Battery Life Optimal Operation of Electric Vehicles

Zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften

(Dr. Ing.)

von der Fakultät für
Wirtschaftswissenschaften
am Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Jennifer Schoch

Tag der mündlichen Prüfung: 15.05.2018
Referent: Prof. Dr. Thomas Setzer
Korreferent: Prof. Dr. Hansjörg Fromm
Korreferent: Prof. Dr. York Sure-Vetter

2018 Karlsruhe

Acknowledgements

I would like to sincerely thank my supervisor Prof. Dr. Thomas Setzer for his great support and for the opportunity to pursue my studies in this field of research. I highly appreciate the valuable discussions and advice as well as the introduction to the analytics community. Furthermore, I want to express my deep gratitude to Prof. Dr. Hansjörg Fromm for fruitful discussions, as well as Prof. Dr. York Sure-Vetter, Prof. Dr. Kai Furmans and Prof. Dr. Martin Klarmann for their valuable comments and for serving on my thesis committee.

I want to express special thanks to my research group Corporate Services and Systems for daily discussions, advice, coffee breaks and friendship. Being embedded in the team at the Institute of Information Systems and Marketing as well as FZI Research Center for Information Technology has helped and inspired me a lot, both, professionally and personally. Furthermore, I want to express my special gratitude to Dr. Nicole Gross who has introduced me to machine learning methods from a very early stage during my bachelor studies.

This work has been completed in cooperation with the BMW Group and I want to express my special gratitude to Dr. Edwin Knobbe for his support and for always believing in the value of data. Additionally, I want to specially thank Dr. Benno Schweiger for such beneficial discussions and feedback as well as the team at BMW Munich for providing their profound domain knowledge.

Finally, my special gratitude goes to my husband David and my family. David always believed in me and encouraged me, while our daughter Emily contributed with the sweetest interruptions. Without the unconditional support and trust of David as well as my parents Gerhard and Sigrid and my sister Nadine, this work would have not been possible.

Contents

List of Figures

List of Tables

List of Abbreviations

AFAP As-Fast-As-Possible Charging

ALAP As-Late-As-Possible Charging

$ALAP_b$ As-Late-As-Possible, including range buffer

BOL Begin of Life

CO2 Carbon Dioxide

DOD Depth of Discharge

ECU Electronic Control Unit

EOL End of Life

EV Electric Vehicle

GPS Global Positioning System

HEV Hybrid Electric Vehicle

I Electric Current

ICEV Internal Combustion Engine Vehicle

MOP German Mobility Panel

MSE Mean Square Error

NEDC New European Driving Cycle

NRMSD Normalized Root Mean Squared Deviation

OBD On-Board Diagnostics

OCV Open Circuit Voltage

OEM Original Equipment Manufacturer

OPT Degradation Optimal Charging

PCA Principal Component Analysis

PHEV Plug-in Hybrid Electric Vehicle

Q Charge Throughput

RMSE Root Mean Square Error

RUL Remaining Useful Life

SOC State of Charge

SOH_C State of Health of Capacity

T Temperature

V2G Vehicle to Grid

VIF Variance Inflation Factor

Part I

Foundations

Chapter 1

Introduction and Motivation

GREENHOUSE gas emissions such as CO_2, that lead to global warming, need to be reduced to preserve the earth for future generations. A key contributer of carbon emissions is individual mobility, where vehicles with an internal combustion engine–Internal Combustion Engine Vehicle (ICEV)–are the main means of transport. Recently, battery Electric Vehicle (EV)s are experiencing increasing popularity, as they offer great opportunities to reduce CO_2 emissions, when charged using renewable energies. The European Union aims to decrease the greenhouse gas emission by 20% compared to level of 1990, to cover 20% of the overall energy consumption by renewables and to increase the energy efficiency by 20% by 2020. To control climate change and global warming, these targets are even increased by 2030 and 2050 (ECF, 2010). Germany approaches these targets through the *Energiewende*. The main targets are phasing out nuclear energy and the expansion of regenerative energies of 40-45% by 2025 and 55-60% by 2035, as well as the increase of energy efficiency (BMU, 2012).

Individual transportation is still mainly dependent on fossile fuels. As a consequence, vehicles propelled by an internal combustion engine with low energy efficiency can hardly meet stricter specifications for fuel economy and emissions. The European guidelines specify CO_2 emissions for new vehicles of currently 120 g/km and 95 g/km by 2020 (UBA, 2012). Latest news from 2016 about the manipulation of emissions for diesel vehicles under test by Volkswagen and other Original Equipment Manufacturer (OEM)s, illustrate the difficulties of meeting these specifications. EVs, in contrast, are treated as zero emission vehicles.

Therefore, EVs and Hybrid Electric Vehicle (HEV)s are on the advance.

1.1 Motivation

Germany intends to have one million EVs on its streets by 2020 (BMU, 2012). Other countries strive towards similar objectives using different incentivization schemes, such as direct buyer's premium or tax reductions. Currently, a broad and fast distribution of EVs is delayed. This results by several drawbacks that EVs have compared to conventional ICEVs. Firstly, the maximum available range is typically lower. Table 1.1[1] summarizes several EVs available in Germany in 2017. Many ICEV can easily cover a range of more than 1000 km per tank filling and refilling within minutes is possible nearly everywhere. As depicted in Table 1.1, even the EV with the highest range–Tesla's Model S–cannot cover 1000 km per battery charge. Other, less costly EVs such as the latest versions of the BMW i3, Renault Zoe, Nissan Leaf and VW E-Golf provide moderate ranges between 250 and 400 km. The ranges given in Table 1.1, however, are based on the New European Driving Cycle (NEDC) and are valid under ideal conditions. Ambient temperature, driving style and speed, topology of the terrain, air-condition and heating as well as other auxiliaries considerably reduce the ideal range.

Another issue comes with recharging. The duration depends on the available power as well as the capacity of the battery and full recharging typically takes several hours. Furthermore, charging infrastructure outside of urban centers is still scarce.

The currently used energy storage system in an EV is typically a lithium-ion battery. Compared to other battery technologies, lithium-ion systems provide a high energy density. Compared to petroleum fuel, however, the energy density is low. Whereas the energy density influences both weight and volume of the battery, the high price is another issue. In summary, the energy storage system in an EV–the high voltage battery–causes limited range, a high price and high

[1] www.renault.de, www.smart.de, www.bmw.de, www.tesalmotors.de, www.nissan.de, www.vw.de

Model	Price [Euro]	Range [km] (NEDZ)	Battery Capacity [kWh]
Renault Zoe	22,100	400	41,0
Smart Fortwo Electric Drive	21,940	160	17,6
BMW i3	36,800	300	33,0
Tesla Model S	112,770	632	100,0
Nissan Leaf	34,385	250	30,0
VW E-Up	26,900	160	18,7
VW E-Golf	35,900	300	35,8

Table 1.1: Overview of currently available EVs in Germany

weight compared to a conventional ICEV.

Besides the mentioned issues, the battery has a limited lifetime and battery degradation is an extensive field of research. The EOL for a battery in an automotive application is typically specified at 80% of the initially available capacity. Although operation is further possible (Saxena et al., 2015), some batteries then show unstable behavior discouraging a further usage of the battery in EVs (Jossen and Weydanz, 2006; Spotnitz, 2003). The battery degrades with cycling but also with time (calendar aging). However, user behavior and in particular charging decisions strongly influence the degradation process.

Initial studies about the behavior of current EV users have revealed that the uncertainty about the possibility to recharge and the range limitations lead to a phenomenon called *range anxiety*. Franke and Krems (2013a) study EV drivers and their charging behavior and found that users prefer to retain range buffers. Since the observed average daily distance is far below the available range, daily recharging is found unnecessary from a technical perspective and consequently the battery State of Charge (SOC) is held in a high range. For battery degradation, and in particular the calendaric aging component, a high SOC is typically disadvantageous and reduces battery life.

The limited life of the battery influences the environmental efficiency of an EV, impacts the maximum range available for the user, and also raises the need for suitable guarantee design for the OEM. Therefore, a smart charging strategy indicating when and how much to charge, needs to be found in order to maximize battery life.

To accomplish this, it is necessary to identify relevant variables and suitable sampling from the vast amount of sensor signals that are available and related to battery degradation. Based on this, a detailed degradation model needs to be developed that maps usage related and environmental conditions to battery degradation.

1.2 Research Questions

The progress of battery degradation highly depends on usage conditions. Therefore, the batteries of EVs of equal age and charge throughput may have experienced degradation differently, and the remaining capacity–and thus remaining range–can vary considerably.

It is known from laboratory aging tests that the major drivers for battery aging are operating conditions related to time, energy throughput, SOC and temperature (Jossen, 2006). The consideration of these factors and their impact on battery aging is essential to propose adjustments of utilization patterns in daily operation that prolong the operational lifetime and maximize the cumulated range of an EV. Thus far, there is only little guidance on how to appropriately consider battery degradation in charging recommendations. First, no models have yet been proposed to estimate the sensitivity of battery lifetime in empirical settings on driving and charging behavior. Second, the phenomenon of range anxiety of EV users–that leads to frequent and full recharging–as well as uncertainty in range predictions need to be considered in any charging strategy recommendation. In particular the trade-off between battery life optimal charging behavior–discouraging to fully charge a battery–and flexibility in mobility requirements, i.e., range–encouraging high charging levels–must be investigated and addressed appropriately.

These aspects are formulated in RQ1.

RQ 1 **Battery Life Optimal Charging**

- **a)** To what extent can battery life be extended by the application of a degradation optimal charging strategy?
- **b)** How to model the trade-off between range flexibility in terms of buffers and battery life in a formal fashion?

I contribute to the literature by introducing a continuous quadratic programming model to calculate the battery life optimal charging strategy OPT. OPT aims at maximizing the time until the end of life (EoL) of a battery, given an empirical driving profile and a degradation model.

However, optimal charging requires deterministic knowledge of future trips and corresponding charging levels–an assumption that is not fully accessible in real-life settings. Therefore, I investigate the more convenient charging heuristic As-Late-As-Possible Charging ($ALAP_b$). In this approach, range buffers are considered, that preserve range for unexpected trips or inaccurate range predictions by answering RQ1b. In simulation studies with real-life assumptions on vehicle parameters and mobility requirements (based on a representative set of empirical driving profiles from the German mobility panel (BMVBS, 2008)), I analyze the trade-off between range flexibility (required by a user) and battery life (compared to OPT) achievable with $ALAP_b$. Based on the results, battery degradation aware charging heuristics can easily derived and applied in real-world settings.

The prerequisite to answer RQ1 is a detailed battery degradation model. The increasing number of EVs in the field offer a vast amount of on-board data gathered through telematics and periodic inspections. Therefore, data related to battery degradation and corresponding usage conditions need to be exploited appropriately in order to derive a battery degradation model that is valid under real-world conditions. Whereas on-board data storage as well as transmission bandwidth are limited at present, it is not sufficiently clear which variables and signals to store and to transmit. Thus, data selection and reduction to be approached firstly.

I close this gap of current literature by providing a decision support for OEMs on how to collect sensor data for accurate prediction of system states in terms of

battery degradation. This aim is summarized in RQ2.

RQ 2 **Data Representation**
What sampling and variable subset of sensor data is recommended, to accurately predict the system states of battery degradation?

By means of a simulation based analysis, dynamic user behavior is simulated based on real-world driving profiles which are parameterized by different driver characteristics and ambient conditions. Using analytical models, a reduced set of features is derived that allows for an accurate prediction of battery degradation in EVs based on standard equipment. This allows for efficient data acquisition in a fleet of EVs in order to derive a degradation model.

The answer from the simulation based analyses of RQ2 are transferred to evaluate real-world data and to develop a predictive model that accurately maps degradation relevant variables to battery aging.

A detailed assessment of battery degradation is crucial to derive a degradation optimal charging strategy, but also the evaluation of the economic benefit of any smart charging strategy, e.g. for Vehicle to Grid (V2G) approaches. By now, in smart charging strategies battery degradation is typically–if at all–considered by simplified or linearized terms. The degradation process, however, has been identified to be a lot more complex.

Recent literature on degradation model development is focused on empirical data fitting based on laboratory tests. However, due to the required time and effort of accelerated aging tests, the tested parameter combinations are usually very limited. Typically, real environmental and usage conditions especially in terms of temperature, differ considerably from test conditions. However, it is not recommended to extrapolate, because it may lead to invalid results.

In order to solve this gap of research I develop an empirical, field data-based battery degradation model by using methods of data analytics. I overcome the issues of non-realistic conditions, especially of temperature that can be observed in laboratory settings, by using field data, gathered from every-day EV usage.

These objectives are formulated in RQ3.

RQ 3 **Degradation Prediction Modeling**

- **a)** What is the accuracy of an empirical, field-data based degradation model?
- **b)** What variable subset of field-data is recommended, to accurately predict battery degradation?

The analysis based on field data of more than 5000 EVs, has the advantage of nearly arbitrary parameter combinations. Therefore, degradation relevant variables can be considered in any possible functional relationship and interaction. In order to analyze the predictive accuracy of the developed degradation models, the test error is quantified using cross-validation based on unseen test data to detect and avoid overfitting. The ambiguously reported functional relationship of degradation and time as well as charge throughput are analyzed systematically using Box-Cox-Transformation. Degradation relevant variables, that improve the predictive accuracy of the models, are systematically selected using Lasso regression.

