Industriesteuerung zur Integration eines Ladeparks mit Batteriespeicher und Photovoltaikanlage in ein Microgrid," Master's thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany, 2021.
- [303] V. Hagenmeyer, H. K. Cakmak, C. Döpmeier, T. Faulwasser, J. Isele, H. B. Keller, P. Kohlhepp, U. Kühnapfel, U. Stucky, S. Waczowicz, and R. Mikut, "Information and Communication Technology in Energy Lab 2.0: Smart Energies System Simulation and Control Center with an Open-Street-Map-based Power Flow Simulation Example," *Energy Technology*, vol. 4, pp. 145–162, 2016.

- [304] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [305] Streamlit Inc., "Streamlit," 2020, accessed on 02.12.2020. [Online]. Available: <https://www.streamlit.io/>
- [306] KEBA Group AG, "KeContact P30 c-series: Die onlinefähige Wallbox für intelligent gesteuertes Laden." 2021, accessed on 16.12.2021. [Online]. Available: <https://www.keba.com/de/emobility/products/c-series/c-serie>
- [307] Tayarani, Jahangir, Nadafianshahamabadi, Golkar, Ahmadian, and Elkamel, "Optimal Charging of Plug-In Electric Vehicle: Considering Travel Behavior Uncertainties and Battery Degradation," *Applied Sciences*, vol. 9, no. 16, p. 3420, 2019.
- [308] N. B. G. Brinkel, W. L. Schram, T. A. AlSkaif, I. Lampropoulos, and W. G. J. H. M. van Sark, "Should we reinforce the grid? Cost and emission optimization of electric vehicle charging under different transformer limits," *Applied Energy*, vol. 276, p. 115285, 2020.
- [309] A. E. Trippe, R. Arunachala, T. Massier, A. Jossen, and T. Hamacher, "Charging optimization of battery electric vehicles including cycle battery aging," in *IEEE PES Innovative Smart Grid Technologies, Europe*. Istanbul, Turkey, IEEE, 2014, pp. 1–6.
- [310] F. Ayadi, I. Colak, I. Garip, and H. I. Bulbul, "Impacts of Renewable Energy Resources in Smart Grid," in *2020 8th International Conference on Smart Grid (icSmartGrid)*. Paris, France, IEEE, 2020, pp. 183–188.
- [311] J. D. Bishop, C. J. Axon, D. Bonilla, and D. Banister, "Estimating the grid payments necessary to compensate additional costs to prospective electric vehicle owners who provide vehicle-to-grid ancillary services," *Energy*, vol. 94, pp. 715–727, 2016.
- [312] S. Bohn, M. Agsten, O. Waldhorst, A. Mitschele-Thiel, D. Westermann, and P. Bretschneider, "An ICT Architecture for Managed Charging of Electric Vehicles in Smart Grid Environments," *Journal of Engineering*, vol. 2013, pp. 1–11, 2013.
- [313] International Organization for Standardization, "ISO/FDIS 15118-20: Road vehicles — Vehicle to grid communication interface — Part 20: 2nd generation network layer and application layer requirements (under

development),” Berlin, Germany, 2022, accessed on 28.01.2022. [Online]. Available: <https://www.iso.org/standard/77845.html>

- [314] C. Spearman, “The Proof and Measurement of Association between Two Things,” *The American Journal of Psychology*, vol. 15, no. 1, p. 72, 1904.
- [315] A. Tharwat, “Classification assessment methods,” *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, 2020.
- [316] A. Lew and H. Mauch, *Dynamic Programming - A computational tool*. Berlin/Heidelberg: Springer-Verlag, 2006.
- [317] S. Meisenbacher, “Bi-directional Charging Optimization for Electric Vehicles considering Energy Prices and Battery Aging,” Master’s thesis, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, 2020.





## A Smart Charging Assistant for Electric Vehicles Considering Battery Degradation, Power Grid and User Constraints

The rise of intermittent renewable power generation increasingly impedes an efficient and reliable utility grid operation. Simultaneously, the number of electric vehicles, which require significant amounts of electric energy to charge, is growing rapidly. The energy and mobility sectors are therefore inevitably coupled, implying that reliable electric mobility depends on robust power supply. Furthermore, vehicle users perceive a limitation of their individual mobility, as electric vehicles currently provide less driving range and require more time to recharge compared with internal combustion engine vehicles. To tackle these challenges, the present thesis presents a novel concept and a software application supporting users when charging their electric vehicles, while considering the interests of all involved stakeholders.



DOI: 10.5445/IR/1000148141