Due to the coverage of real environmental and usage conditions, I develop an empirical, field data-based degradation model that leads to the most promising base for a valid evaluation of smart charging strategies, through which I am able to fill this gap of research.

1.3 Structure of the Thesis

For the purpose of a thorough analysis and evaluation of a degradation optimal smart charging strategy, this thesis is structured as depicted in Figure 1.1. The first part introduces the foundations of this thesis. Following the introduction and motivation, characteristics of EVs and the basic concepts of lithium-Ion batteries are discussed. An introduction to battery degradation including an overview of battery degradation models, currently available in literature, is followed by an outline of current findings on EV user behavior.

The prerequisites to answer RQ1 are the answers to RQ2 and RQ3. Therefore, before degradation optimal charging is investigated in Part III of this thesis, data

representation and prediction of battery degradation are analyzed in Part II.

In Part II, dynamic user behavior is simulated based on real-world driving profiles parameterized by different driver characteristics and ambient conditions by means of a simulation-based approach. This includes considerations on data storage and transmission bandwidth that are limited technically and economically. Following that, a predictive model, i.e. an empirical battery degradation model is developed based on a fleet of more than 5000 EVs operated without controlled environmental conditions. The predictive power of different degradation models is compared to literature-inspired models by means of cross-validation. Furthermore, relevant variables and interaction terms are identified using the Lasso method.

The findings of Part II allow to develop the degradation optimal charging strategy and answer RQ1. Based on a comprehensive battery aging model, I introduce a continuous quadratic programming model to derive battery life optimal charging. The strategy indicates when and how much to charge to maximize the potential range throughout the battery life. Since optimal charging would require deterministic knowledge of future trips and corresponding charging levels I investigate a more convenient charging heuristic, which implies different range buffers and investigate the trade-off between battery degradation and range flexibility.

The fourth and final part of this thesis includes conclusions as well as an outlook on future work and possible extensions.

This work is based partially on published and working papers. The selection and reduction of data in a simulation-based analysis is partially based on Schoch et al. (2017). Furthermore, the predictive modeling of battery degradation is partially based on Schoch et al. (2018) and the prescriptive analytics of a degradation optimal charging strategy is partially based on Schoch et al. (2018).

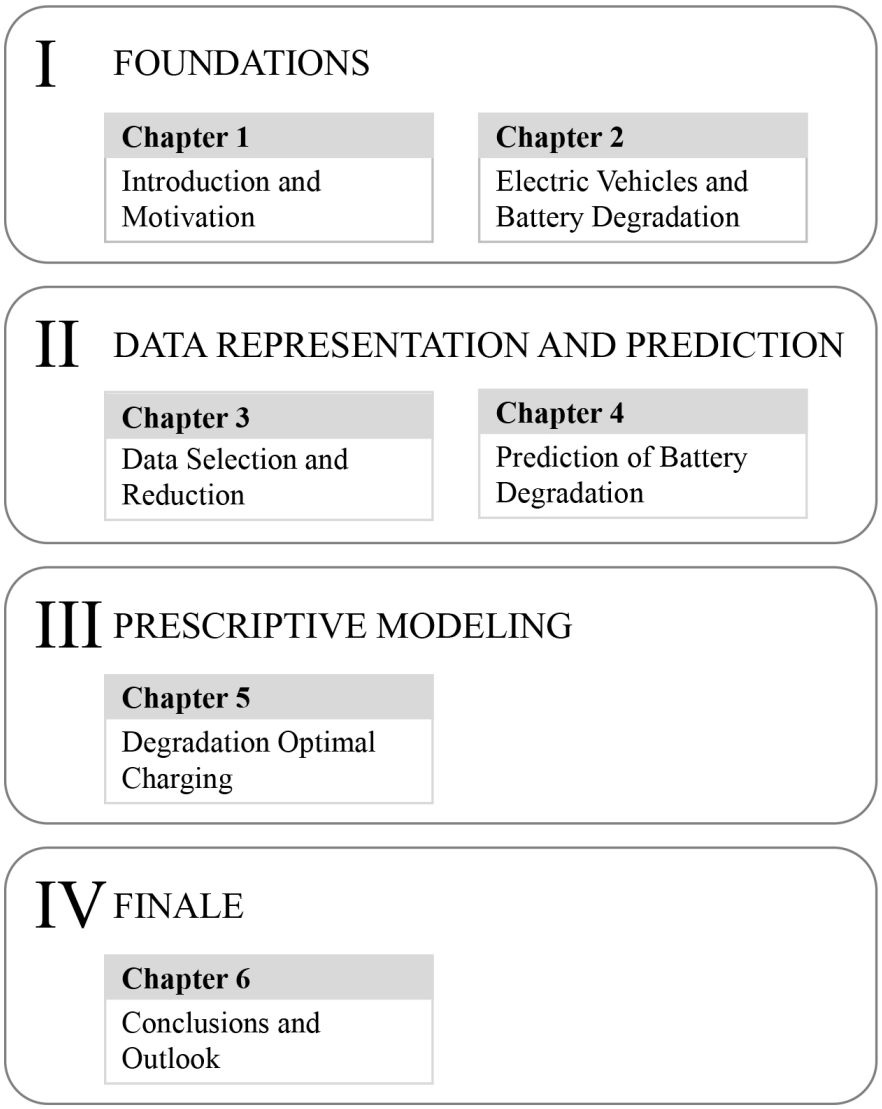

Figure 1.1: Thesis Structure

Chapter 2

Electric Vehicles and Battery Degradation

ELECTRIC vehicles have first been introduced in the early 20th century and were even preferred over vehicles with an internal combustion engine. ICEVs used to be unreliable, smelly and needed to be manually cranked to start. However, as soon as cheap oil was available and the self starter was invented after 1910, ICEVs became the preferred means of transport up until today (Larminie and Lowry, 2004).

2.1 Characteristics of Electric Vehicles

This development was promoted by the drawbacks of EVs that result from the energy storage system–the high voltage battery. The battery limits the EV's range, due to it's low specific energy. Petroleum fuel has a specific energy of approximately 9000 Wh/kg, of which 1800 Wh/kg are usable due to the rather low energy efficiency of the combustion engine of η in the range of $20 - 50\%$ (Larminie and Lowry, 2004; Soimar and Kluger, 2000). A lithium-ion battery, which corresponds to the currently preferred battery technology, in contrast, has a specific energy of around 200 Wh/kg (Linden and Reddy, 2011). The electric motor is much more efficient with $\eta \approx 90\%$ as compared to the combustion engine and, by using regenerative breaking, energy can be reused. Nevertheless, a battery storage system that provides a range comparable to that of an ICEV is

very heavy. Other drawbacks of an EVs include long recharging times and high costs (Larminie and Lowry, 2004).

Recently, EVs have gained new popularity, due to increasing oil prices and environmental issues such as CO2 and NOx emissions. Germany, for instance, intends to have one million EVs on it's streets by 2020 (NPE, 2011). Advancements, especially in the lithium-ion battery technology enable large scale production and distribution of EVs. To meet the emission standards and to solve environmental issues as well as the growing energy insecurity, the transition between ICEVs and EVs is promoted by HEVs that combine both technologies (Liu et al., 2013).

Whereas EVs are solely powered by an electric motor that draws it's energy from the battery, alternative concepts include HEVs and Plug-in Hybrid Electric Vehicle (PHEV)s. A HEV is composed of a combustion engine as well as an electric motor. The battery is charged during driving phases or by regenerative breaking and the electric motor is typically applied to power the vehicle in certain speed ranges. The PHEV, in contrast, can be recharged by plugging into the electric grid. PHEVs are operated in charge depletion or charge sustaining mode. Charge depletion mode corresponds to an all-electric operation, while power demands are met. Once the lower bound of state of charge (SOC) is reached, the energy from the combustion engine is used to propel the vehicle and to keep the SOC constant in charge sustaining mode (Linden and Reddy, 2011).

To date EVs are not price-competitive with ICEVs, due to the high cost of the lithium-ion battery of approximately 300 $ per kWh (Nykvist and Nilsson, 2015). Following the US department of energy which aims at lowering the battery cost to 125 $ per kWh by 2022, EVs are expected to be cost-competitive at 150 $ per kWh (Department of Energy, 2017).

The lithium-ion battery is the most expensive part of an EV, which moreover degrades with both time and charge throughput. It is crucial to reduce the battery degradation to a minimum. This work therefore focuses on the minimization of battery degradation related to EV usage conditions. For this purpose the following paragraphs provide an overview of the lithium-ion battery system,

degradation drivers, as well as currently observed user behavior.

2.2 Fundamentals of Batteries

A battery is composed of one or more cells and converts chemical energy into electrical energy by means of an electrochemical oxidation-reduction (redox) reaction. For primary cells, this process is irreversible, while secondary cells can be recharged. A cell is composed of the negative electrode (anode), the positive electrode (cathode), the ionic conductor (electrolyte) and the separator. The active material, necessary for the discharge reaction, is contained in the electrodes.

The negative electrode is oxidized by giving up electrons to the external circuit at discharge. The same amount of electrons is accepted by means of reduction of positive active material at the positive electrode (electrons are accepted). Electrochemical reactions take place at the interface between electrode and electrolyte. Therefore, the active material is typically porous to provide a large surface to be covered by electrolyte.

Assuming a metal as the anode material and chlorine (Cl_2) as cathode material, the discharge reaction is expressed by Equations (2.1) and 2.2 for the negative and positive electrode, respectively.

$$Zn \rightarrow Zn^{2+} + 2e \tag{2.1}$$

$$Cl_2 + 2e \rightarrow 2Cl^- \tag{2.2}$$

Whereas the negative electrode (Zn) is oxidized by giving up electrons ($Zn^{2+} + 2e$), the positive electrode is reduced by gaining electrons ($2Cl^-$). As depicted in Figure 2.1 (left), electrons pass an external load. In the electrolyte–typically a liquid–an ionic current closes the electric circuit. However, the electrolyte has limited conductivity and exhibits an internal resistance.

Applying a voltage source instead of a load to the external circuit, the direction of the current turns and all processes within the cell are reversed and the cell is recharged. This process is depicted in Figure 2.1 right (Jossen and Weydanz, 2006; Linden and Reddy, 2011).

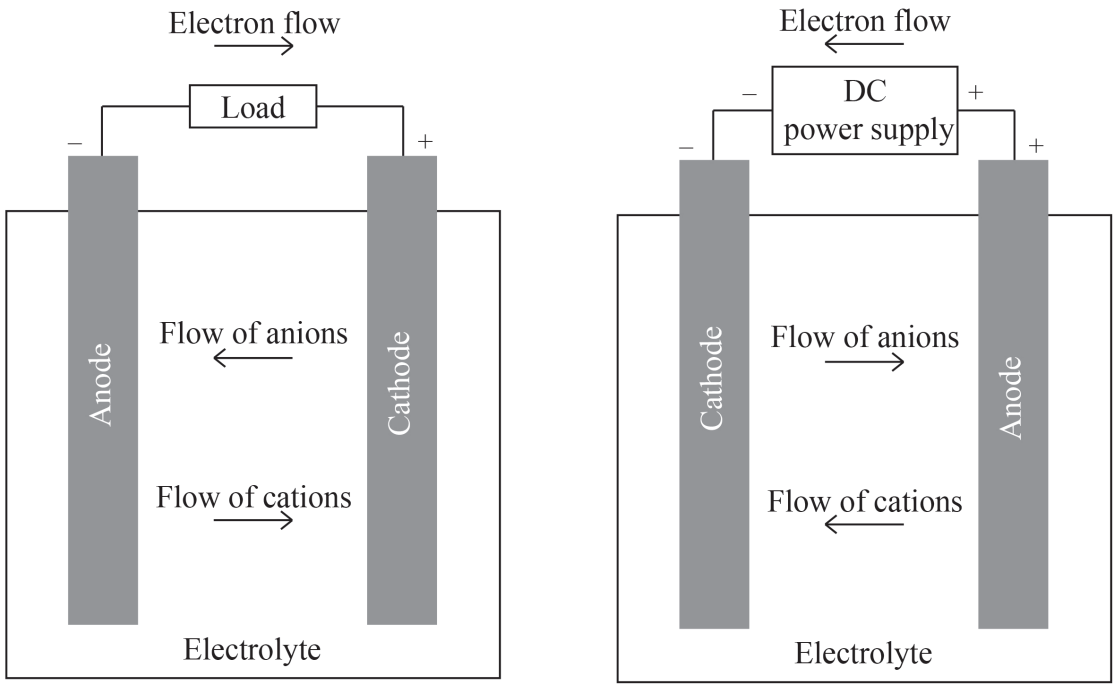

Figure 2.1: Electrochemical operation of a cell in discharge mode (left) and charge mode (right) (adapted from Linden and Reddy (2011))

The theoretical capacity of the cell results from the proportionality of active material and electric load (one mol of Zn yields two mols of electrons). In practical applications, however, the theoretical capacity is not reached, since not all active material can be discharged and further electrochemical processes take place in parallel (e.g. self discharge or degradation) (Jossen and Weydanz, 2006; Linden and Reddy, 2011).

An important characteristic of a battery is it's specific energy–when referred to the mass–or the energy density–when referred to the volume. Lithium-ion batteries compared to other battery technologies, exhibit both a high specific energy as well as a high energy density. However, the energy density is reduced with increasing energy density. Consequentially, high energy cells are suitable for EV applications, where the discharge process is distributed over a long period of time and high mileage ranges are desired. High power cells, on the other hand, are well suited for hybrid applications, where peak charges and discharges for acceleration or regenerative breaking are necessary (Jossen and Weydanz, 2006;

Linden and Reddy, 2011).

2.2.1 Lithium-ion Batteries

To date, lithium-ion batteries are the preferred technology for EV storage systems. High suitability, as compared to other battery technologies, result for example from the long cycle life, long shelf life, low self-discharge rate, high-rate as well as high power discharge capability and no memory effect. The first lithium-ion battery with lithium cobald oxide ($LiCoO_2$) as the positive electrode has been introduced by Sony in 1991. The lithium-ion battery is also known as "swing" or "rocking chair" battery, due to the fact that lithium ions (Li^+) are exchanged between the positive and negative electrode. The positive electrode material is typically a metal oxide and the negative electrode a graphic carbon. Apart from lithium cobald oxide LCO ($LiCoO_2$) other positive electrode materials include LFP ($LiFePO_4$), LMO (Spinel) ($LiMn_2O_4$), NMC ($Li(NiMnCo)O_2$), NCA ($Li(NiCoAl)O_2$) and LTO ($Li_{4/3}Ti_{5/4}O_4$ spinel) (Jossen and Weydanz, 2006; Linden and Reddy, 2011).

Functionality of Lithium-Ion Batteries

In the cycling process, lithium-ions (Li^+) are transported through the liquid electrolyte. At charging, lithium-ions are deintercalated from the positive electrode and pass through the electrolyte. By reacting with electrons from the external circuit, resulting lithium atoms are intercalated into the host, in between the graphite layers. The process is reversed at discharging, where lithium atoms release an electron and leave the host (charge transfer). The resulting lithium-ions are transported from the negative to the positive electrode through the electrolyte. The lithium-ions are intercalated into the host at the positive electrode. The process is illustrated in Figure 2.2.

The discharge reaction for an exemplary LMO postitive electrode material can be expressed by Equations 2.3 and 2.4 for the positive and negative electrode,

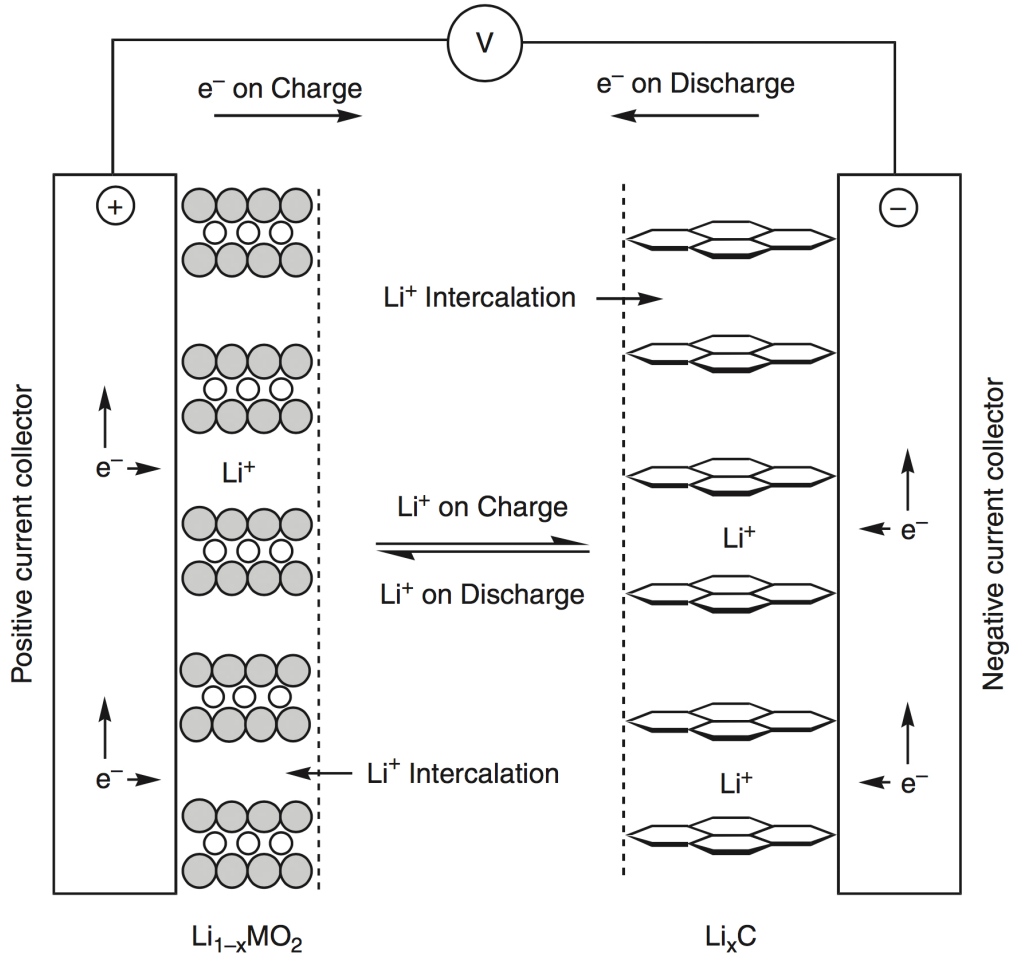

Figure 2.2: Functionality of a LMO lithium-ion battery (Linden and Reddy, 2011))

respectively (Jossen and Weydanz, 2006; Linden and Reddy, 2011).

$$Li_{1-x}MO_2 + xLi^+ + xe^- \rightarrow LiMO_2 \tag{2.3}$$

$$Li_yC \rightarrow C + yLi^+ + ye^- \tag{2.4}$$

The lithium-ions represent the active species in a cell, however typically not the active material itself. Therefore, a wide variety of active materials that accept lithium-ions can be selected for the positive and negative electrode. Materials with a potential close to that of lithium metal is well suited for the negative electrode (e.g. graphite). The material for the positive electrode, on the other hand, should have a high potential compared to the lithium metal in order to allow for

a high cell capacity. The cell voltage depends on the chosen electrode materials and their potentials and typically lies between 3.0 and 4.2 V. However, due to the limited amount of lithium within the cell as well as different concentrations of cathode and anode potential, the electrode capacity cannot be used entirely (Jossen and Weydanz, 2006; Linden and Reddy, 2011).

Lithium-ion batteries can be constructed in round cells, where the most frequently used type is 18650, corresponding to a diameter of 18 mm and a length of 65,0 mm. Anode- and cathode material as well as a separator filled with electrolyte are rolled up and packed in a housing. Other geometries include prismatic (flat rolled) and pouch cells (stacked electrodes).

Characteristics of Lithium-Ion Batteries

Lithium-ion cells are characterized by the Open Circuit Voltage (OCV) that is derived by experiment. It maps the lithium concentration or alternatively the SOC in percent to the cell voltage under rest conditions. The OCV-curve is typically S-shaped. The cell is fully charged for instance at 4.2 V and fully discharged to the terminal voltage between 2.7 and 3.0 V, depending on the application. After the terminal voltage is nearly reached, further discharging of the cell leads to a rapid voltage drop that should be avoided. The internal ohmic resistance results from charge transfer at the interface and from losses due to ionic transport in the electrolyte and conductivity in the porous electrodes. As a result, increasing discharge currents decrease the cell voltage compared to rest conditions. Therefore, the terminal voltage is reached earlier and the extractable capacity of the cell is reduced. Something similar applies for charging. Increased charging currents increase the cell voltage, such that the upper voltage bound is reached before the cell is fully charged. To be able to access the total available capacity of a cell, charging protocols such as CC-CV (constant-current constant-voltage) are applied. After charging with a constant (high) current, the voltage is held constant at the upper voltage bound and the current is decreased until it drops below a predefined value. Thus, applying CC-CV charging, the cell can be nearly fully charged even with an initially high current.

Whereas temperature influences the conductivity of the electrolyte, the ex-

tractable capacity of the cell decreases typically for temperatures below $0°C$. High temperatures above $30°C$ on the other hand increase degradation (Jossen and Weydanz, 2006).

Degradation is a considerable issue for lithium-ion batteries. Especially for traction batteries that require both high energy density as well as high specific energy, which makes them expensive. The following paragraph therefore describes electro-chemical processes that cause degradation and as well as degradation models.

2.3 Battery Degradation

In lithium-ion batteries, reversible as well as irreversible loss of capacity arises. Reversible capacity loss is caused for example by self discharge of the cell, typically in the range of few percents per month. The full capacity of the cell can be reached by recharging the cell. Irreversible capacity loss, on the other hand, damages the cell and should be minimized. Compared to other cell chemistries lithium-ion batteries have limited calendaric life independent of usage. Additionally, capacity loss is driven by cycling. Capacity loss is furthermore indirectly driven by the increasing internal resistance, which corresponds to power loss.

An increasing internal resistance, which is driven by cycling but also over time, is caused by the formation of a passivation layer at the electrode-electrolyte interface at the anode–the solid electrolyte interface (SEI). The layer is initially formed in the first cycles–typically by the cell manufacturer–and protects the electrolyte from further decomposition by reacting with the graphite. With further cycles and time, however, the SEI layer grows and results in an increased internal resistance. The process is driven by high temperatures and high currents as well as high cell voltages.

Charging at high currents and low temperatures can cause lithium-plating, where lithium metal is deposited at the anode surface. This causes a loss of active lithium within the cell and may even lead to a short circuit.

Violation of upper and lower voltage bounds (overcharge or deep discharges)

and surpassing the upper temperature bound, furthermore, increase degradation. These constraints given by the cell manufacturer, however, are typically adhered to by the battery management system. Moreover, charging and discharging causes mechanical stress due to volumetric changes within the cell and leads to capacity loss.

Capacity loss over the number of cycles can be divided into three phases. The first phase is characterized by high capacity loss due to the formation of the SEI layer, as described above, which is done by the manufacturer. The capacity loss stabilizes in phase two until the EOL definition (60-80% of remaining capacity) (for automotive applications) is reached. Subsequently, the capacity loss dramatically increases in phase three (Jossen and Weydanz, 2006; Spotnitz, 2003). In the following part of this thesis I will concentrate on the degradation process of phase two, which is relevant for the application in EVs.

2.3.1 Usage Related Degradation Drivers

In literature the reasons for battery degradation such as SEI layer growth, loss of active lithium and lithium plating are analyzed electro-chemically by equivalent circuit models. In this thesis, however, I focus on the relationship between usage conditions, such as SOC, temperature and battery degradation. Therefore, I will now discuss what is known about the complex processes behind the two components that cause capacity fade, namely the *calendaric aging* under storage and *cyclic aging* under usage (Jossen and Weydanz, 2006). I will focus on lithium-ion batteries, the predominantly used chemistry in currently available EVs, and assume that manufacturer specifications are met by the battery management system.

Calendaric aging has been found to depend on the SOC and temperature (T). The SOC is measured in percent of the actual capacity, which depends on the cell voltage (measured in volts). This relationship is determined by the OCV curve, which is typically nonlinear. High SOC as well as high temperature increase the decomposition of cell composites and increase degradation. The relationship between aging and temperature is described by the Arrhenius law. The

Arrhenius law describes the speed of a chemical reaction, and is typically used to describe the thermal behavior of batteries. It indicates the rapid increase of reaction speed with increasing temperature (Spotnitz, 2003; Kaebitz et al., 2013; Schmalstieg et al., 2014). The relationship is described by an exponential function.

$$C_{cal} \propto (\sim exp(-\frac{E_a}{RT})) \tag{2.5}$$

Thereby, R is the universal gas constant, E_a is the activation energy necessary for the reaction and T is the temperature.

Furthermore, calendaric aging is described by a monotonically declining function with time (Jossen and Weydanz, 2006). Thus, at the begin of life the degradation rate is higher and linearizes with increasing time. However, the resulting capacity reduction is increased with higher SOC and T. The Arrhenius law indicates a halving of lifetime with an increase of $10°C$ in temperature (Marongiu et al., 2015; Sarasketa-Zabala et al., 2016; Ecker et al., 2014; Schmalstieg et al., 2014; Kaebitz et al., 2013). Calendaric aging is inevitable as it is related to time, however higher temperatures increase calendaric aging rapidly.

$$C_{cal} \propto f\left(exp\left(-\frac{E_a}{RT}\right), f(SOC), t^{k_1}\right) \tag{2.6}$$

With $0 < k_1 \leq 1$.

Whereas literature typically agrees upon the Arrhenius-relationship (defining the relationship with temperature), $f(SOC)$ is either described as a linear function (Schmalstieg et al., 2014), or in terms of an exponential function of SOC (Ecker et al., 2012, 2014; Sarasketa-Zabala et al., 2016; Marongiu et al., 2015). For the relationship with time however, k_1 is identified as 0.5 corresponding to a square root relationship (Marongiu et al., 2015; Sarasketa-Zabala et al., 2016; Ecker et al., 2012, 2014; Kaebitz et al., 2013; Wang et al., 2014), or $k_1 = 0.75$ (Schmalstieg et al., 2014).

Cyclic aging highly depends on the chemical composites of the battery. However, a large Depth of Discharge (DOD)–the SOC range in which cycling occurs–has been found to increase degradation, while a low DOD around a moderate

state of charge ($\overline{SOC}$ around 50%) was proven to enable a longer battery life. $\overline{SOC}$ is the average SOC within a cycle. As an example, it would be beneficial to cycle a battery around 40 - 60% SOC, with a small DOD of 20% instead of taking advantage of the full range between 0 to 100%. However, $\overline{SOC}$ is at the optimum for reduced degradation of 50% for both examples. Additionally, high charge and discharge currents (I) increase degradation (Jossen and Weydanz, 2006; Bashash et al., 2011; Linden and Reddy, 2011). Throughout the first cycles the gradient of the SOH_C curve is rather high, while the rate of degradation stabilizes with charge throughput (Jossen and Weydanz, 2006). This relationship has frequently been described by a square root function of charge throughput Q ($\sim \sqrt{Q}$) (Kaebitz et al., 2013; Barré et al., 2013; Ramadass et al., 2003) or as a linear function of Q (Wang et al., 2014).

$$C_{cyc} \propto f\left(f(\overline{SOC}), f(DOD), f(I), Q^{k_2}\right) \tag{2.7}$$

With $0 < k_2 \leq 1$.

The most detailed degradation models, that can be found in literature to date, are summarized according to cell type (geometry), chemistry, capacity and considered parameters in the model in Table 2.1. While $\overline{SOC}$ is only provided in the degradation model developed by Schmalstieg et al. (2014), it is considered by a quadratic relationship. DOD is included in the work of Schmalstieg et al. (2014) linearly and by Marongiu et al. (2015) as well as Sarasketa-Zabala et al. (2016) exponentially. Marongiu et al. (2015); Sarasketa-Zabala et al. (2016); Ecker et al. (2014); Wang et al. (2014) perform accelerated aging tests with a varying C-rate (a C-rate of 1 C would correspond to the current required to fully charge the battery within one hour, e.g. the rate of 1 C for a 2 Ah battery equals 2 A). However, only Wang et al. (2014) have found a functional relationship that unambiguously explains the declining effect of C-rate on capacity. The dependency on charge throughput Q, however, is modeled explicitly with an exponent of $k_2 = 0.5$ (Schmalstieg et al., 2014) and $k_2 = 0.87$ (Sarasketa-Zabala et al., 2016). Other degradation models include Q implicitly, e.g. by the current I multiplied with time (Wang et al., 2014) or the number of equivalent full cycles (Marongiu et al., 2015).

	Cell Characteristics		Parameters				
Source	**Type & Chemistry**	**Capacity**	**SOC**	**T**	$\overline{\textbf{SOC}}$	**DOD**	**C − Rate**
Marongiu et al. (2015)	NMC pouch cell	40 Ah	✓	✓	x	✓	x
Marongiu et al. (2015)	LFP cylindrical	8 Ah	✓	✓	x	✓	x
Sarasketa-Zabala et al. (2016)	26650 LFP	2.3 Ah	✓	✓	x	✓	x
Ecker et al. (2012)	NMC pouch cell	6 Ah	✓	✓	x	x	x
Ecker et al. (2014)	18650 NMC	2.15 Ah	✓	✓	x	x	x
Schmalstieg et al. (2014)	18650 NMC	2.15 Ah	✓	✓	✓	✓	x
Kaebitz et al. (2013)	NMC pouch cell	10 Ah	x	✓	x	x	x
Wang et al. (2014)	NMC/LMO 18650	1.5 Ah	x	✓	x	x	✓

Table 2.1: Overview of degradation models. ✓: Parameter is considered in the degradation model, x: Parameter is constant

The authors Marongiu et al. (2015); Sarasketa-Zabala et al. (2016); Wang et al. (2014) and Schmalstieg et al. (2014) consider the calendaric and cyclic part of aging to be additive, as indicated by Equation 2.8, while Ecker et al. (2014) and Kaebitz et al. (2013) do not consider the calendaric degradation term at all.

$$C = 1 - C_{cal} - C_{cyc} \tag{2.8}$$

All degradation models presented in Table 2.1 find that increasing temperature and SOC lead to increased degradation. However Wang et al. (2014) find indifferent behavior between degradation and SOC in combination with various C-rates, while Kaebitz et al. (2013) do not include SOC in the degradation model.

Marongiu et al. (2015); Sarasketa-Zabala et al. (2016) as well as Schmalstieg et al. (2014) identify elevated degradation associated with an increasing DOD. Only the work presented in (Schmalstieg et al., 2014) includes $\overline{SOC}$ and find a quadratic relationship with degradation, exhibiting a minimum around a $\overline{SOC}$ of 50%. The only model in Table 2.1 that includes C-rate introduced by (Wang et al., 2014), does test SOC, DOD and $\overline{SOC}$, but does not include these parameters in the final model, due to the ambiguous relationship with degradation.

In summary, all degradation models analyzed follow the same structure. Typically, a calendaric and a cyclic term is included and both are considered to affect battery capacity additively. Literature mostly agrees upon the shape of the calendaric term, including Arrhenius dependency, a linear or exponential rela-

tionship with SOC and a square root or $t^{0.75}$ function of time. The cyclic term, typically excludes certain variables as indicated in Table 2.1, and the dependency on charge throughput Q is modelled by a square root function or proportional to $Q^{0.87}$. Therefore, I do not expect the major effects and relationships to vary strongly between different degradation models presented here.

2.4 Electric Vehicle User Behavior

Battery degradation is strongly influenced by the usage of an EV related not only to the driving but also the charging behavior. While, for instance, high charging states are known to be associated with increased degradation, recent publications identify phenomena like range anxiety that lead to frequent and full recharging.

Uncertainty about the possibility to recharge and the range limitations have led to the observation of *range anxiety* in field studies (Franke and Krems, 2013b) and user interviews (Eberle and von Helmolt, 2010; Neubauer and Wood, 2014). Franke and Krems (2013a) study EV drivers and their charging behavior and found that users prefer to retain range buffers. While the observed average daily distance is far below the available range, daily recharging is found unnecessary.

Rolim et al. (2012) observed that a large portion of users considered in a similar setting are mostly charging at home and overnight. Other empirical studies from the US that estimate the potential of EV applicability show that the average driving distance per day is between 46 km (NHTS[1]) and 52 km (Pearre et al., 2011). Further empirical evidence from field trials in Japan shows that the majority of private users reconnects a vehicle for full-charging after the last trip of the day to have the full range available the next day (Sun et al., 2015). This behavior was also observed in field trials reported in Zoepf et al. (2013) and Jabeen et al. (2013).

While this frequent full-recharging covers the drivers' potential (not necessarily actual) mobility requirements, higher charging states drive battery degradation and shorten the lifetime of the battery. In this spirit, a trade-off between

[1]National Household Travel Survey, http://nhts.ornl.gov

degradation minimization and range maximization exists. As of yet, this trade-off has not been modeled in a formal fashion, and charging aimed at maximizing battery life so far received scant attention in the literature.

2.4.1 Smart Charging

Field studies in the realm of V2G include battery degradation to evaluate price optimal charging. EVs of private customers are–on average–used only around one hour a day (Kempton and Tomic, 2005). Hence, most EVs have temporal flexibility regarding their recharging phases. A significant body of research has been devoted to exploit this flexibility for example to reduce charging costs incurred given a variable pricing scheme (Valentine et al., 2012; Schuller et al., 2014; Wei and Guan, 2014; Flath et al., 2014). Since V2G introduces additional cycles to the battery this also needs to be accounted for in the economic assessment. Current literature in the field of V2G considers battery degradation by simplified assumptions, such as the calendaric part of aging (Dietze, 2015) or penalties for high power charging and discharging (Peterson et al., 2010a). Other, more recent work from (Wang et al., 2016) combines semi-empirical battery aging, vehicle and ambient temperature models in order to determine the effect of a constant V2G operation mode for peak shaving or frequency regulation in the Californian energy market.

A further strain of work is looking into the optimization of EV infrastructure or their operation in commercial distribution fleet applications. Almuhtady et al. (2014) for instance introduce a degradation-based resource allocation policy to optimally utilize swappable batteries on a fleet level. They thus perform an optimization of the resource usage given transportation requirements. However, battery degradation is merely considered by constant parameters for battery usage. Sweda et al. (2016) are looking into battery degradation aware siting of charging infrastructure along major highways. Furthermore, they elaborate on the performance of different heuristics to determine these charging policies. Since their focus is on optimal infrastructure deployment and usage, battery degradation is considered in terms of the number of overall cycles and a

punishment of high battery SOCs i.e. overcharging. Whereas overcharging–i.e. crossing the upper voltage bound set by the battery manufacturer–is typically not an issue in real-world settings due to proper charging regulations by the battery management system Sweda et al. (2016) do not capture battery degradation in all detail.

In summary, most of the mentioned sources consider battery degradation merely as an additional cost factor, but do not focus on reducing it as a primary objective.

Part II

Data Representation and Prediction

Chapter 3

Data Selection and Reduction

BATTERY degradation is typically analyzed in laboratory settings. However, considered parameters and their combinations in a laboratory environment are limited and far from the dynamics of a real-world scenario. To overcome this issue, the analysis of field data is crucial. The EVs already on the streets offer the potential–not only–for the analysis of degradation that is relevant for signals from components which are exposed to real environmental conditions and user behavior. However, by now it is unclear how to deal with the potential huge amount of data and signals in real-world environments.

EVs are increasingly used in mobility services such as car-sharing. Often, these services are offered and operated by OEMs themselves, taking Drive Now or Car2Go as examples. OEMs are seeking to reduce costs, improve quality and customer satisfaction by offering advanced services. Managerial actions are manifold, ranging from guidance and incentive schemes on how to use a mobility service in a way that extends its lifetime (thereby exploiting potentials to offer the service at lower fees) to predictive maintenance to avoid service level degradation or even car breakdowns during service usage. One primary means of achieving these goals is the exploitation of the vast amount of on-board data gathered from vehicles in the field through telematics or at periodic inspections. In the near future, many vehicles will be transmitting on-data stored in the Electronic Control Unit (ECU) by telematics. This development is supported by the EU-guideline for eCall that needs to be fulfilled by 2018 (EU, 2007).

OEMs are seeking for a deeper understanding of system faults and the re-

maining useful life (RUL) of components. This includes the forecasting of failure rates of technical devices, guarantee and service design (Liu et al., 2013). However, predictive maintenance aims at forecasting the optimal maintenance interval based on performance or parameter monitoring (Deutsches Institut fuer Normung, 2015) as well as minimizing the occurrence of faults, to increase consumer satisfaction.

Vehicle sensor data is acquired and processed by the respective ECU and On-Board Diagnostics (OBD) are performed for the sake of vehicle design, validation and verification, to identify warranty relevant information and for the detection of system faults. Meeting the requirements for real-time processing, the ECU is an embedded system, which has very limited data storage capabilities in an order of magnitude of kB to MB (Sung and Shin, 2015). On the contrary, data loggers that allow for a recording and storage of sensor signals with a high frequency, are limited to the development phases, and therefore rarely represented in series vehicles (Zhang et al., 2009), (Prytz, 2014). Overall, the collection of required sensor signals for the development of new customer services is highly limited by the data storage capabilities of today's ECUs as well as the transmission capacity of telematics. Hence, to reveal the potentials of smart data analytics, intelligent methods are required to extract the information from sensor data that is most relevant to a respective descriptive or predictive analytical task. In this work I focus on the collection of data of EVs in the context of battery degradation.

Battery degradation, is highly driven by user behavior, in terms of driving, charging and environmental factors such as the ambient temperature, as well as the battery management system. An outline on these aspects can be found in Section 2.3. Whereas functional dependencies and interactions of degradation relevant variables are not yet fully understood, it is key to make use of the already large amount of EVs yet in the field to overcome this lack of knowledge. Comprehension of the interplay between dynamic user behavior and battery degradation, is not only crucial for warranty specification, but also for the development of services such as predictive maintenance, eco driving assistance systems or V2G approaches. Therefore, I provide decision support for OEMs on

how to collect sensor data for accurate prediction of system states in terms of capacity fade.

3.1 Sensor Data Acquisition

Due to an increasing number of sensors in different fields such as the automotive industry, industrial production, health sector, mobile devices, fitness and life tracking (quantified self) (Alhonsuo et al., 2016), suitable data acquisition and processing is becoming increasingly relevant in order to make use of the data. However, data reduction is necessary to meet three major issues. Firstly, the challenges of energy consumption of sensors at high sampling frequencies. Secondly, the communications costs that arise when data is transmitted to the base station. Thirdly, the limited data storage on embedded systems (Aggarwal, 2013; Prytz, 2014). Reducing the amount of data can be achieved with different approaches of supervised and unsupervised approaches. Reduction of data (unsupervised) can be achieved for example with principal component analysis and Fourier- and Wavelet-transformations.

In contrast, if the goal is to explain or predict a particular target variable, such as the capacity fade, by using the remaining data variables as explanatory features, the nature of the data reduction problem changes. This case is a regression setting where the loss function is solely related to the error when approximating or predicting the target variable (supervised reduction of information). Here, for instance, methods to select relevant subsets of sensor-signals are advised, using for example shrinkage methods such as the Lasso regression. Also, a coarser-grained representation of the explanatory variables might be beneficial, given a low increase of predictive error. Filtering data by means of sampling techniques has also been successfully applied in regression settings (Aggarwal, 2013).

Aiming at predicting the battery degradation as accurate as possible, under the given restrictions of on-board data storage as well as transmission capabilities, transformations and selection of relevant variables need to be performed and evaluated. To evaluate the trade-off between predictive accuracy and sampled and shrinked subsets of features, I introduce a simulation model based on

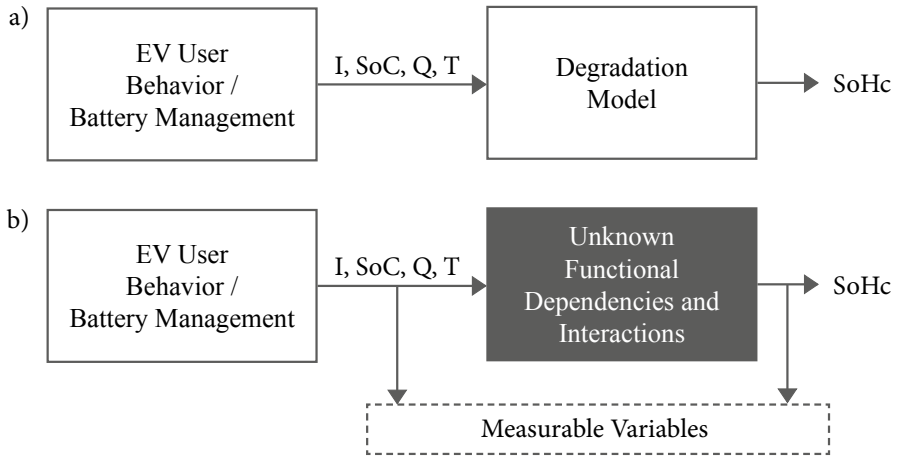

Figure 3.1: Battery stress factors follow from user behavior and battery management system and the corresponding SOH_C results from the degradation model.

real-word driving profiles and a degradation model from literature in the following Section.

3.2 Degradation Simulation Model

Degradation relevant variables, as described in detail in Section 2.3, include the battery current I, the SOC and derived from these variables, the cycle related variables DOD, $\overline{SOC}$ as well as the charge throughput Q. Furthermore, temperature T and time, i.e. the batteries age, need to be considered.

Figure 3.1a indicates how battery stress factors, such as I, SOC, Q and T result from user behavior and the strategy in the battery management system. The degradation model of the respective type of battery reacts on the stress factors and outputs the respective SOH_C. Figure 3.1b depicts the measurable variables, i.e. the stress factors and the SOH_C. The degradation model, however is not known for currently available EVs. In order to evaluate the predictive accuracy of a degradation model that is learned from field data, I create a ground truth of the target variable SOH_C, by using a battery degradation model from literature.

The following subsections detail the simulation of realistic EV user behavior, the parameterization of driver types and ambient conditions as well as the

degradation model.

3.2.1 Trip Generation and Parameterization

The simulation of user behavior, throughout the expected battery life of several years, requires a data set of driving profiles of such length with high resolution. Such a set of driving profiles is generated based on a combination and extension of data from the German mobility panel (BMVBS, 2008) as well as GPS data logs from a publicly available Uber Data Set including 25,000 taxi trips within the San Francisco Bay area (Inc., 2013).

The German Mobility Panel (MOP) is based on the reporting of driving behavior in terms of distance travelled and vehicle location of more than 17,000 households over a period of one week with a resolution of 15 minutes. The mobility panel is separated by the socio-economic background of the participants. However, in this work I focus on the groups of full-time employees and retired persons. These include both the largest number of different profiles as well as the total number of trips. Nine different locations are included in the MOP dataset: *home, work, businesstrip, company trainingcenter, leisure, second home, service, shopping and vacation*.

In order to create driving profiles throughout the lifetime of a EV battery, the one week MOP driving profiles need to be extended to several years. Hence, based on the MOP dataset, the usage profile of an EV is created. Time, distance and locations of trips, arrivals and departures are derived from empirical distributions resulting from MOP data.

This process is depicted in terms of a flow chart in Figure 3.2. User behavior and environmental conditions are initialized in terms of aggression level, driver type, temperature and the charging strategy. The driving profile, in terms of velocity, is determined based on the Uber data set. Therefore, GPS logs are transformed to distances, with a resolution of one second. The resulting speed profiles are then clustered based on their specific speed and acceleration levels to create different levels of aggression. Increased maximum speed and an increased gradient of speed (acceleration) correspond to increased aggressive-

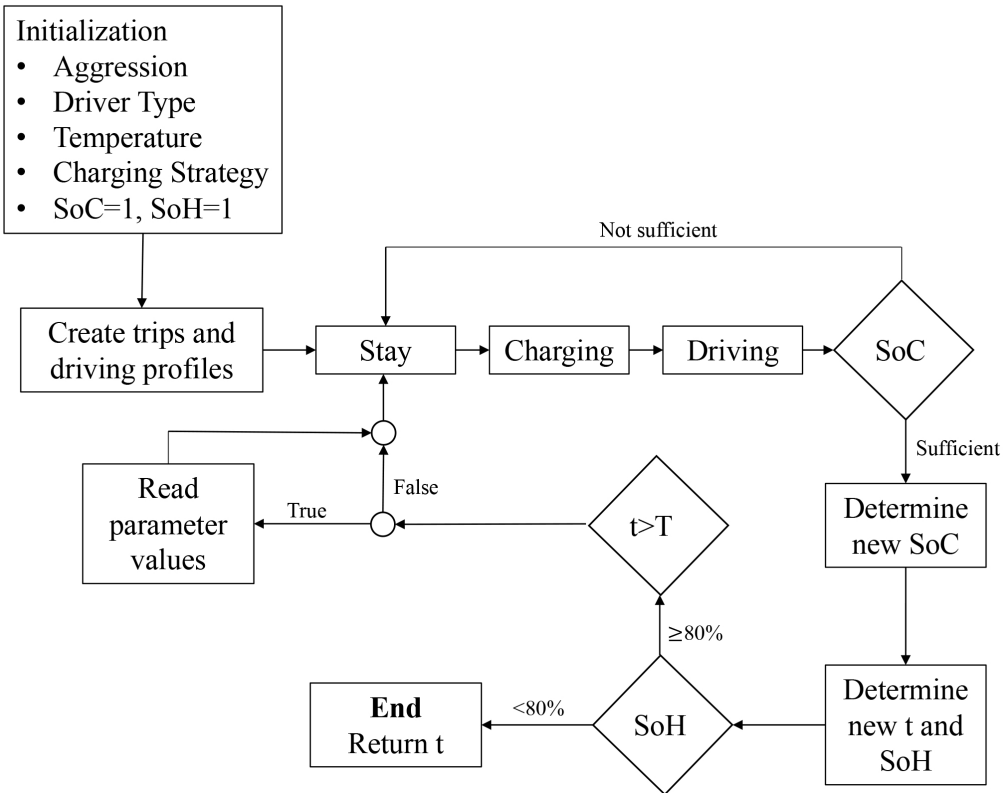

Figure 3.2: Flowchart of driving profile creation

ness. The driver type differentiates between MOP data from retired persons and full-time employees. Ambient temperature is considered a parameter due to its considerable influence on degradation. The temperature profiles employed in this work are based on the year 2015 with a resolution of one hour and are repeated annually. The data on air temperature in Munich, Madrid and Phoenix was taken from several weather data bases (Wetterdienst, 2014; TuTiempo.net, 2014; Underground, 2014). I differentiate between four different charging strategies: *Just-in-time* charging corresponds to a strategy for charging the EV as late as possible, whereat all trips need to be feasible with the available SOC. *AFAP* (as fast as possible) charging, corresponds to a maximization of SOC. With *corridor charging* two bounds are defined for the start and end of charging, *lower bound charging* instead only considers a lower bound.

Following the initialization, as illustrated in Figure 3.2, driving profiles and corresponding trips are generated. Due to the SOC restrictions arising with an

EV, the created trip, including charging and driving, is only performed if the battery is sufficiently charged. If the SOC is not sufficient to perform the trip, another stay is included. Otherwise, the new SOC, time (depending on the duration of the trip) and SoH are determined. This process is repeated until the SoH hits the EoL criterion of the battery. However, if the time T is reached, the initialized parameters can be changed to account for varying user behavior throughout the battery life.

Driving profiles, i.e. trips, locations of stay and charging are generated based on empirical distribution of MOP datasets of full-time employees and retired persons. Three empirical distributions are created, the duration of a stay, the destination and the distance.

Duration of a stay: Based on the collection of all one week MOP profiles, histograms on the duration of stay are created for each time slot of the day differentiated by weekdays and weekends. Following that, empirical distributions are created for each 15 minute time slot of a day, differentiated by weekdays and weekends, resulting in $2 \cdot 4 \cdot 24 = 192$ tables for any of nine available locations. Therefore, for each time of the day, there is an empirical distribution for the duration of stay in each of nine possible locations. For example, at 8 AM on a workday, the duration of stay at WORK is typically longer than at SHOPPING.

Destination: Similar to the approach for *duration of stay* $2 \cdot 96$ tables are created for weekdays and weekends. Furthermore, the empirical, relative frequencies of occurrences of trips from a start location to an end location are added up to empirical distributions. As an example, on a workday at 8 AM the most frequent destinations of full-time employees are BUSINESSTRIP and WORK.

Distances: For each start and end location ($9 \cdot 9$), where start and end location might be identical, relative frequencies are cumulated to empirical distributions. The distance from HOME to SHOPPING, for example, is shorter than the distance from BUSINESSTRIP to HOME.

With the start of the simulation each specific trip is assigned a distance by drawing a random number. That distance remains constant for a given amount of time, typically one year. I have chosen this design to account for the constancy of many daily distances, for example the trip from home to work or shopping,

assumed to be typically similar for a certain period of time. The duration of a stay as well as the next destination are chosen randomly after each trip, based on the empirical distributions. However, SOC restrictions are taken into account, when a driving sequence is calculated and it is only allowed to charge the vehicle at defined locations according to the charging strategy.

Table 3.1 depicts parameters and values used to generate different driving profiles. In total, subsequent analysis are based on $\binom{2}{1} \cdot \binom{5}{1} \cdot \binom{4}{1} \cdot \binom{3}{1} = 120$ different combinations of the parameters considered.

Parameters	*Values*
driver type	Fulltime; Retired
Aggressiveness cluster	1; 2; 3; 4; 5
Charging strategy	Just-in-Time; AFAP; Corridor; Lower Bound
Ambient temperature	Munich; Madrid; Phoenix

Table 3.1: Parameters and values for driving profile generation.

3.2.2 Degradation Model

In the following, I provide the prerequisites for subsequent research by identifying a suitable representation of degradation relevant variables by meeting the constraints of data storage and transmission capacities. Furthermore, this work presents methods on how to transform and process EV degradation related variables in order to achieve a high predictive accuracy.

In a real-world scenario the case of Figure 3.1b applies. The variables arising from user behavior limited by the battery management system as well as the resulting SOH_C are measurable. But the underlying degradation model with it's functional dependencies and interactions is unknown. In order to create a ground truth for further evaluations, a degradation model from the literature is employed to generate SOH_C based on simulations of user behavior.

A review of degradation models is presented in Part I of this thesis. The degradation model developed by (Schmalstieg et al., 2014) includes all relevant variables except for C-rate, and is therefore found to be the most useful to simulate

the usage related degradation progress. The model consists of a calendaric as well as a cyclic component, which are represented in Equation 3.2 and Equation 3.3, respectively. Calendaric and cyclic aging lead to a monotonically decline of the initially available capacity with $t^{0.75}$ and the square root of Q, respectively. Equation 3.1 summarizes the relationship.

$$\begin{aligned} Capacity = 1 &- \alpha_{cal}(T,v) \cdot t^{0.75} \\ &- \beta_{cyc}(\overline{v}, DOD) \cdot \sqrt{Q} \end{aligned} \tag{3.1}$$

$$\alpha_{cal}(T,v) = (7.543 \cdot v - 23.75) \cdot 10^6 e^{-\frac{6976K}{T}} \tag{3.2}$$

$$\begin{aligned} \beta_{cyc}(\overline{v}, DOD) = 7.348 \cdot 10^{-3}(\overline{v} - 3.667)^2 + \\ 7.6 \cdot 10^{-4} + 4.081 \cdot 10^{-3} DOD \end{aligned} \tag{3.3}$$

The battery capacity deployed in the analyses of (Schmalstieg et al., 2014) is much lower (2.15 Ah) than that of a typical traction battery in a EV (in this work I assume a battery capacity of 18.8 kWh - Table 3.2). However, interconnecting many cells in series, results in an overall capacity, meeting the requirements for a traction battery. With a nominal voltage, a number of $18800Wh/(2.15Ah \cdot 3.6V) \approx 2430$ cells need to be connected in series to model the considered traction battery of 18.8 kWh. Practically, the battery stress factors are divided by the number of cells.

3.2.3 Simulated Data Set

The energy required for propulsion results from the sum of energy required for acceleration, rolling and air resistance ((Peterson et al., 2010b), (Linden and Reddy, 2011)). Furthermore, I assume the power drawn from the battery to cor-

respond to the power required to propel the vehicle (Equation 3.4).

$$P_{bat} = P_{propulsion} \tag{3.4}$$

Vehicle specific parameters required to derive the battery current from a driving profile (velocity, V and acceleration, a) include drag coefficient c_w, vehicle frontal area A_{veh}, vehicle mass m_{veh}, nominal battery voltage $U_{nominal}$ and battery capacity C_{Bat}. Moreover, constants are required and include air density ρ, rolling resistance coefficient c_r and gravitational constant g. Table 3.2 summarizes parameters and constants.

$$P_{propulsion} = [F_{acc} + F_{drag} + F_{roll}] \cdot V \tag{3.5}$$

$$\begin{aligned} F_{acc} &= m_{veh} \cdot a \\ F_{drag} &= \frac{\rho}{2} c_w A_{veh} V^2(t) \\ F_{roll} &= c_r \cdot m_{veh} \cdot g \end{aligned} \tag{3.6}$$

The battery current finally results from the definition of electric power:

$$I_{battery} = \frac{P_{Propulsion}}{U_{nominal}} \tag{3.7}$$

Parameters		*Constants*	
c_w	0.29	ρ	$\rho(T)$ $\frac{kg}{m^3}$
A_{veh}	2.38 m^2	c_r	0.013
m_{veh}	1195 kg	g	9.81 $\frac{m}{s^2}$
$U_{nominal}$	360 V		
C_{Bat}	18.8 kWh		

Table 3.2: Assumed vehicle specific parameters and constants.

The resulting battery current is derived from the velocity profiles and is divided by the number of cells (cells are assumed to be connected in series) as de-

scribed in Section 3.2.2. The SOC results from Ampere-hour counting based on charge (positive) and discharge (negative) battery current. Similarly, the charging throughput is derived, employing absolute values for Ampere-hour counting. The degradation model derived from Schmalstieg et al. (2014) deploys the SOC in terms of the cell voltage v. The SOC is in this work assumed to be linearly related to v and mapped from $[0,100]\% \rightarrow [3.2,4.1]$ V. Thus, [3.2, 4.1] V, corresponding to the upper and lower cell voltage bounds, given by the cell manufacturer.

$\overline{SOC}$ and DOD are derived from SOC. However, one cycle is defined such that it contains at least one time slot of driving as well as charging, which starts/ends before the next trip. DOD corresponds to the SOC delta within one cycle and $\overline{SOC}$ is calculated as the $min(SOC) + DOD/2$ within a cycle.

The procedure of trip generation in each time slot, followed by deriving the battery current, and the calculation of battery degradation is repeated until the EoL criterion of 80% is reached. Cumulating time slots correspond to the respective battery age t. The battery temperature is assumed to correspond to the ambient temperature (T).

An overview of the simulated dataset of 120 combinations of the parameters charging strategy, drivers occupation, level of aggressiveness and temperature is given in the descriptive analysis of the following Subsection.

3.2.4 Descriptive Analysis and Initial Prediction Model

Resulting from the simulations, on average the lifetime of a simulated EV battery is 10 years and 80,307 km are covered. This corresponds to 3,931 Ah of throughput. Comparing the covered distance and the overall battery lifetime at the point of reaching the EoL criterion, Figure 3.3 depicts considerable differences comparing full-time employees and retired persons. For each parameter combination that includes 'retired', the covered distance at the same lifetime is in nearly all cases lower than that of 'employees'. For example a lifetime of 10 years leads to approximately 50,000 km covered for 'retired', and approximately 100,000 km for 'employees'. This finding becomes especially interesting

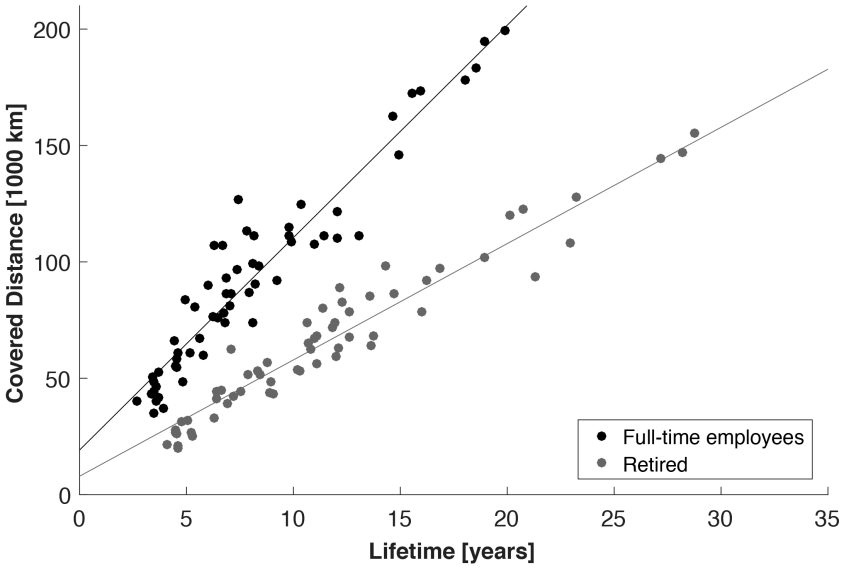

Figure 3.3: Relationship between the lifetime and the distance covered of each parameter combination at the EoL criterion.

when thinking of the guarantee design of currently available EVs. The guarantee that OEMs currently provide, is expected at least with 5-8 years (Nissan Leaf 24 kWh: 5 years or 100,000 km, www.nissanusa.com; BMW i3 18.8 kWh: 8 years or 100.000 km, www.bmw.com; Tesla Model S 85 kWh: 8 years and no range limitation, www.teslamotors.com).

Most OEMs tailor the guarantee on the battery's age or covered distance, but as can be seen from Figure 3.3 the variables considerably diverge depending on the driver type. From the perspective of a full-employed person, it it is beneficial to consider an EV for purchase that guarantees a certain battery lifetime instead of a distance covered. The opposite applies for retired persons.

Analyzing the influence of each parameter value, two linear regression models with categorical variables have been fitted according to Equation 3.8 and Equa-

tion 3.9, for the lifetime and the distance covered, respectively.

$$\begin{aligned} t(EoL) = &\beta_0 + \\ &\beta_1 \cdot DriverType + \\ &\beta_2 \cdot ChargingStrategy + \\ &\beta_3 \cdot AggressivenessCluster + \\ &\beta_4 \cdot T \end{aligned} \tag{3.8}$$

$$\begin{aligned} Distance(EoL) = &\tilde{\beta}_0 + \\ &\tilde{\beta}_1 \cdot DriverType + \\ &\tilde{\beta}_2 \cdot ChargingStrategy + \\ &\tilde{\beta}_3 \cdot AggressivenessCluster + \\ &\tilde{\beta}_4 \cdot T \end{aligned} \tag{3.9}$$

	$t(EoL)$	$Distance(EoL)$
Coefficient	**Estimate** β	**Estimate** $\tilde{\beta}$
Intercept	19.01^{***}	172180^{***}
ChargingStrategy:AFAP	-10.59^{***}	-68182^{***}
ChargingStrategy:Corridor	-8.54^{***}	-53599^{***}
ChargingStrategy:LowerBound	11.09^{***}	-72955^{***}
AggressivenessCluster:2	-1.69^{**}	-24501^{***}
AggressivenessCluster:3	$-0.15(ns)$	-8026^{**}
AggressivenessCluster:4	$0.62(ns)$	$2200(ns)$
AggressivenessCluster:5	$-0.14(ns)$	$-3507(ns)$
DriverType:Retired	4.0^{***}	-24477^{***}
T:Madrid	-2.74^{***}	-21248^{***}
T:Phoenix	-6.67^{***}	-51304^{***}

Table 3.3: Coefficient estimates to explain lifetime in years and distance covered in km corresponding to Equations 3.8 and 3.9. Significance codes: *** 0.001; ** 0.01; * 0.05; 0.1; ns: not significant

The intercept β_0 and $\tilde{\beta}_0$ of both presented regressions with categorical variables corresponds to the reference scenario with *ChargingStrategy*: Just-inTime,

AggressivenessCluster: 1, *DriverType*: Fulltime and the temperature *T*: Munich (Table 3.3). Coming from the reference scenario with an average lifetime of 19.01 years, battery lifetime is reduced significantly by 10.59 and 8.54 years for AFAP and Corridor charging, respectively. On the contrary, Lower bound charging significantly increases lifetime by 11.09 years. Comparing *AggressivenessClusters* indicates that only cluster 2 yields significant reduction of lifetime of 1.69 years, affecting the lifetime much less than the *ChargingStrategy*. Retired on average lead to an increase in lifetime of 4 years compared to fulltime profiles. Both temperature profiles derived from the ambient temperature in Madrid and Phoenix lead to a decrease of lifetime of 2.74 and 6.67 years, respectively. Looking at the distance covered, any parameter combination deviating from the reference scenario leads to a reduction of the distance throughout the battery's lifetime, as indicated by Table 3.3. However, coefficients for *AggressivenessClusters* 4 and 5 are non-significant.

Aiming at the evaluation of predictive accuracy, different transformations and shrinkage of features will be analyzed in the following.

3.3 Initial Prediction Model

In this Section transformed, selected and compressed versions of relevant stress factors are evaluated on their predictive accuracy on battery degradation.

In order to predict the SOH_C two approaches are applied. First, the dependent variable corresponds to the monotonously decreasing SOH_C progress. Second, the delta of SOH_C between two subsequent time slots is used as the dependent variable. In the following, the first and second approach are called *global* and *delta* model, respectively.

The features created from the trip generation, as summarized in Section 3.2.1, and the resulting battery stress factors are shown in Table 3.4.

In order to evaluate the predictive accuracy of features described in Table 3.4 linear regression models are employed. A 10-fold cross validation was carried out to evaluate the out-of-sample prediction error. Models are compared based on their Normalized Root Mean Squared Deviation (NRMSD) and results are

Feature	Description	Frequency
t	Battery age	trip
dist	Covered distance	trip
N_{trip}	Total number of trips	trip
f_{trip}	Frequency of trips in trips per year	trip
Q	Charge throughput	trip
DOD	Depth of discharge per cycle	cycle
$\overline{SOC}$	Average voltage per cycle	cycle
$loc_{beforeTrip}$	Location before trip	trip
$SOC_{beforeTrip}$	SOC before trip	trip
$SOC_{afterTrip}$	SOC after trip	trip
km_{trip}	Length of trip in km	trip
km_{cycle}	Distance covered per cycle in km	cycle
$Q_{perMeter}$	Average consumption per meter	trip
$Q_{perTrip}$	Average consumption per trip	trip
SOC_{rest}	SOC during rest	trip
SOC_{trip}	SOC during driving	trip
SOC_{Δ}	SOC consumption per trip	trip
T_{rest}	Average Temperature during rest (*)	trip
T_{charge}	Average temperature during charging (*)	cycle
V	Average velocity (*)	trip
acc	Average acceleration (*)	trip

Table 3.4: Overview of features, (*) minimum, maximum, mean, median, 25 and 75% quartiles are considered

depicted in Table 3.5. For variable selection and shrinkage the Variance Inflation Factor (VIF), Lasso, Ridge and Elastic Net regression are applied. Furthermore, variable transformation and selection of linear combinations of variables is performed using a combination of principal component analysis and VIF.

Comparing global regression models, none of the shrinked or in dimensionality reduced models outperform the full model–containing 39 features in total according to Table 3.4–in terms of test NRMSD. However, *Global Lasso* and *Global Elastic* result in a comparable predictive accuracy compared to the *Global* model, requiring only 24 and 27 out of 39 features, respectively. Similar to the observations for global regression models, *Delta Lasso* and *Delta Elastic* result in low RMSD but do not outperform the *Delta* model including all 40 features (cf. Table 3.4). Delta models are based on the differentiated and log-transformed $SoHc$.

Modell	**Features/ Dimensions**	**RMSD**	**NRMSD**
Global	39	0.0097	0.0486
Global VIF	15	0.0131	0.0657
Global Lasso	24	0.0105	0.0521
Global Ridge	39	0.0128	0.0640
Global Elastic	27	0.0105	0.0524
Global PCA	12	0.0164	0.0822
Global Cycle	39	0.0143	0.0531
Global Cycle VIF	15	0.0203	0.1017
Global Cycle Lasso	29	0.0149	0.0748
Global Cycle Ridge	39	0.0177	0.0885
Global Cycle Elastic	30	0.0150	0.0750
Global Cycle PCA	13	0.0268	0.1344
Delta	40	0.3909	0.0418
Delta VIF	23	0.3942	0.0422
Delta Lasso	32	0.3912	0.0418
Delta Ridge	40	0.4025	0.0430
Delta Elastic	33	0.3912	0.0418
Delta PCA	15	0.6609	0.0707

Table 3.5: Test error (derived from cross validation) for different regression approaches.

Delta models outperform global regression models in terms of NRMSD. However, *Delta Lasso* and *Delta Elastic Net* models result in NRMSD comparable to that of the full model and require only a subset of 32 and 33 variables of the originally 40 variables.

Whereas NRMSD allows for the comparison of results in different scales, global and delta models can be compared based on NRMSD. However, based on NRMSD the delta models overall show better predictive performance as compared to global models.

Global models generally are based on features generated per trip. Delta models, however, imply cycle based feature updates. According to the definition of a cycle, serval trips can be included within one cycle and the update frequency is reduced. Therefore, also global models are evaluated by using a cycle based feature update frequency, as depicted in Table 3.5, but did not outperform delta or global models.

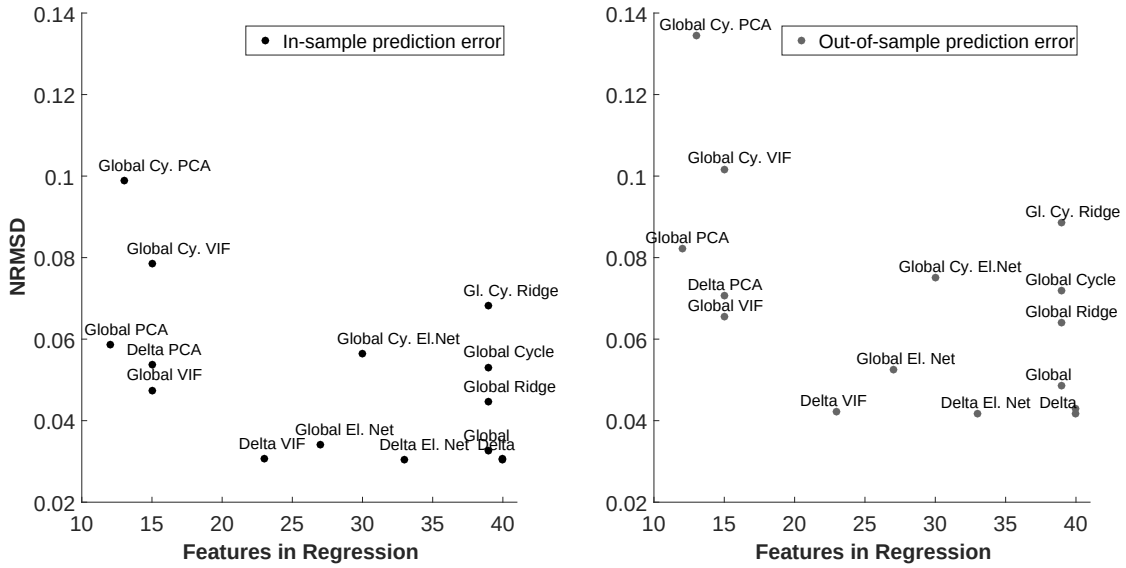

Figure 3.4: In- and out-of-sample NRMSD based on different prediction models

NRMSD for in-sample (training error) and out-of-sample evaluation are depicted in Figure 3.4. Comparing in- and out-of-sample errors does not show strong deviations and I can conclude that non of the models considered leads to considerable overfitting.

The models presented in Table 3.5 either include all variables derived from this simulation or are based on a shrinked subset of variables or linear combinations of models with reduced dimensionality. However, shrinked models that underwent Lasso regression or variable selection using VIF, do no longer include all variables. These models are therefore compared to relevant stress factors that were used for simulation in Table 3.6.

In Table 3.6 only the most promising models from Table 3.5, in terms of NRMSD, are depicted. Each model that underwent variable selection by using VIF is missing variables related to one to two different, relevant stress factors, while only temperature is included in each model. *Global Lasso* and *Global Elastic*

Model	**t**	**Q**	$\overline{SOC}$	SOC	**T**	DOD
Global VIF		x	x	x	x	
Global Cycle VIF	x			x	x	x
Delta VIF	x	x		x	x	x
Global Lasso	x	x			x	x
Global Cycle Lasso	x	x	x		x	x
Delta Lasso	x	x	x	x	x	x
Global Elastic Net	x	x			x	x
Global Cycle Elastic Net	x	x	x		x	x
Delta Elastic Net	x	x	x	x	x	x

Table 3.6: Degradation factors included in shrinked prediction models

Net as well as *Global Cycle Lasso* and *Global Cycle Elastic Net* models miss two and one relevant stress factor, respectively. However, the model performing best in terms of NRMSD–*Delta Lasso*–as well as *Delta Elastic Net* include features related to all relevant stress factors. *Delta Lasso* explicitly includes all features except for: $SOC_{beforeTrip}$, the mean and 75% quartile of T_{rest}, the 25% and 75% quartiles, median and mean of T_{charge} and the 25% quartile of acc. $SOC_{beforeTrip}$ is highly correlated with $SOC_{afterTrip}$ (0.92), SOC_{rest} (0.85) and SOC_{Trip}, such that the information content is reduced. The statistical moments of T_{rest} and T_{charge} are correlated up to 0.99 such that the selection of moments it not surprising. The 25% quartile of acc does not show an absolute correlation greater than 0.67, but might often be close to zero, explaining the low predictive relevance of this feature.

3.3.1 Evaluation of Data Volume

By now, I have evaluated the predictive accuracy of different models given the number of predictors or dimensions included in the model. However, I aim at minimizing the required data storage for the underlying subset or representation of variables. Hence, the data volume is evaluated in this Section.

Data reduction is initially achieved by sampling based on trips or cycles. Assuming a sampling of 1 Hz of four relevant signals (SOC, I, T, Q) corresponds to $(4 \cdot 24 \cdot 60 \cdot 60s \cdot 1Hz = 354600)$ data points per day. Having 2.4 and 1.7

trips per day for fulltime employees and retired persons, respectively, the number of data points reduces considerably by factor $354600/(40 \cdot 2.4) = 3600$ and $354600/(40 \cdot 1.7) = 5082$, with 40 variables in total.

I investigate on the accuracy of models by predictions in terms of the lifetime in years and distance covered in km at EoL ($SOH_C = 80\%$). Results are presented in Table 3.7 using the most promising models of Table 3.5, considering the models with all features included as well as VIF and Lasso models.

Model	**Data volume**	**Prediction error**	
	[kByte/day]	**age [years]**	**Dist. covered [km]**
Global	410	2.48	17,416
Global VIF	146	2.64	17,239
Global Lasso	244	2.49	17,277
Global Cycle	291	1.7	12,077
Global Cycle VIF	109	2.15	15,370
Global Cycle Lasso	164	1.7	12,268
Delta	290	1.72	12,843
Delta VIF	156	1.85	13,516
Delta Lasso	212	1.72	12,842
Parameter model	0	3.7	23,151

Table 3.7: Prediction error in lifetime and distance covered

Evaluating the simplest model as a benchmark, a regression is performed based on the parameter configuration according to Table 3.1, which is indicated by *Parameter model* in Table 3.7. Throughout the battery lifetime, one constant combination of parameters needs to be derived from driving and charging style and the ambient temperature conditions. Therefore, the required data volume is nearly zero. Any other model, indicated in Table 3.7, requires a considerably larger data volume due to updates of trip or cycle based variables. Comparing the predictive accuracy of EoL prediction in terms of lifetime and distance covered, the cycle based, shrinked global model *Global Lasso Cycle* yields the best predictive accuracy–with an average prediction error of 1.7 years and 12,268 km–under minimal data volume of 164 kB per day. The required data volume is well in line with the data storage capabilities of a standard ECU for battery management systems, laying in an order of magnitude of kB to MB. Similar results can

be achieved by applying the *Delta Lasso* model.

3.4 Conclusions and Limitations

A simulation of battery degradation has been developed, that considers dynamic user behavior. Based thereupon, I am able to derive implications for battery EV guarantee design from an OEMs point of view and guarantee (corresponding to EV) choice from an users point of view, that may differ considerably depending on the driving habits of users. Furthermore, different models have been evaluated based on their predictive accuracy and required data storage.

I found that Lasso regression models perform best–compared to dimensionality reduction using PCA and feature selection using VIF–in order to select features with a high predictive accuracy. Moreover, Lasso regression models allow for considerable data storage reductions. A higher predictive accuracy can be achieved based on *Delta* models as compared to *Global* models. Resulting subsets of features can be stored on-board a standard ECU assuming daily submission through telematics.

Using analytical models I have derived a reduced set of features that allows for an accurate prediction of battery degradation in EVs based on standard equipment. This allows for efficient data acquisition in a fleet of EVs, for example of a car sharing service provider, assuming daily data transmission to a home station through telematics.

Such a resulting database allows for detailed analysis of EV user behavior and the related battery degradation. Using prescriptive analytics, optimal behavior can be recommended to the user, which will increase the overall efficiency of EVs including battery lifetime as well as the available range. Car sharing providers may use these insights to map different users, depending on their driving and charging behavior, to the best suited type of EV. The location of newly build charging station can be optimized based on data gathered from a fleet of EVs.

From an OEMs point of view, the data allows accurate predictions of the time to EoL and the development of predictive maintenance approaches. Accurate models will result in greater customer satisfaction and therefore increase the re-

tention. It will also cause customers to use the OEMs proprietary service garages and increase revenue.

The presented analysis is simulation based, and can be enhanced through real-world measurements of degradation related signals.

Chapter 4

Prediction of Battery Degradation

Degradation modeling is typically done by empirical data fitting based on accelerated tests in a laboratory environment. These experiments require tremendous efforts in terms of equipment and time, which limit the combinations of tested parameters as well as the number of observations under equal conditions. Especially temperature is typically elevated under laboratory tests, such that environmental and usage conditions differ considerably compared to real-world scenarios. Due to the complex nature of battery degradation, extrapolation may lead to invalid results. Therefore, in the following Sections an empirical, field-data based battery degradation model is developed, based on a large amount of field-data, resulting from every-day EV usage and real environmental conditions.

4.1 Data Set and Descriptive Analysis

This study is based on anonymous data gathered of a fleet of more than 5000 BMW i3 EVs with a battery capacity of 18.8 kWh (60 Ah) each. Vehicles are operated by private users and in commercial fleets across the world between 2013 and 2017. The data is preprocessed and aggregated on-board the vehicle by the electronics control unit–the battery management system–followed by a periodic, telematics-based transfer. The database contains information about battery stress factors such as cell temperature, SOC, charging events, the covered milage and an estimate for the state of health of capacity SOH_C corresponding to aging.

Currently available articles that are based on a fleet of EVs, do not include analysis of battery degradation (Smart and Schey, 2012; Rodgers et al., 2014; Devie et al., 2012; Franke et al., 2012; Corchero et al., 2014). Solely, Barré et al. (2014) estimate the current capacitive as well as resistive SOH based on logged data from one vehicle. To the best of the authors knowledge, this database therefore composes the largest database, in terms of number of vehicles as well as number of data points that includes SOH_C on-board estimation.

4.1.1 Idle times

EVs offer a large potential to store energy, as they are in idle mode for 96% of the time in a day (Kempton and Tomic, 2005). Making use of flexibility in the timing of recharging, the development of smart charging strategies for EVs is emerging recently. Smart charging typically includes some optimization procedure with the objective of cost minimization in case of V2G (Schuller et al., 2014; Valentine et al., 2012) or load balancing approaches (Peças Lopes et al., 2009), maximization of renewable energy usage or minimization of battery degradation (Schoch, 2016). However, evaluating the economic benefit of any smart charging strategy–especially for V2G approaches–a detailed assessment of battery degradation is crucial. By now, battery degradation is typically–if at all–considered in terms of additional cycling, SOC minimization (Dietze, 2015), charging power penalization (Peterson et al., 2010a; Schuller et al., 2014) or simplified, linearized terms. In contrast, the vast amount of literature in the field of accelerated aging testing indicates that the degradation process is by far not as trivial.

Applying any kind of charging coordination, sufficient idle time to allow for the required flexibility is required. By now, (smart) charging strategies are developed based on driving profiles acquired from conventional vehicles with an internal combustion engine (ICEV) (Schuller et al., 2014; Flath et al., 2014; Peterson et al., 2010a; Pearre et al., 2011; Neubauer et al., 2012). However, it is expectable that EV user behavior in terms of driving will differ from that of ICEVs. Based on the dataset at hand the ratio of time in which driving or charging occurs and the total time of operation can be analyzed. One minus that ratio then allows

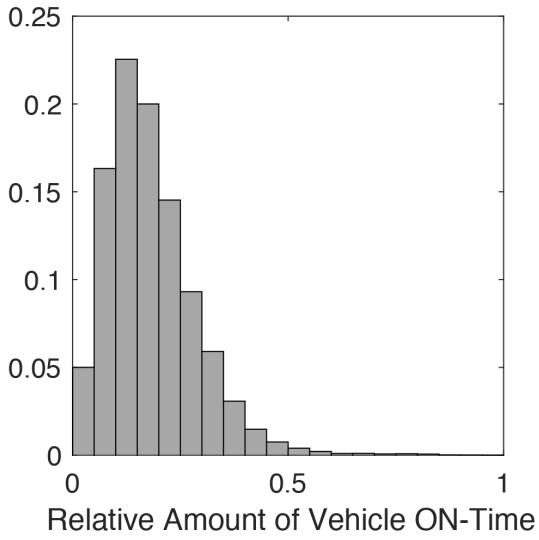

Figure 4.1: Histogram of the ratio between vehicle on-time and the total time.

to provide a sound estimate of idle times in which neither driving nor charging takes place.

The histogram in Figure 4.1 depicts the ratio between the time of driving or charging (total time on) and the total time. On average a vehicle spends 18% of the time of a day either charging or driving (median 16%). This corresponds to a theoretic daily time spent in idle mode of 82% and a mean charging time window (charging flexibility) of more than 19 hours, presuming the availability of charging opportunities.

4.1.2 Comparison Between Test and Real-World Conditions

For accelerated aging tests performed in a laboratory environment, combinations of parameters are typically very limited due to efforts in time and equipment. This results in a gap between real-world and tested conditions that is investigated in the following.

Test Conditions for Calendaric Aging

Figure 4.2 shows the considered ranges for temperature T and SOC. The SOC range between 0 and 100% is covered by the work of Schmalstieg et al. (2014). Kaebitz et al. (2013); Marongiu et al. (2015) consider a slightly smaller range of 20-100%, while Sarasketa-Zabala et al. (2016) only conduct tests in the range of 30 to 90% SOC. Temperature is considered between 35 and 50°C (Schmalstieg et al., 2014), 25 and 60°C (Kaebitz et al., 2013), 40 and 70°C (Marongiu et al., 2015) and 30 to 50°C (Sarasketa-Zabala et al., 2016). In summary, calendaric tests in a laboratory setting are performed in SOC ranges nearly covering the complete range of 0-100%, while ranges of temperature lie between 25 to 70°C.

Test Conditions for Cyclic Aging

Figure 4.3 illustrates the laboratory test settings for ΔDOD and $\overline{SOC}$. All valid combinations of ΔDOD and $\overline{SOC}$ lie in the 2-simplex defined by the vertices (ΔDOD, $\overline{SOC}$) (0, 0), (100, 0) and (50, 100). Therefore, the maximum ΔDOD of 100% can only be reached in combination with a $\overline{SOC}$ of 50%, to account for SOC bounds of 0 and 100%, which cannot be crossed.

From Figure 4.3 it becomes evident, that the degradation model presented by Schmalstieg et al. (2014) covers a wide range of cyclic parameter combinations of ΔDOD and $\overline{SOC}$. The models presented by Kaebitz et al. (2013), Marongiu et al. (2015) and Sarasketa-Zabala et al. (2016) instead sample the range of possible combinations of cyclic parameters very sparsely.

Real-World Conditions for Calendaric Aging

In order to evaluate the differences between parameterizations in a laboratory setting and real environmental and usage conditions, I consider empirical cumulative density functions of temperature and SOC resulting from the analysis of field data. As depicted in Figure 4.4, the observations of SOC lie in the range between 0 and 100%. However, due to the high frequency that users choose to recharge and the high SOC at begin of charge, only 14% of observations lie below 50% SOC. In summary, compared to conditions in the field, SOC range is

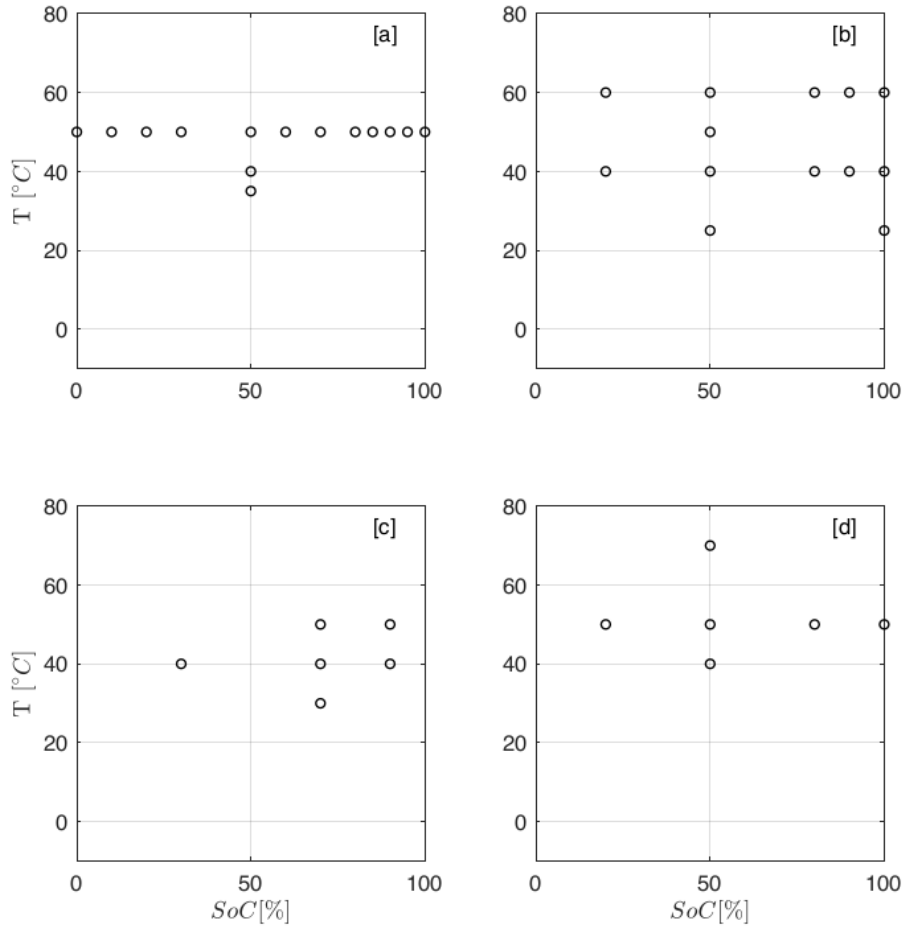

Figure 4.2: Summary of calendar test conditions in the literature [a]-[d] (Schmalstieg et al., 2014; Kaebitz et al., 2013; Sarasketa-Zabala et al., 2016; Marongiu et al., 2015).

sufficiently covered in the degradation models from laboratory tests presented above.

As Figure 4.4 indicates, more than 80% of observations of temperature from field data lie below 25°C. Hardly any (less than 1%) of temperature measurements can be found above 30°C, and most observations lie in the range between 10-20°C. Comparing field data-based measurements to those considered in laboratory settings, lying between 25 and 70°C as presented above, temperature ranges diverge considerably.

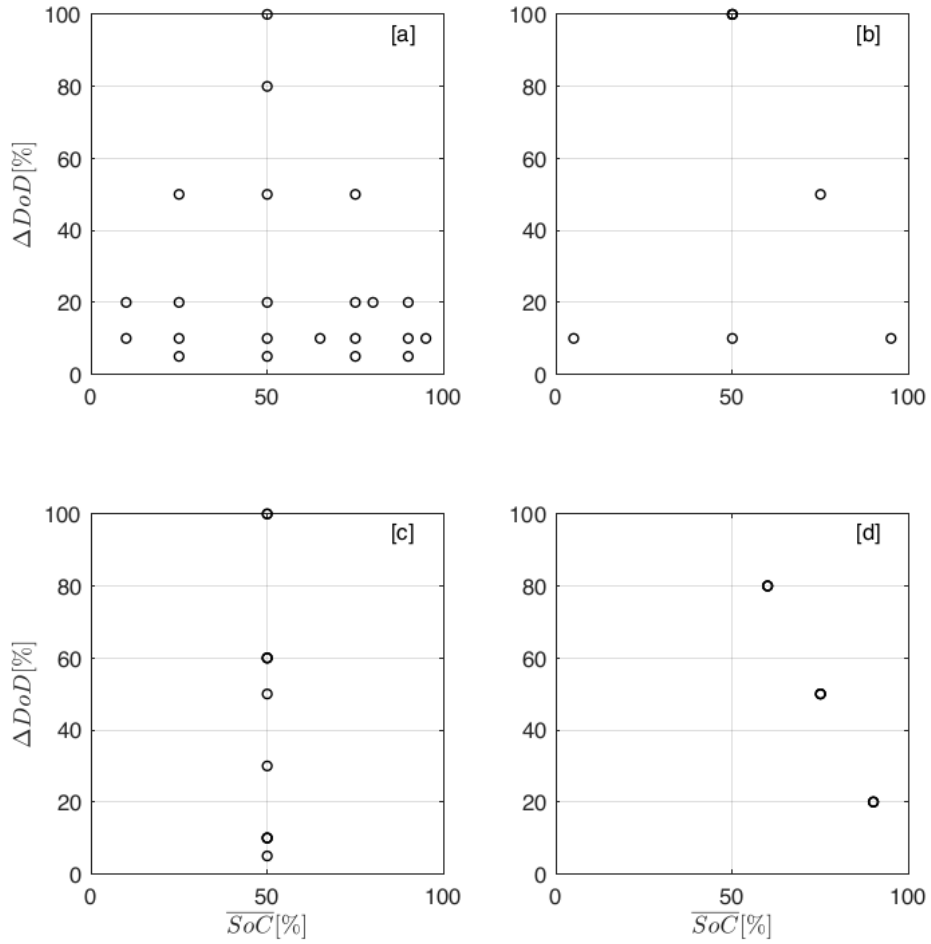

Figure 4.3: Summary of cyclic test conditions in the literature [a]-[d] (Schmalstieg et al., 2014; Kaebitz et al., 2013; Sarasketa-Zabala et al., 2016; Marongiu et al., 2015).

Real-World Conditions for Cyclic Aging

Figure 4.5 depicts a two dimensional, empirically observed histogram of combinations of *DOD* and $\overline{SOC}$. As illustrated by the histogram, the aforementioned 2-simplex of feasible combinations of *DOD* and $\overline{SOC}$ is nearly, completely covered. However, the edge between the right outer vertex and the top vertex shows a density maximum of entries. This observation is inline with the SOC distribution illustrated in Figure 4.4 and indicates frequent and full recharging at high SOCs. Obviously, the observed charging behavior, and thus SOC, *DOD* and $\overline{SOC}$ distributions, result from the current absence of any smart charging strat-

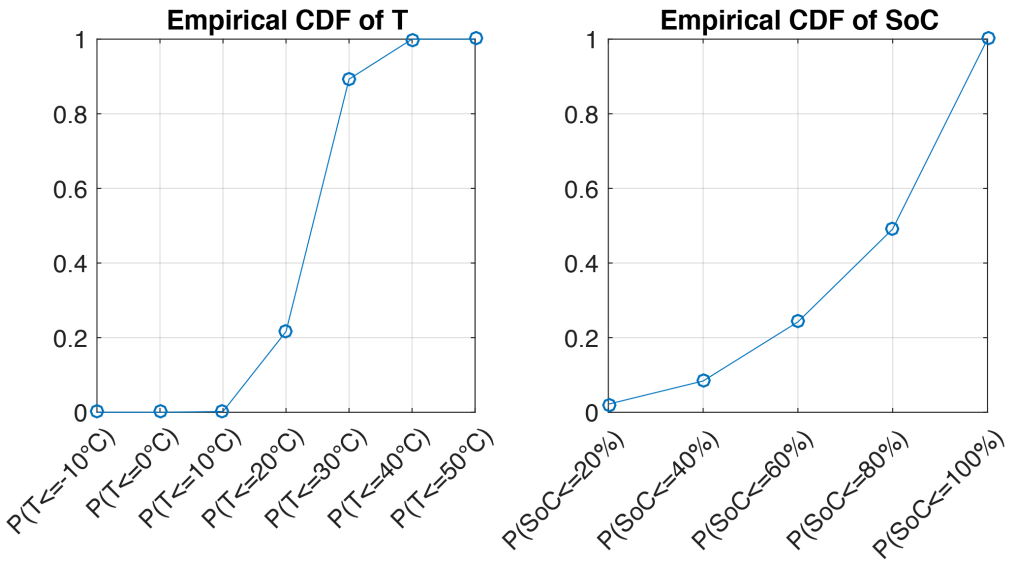

Figure 4.4: Empirical cumulative density functions (CDF) of temperature and SOC.

egy. Therefore, recharging to a SOC less than 100% would be mostly inconvenient for users without automation, but also limit range. Frequent recharging at high SOCs at the beginning of the charging process can be explained by range anxiety (Neubauer and Wood, 2014; Franke and Krems, 2013b) and habituality (Rolim et al., 2012). However, the observed characteristics related to SOC might change with the introduction of smart charging strategies.

In summary, test conditions in terms of SOC, DOD and $\overline{SOC}$ correlate well with usage conditions observed from field data. In contrast, temperatures differ considerably and hardly overlap. Whereas elevated temperature is required to achieve acceleration of aging, the external validity of the results is questionable and it is unclear to what degree results can be extrapolated to realistic temperature conditions.

Aiming to approach this question, I extrapolate degradation models provided in the literature (and presented in Table 2.1) to realistic temperature conditions, while focusing on the temperature dependent calendaric term. I calculate the time to end of life (EoL) for different combinations of temperature and SOC and will discuss the results subsequently.

Extrapolations of the time to EoL are illustrated in Figure 4.6 for calendaric

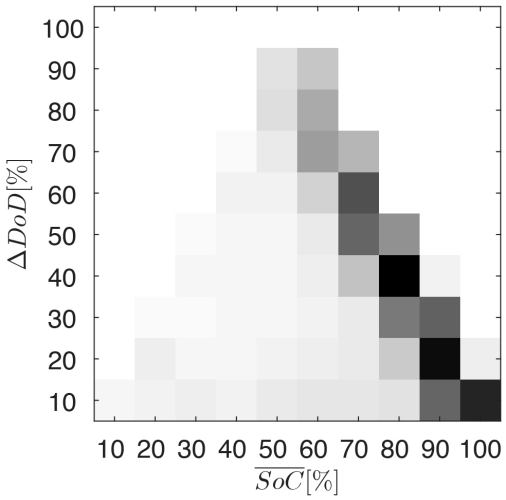

Figure 4.5: Empirical, field-data based, two dimensional histogram of combinations of DOD and $\overline{SOC}$. Darker colored tiles: High number of observations, White tiles: No observations.

degradation presented in Schmalstieg et al. (2014); Marongiu et al. (2015); Ecker et al. (2012) and tested parameter combinations are indicated by red stars. As I have found in Section 4.1.2, the observed temperature under real environmental and usage conditions lies in the range of -10 to 50$^{\circ}C$, with most observations between 10 to 20$^{\circ}C$. Extrapolations to such temperature ranges would lead to extraordinarily high lifetimes of more than 1000 years, especially in combination with a low SOC for the degradation model presented in Schmalstieg et al. (2014). Calendaric degradation models presented in Marongiu et al. (2015) and Ecker et al. (2012) result in maximum battery lifetimes of approximately 60 and 100 years. However, as aforementioned, the validity of such results is unclear.

Validity is further questioned by literature. Different aging mechanisms have been found above and below 20$^{\circ}C$ Lam et al. (2011) and an inverted Arrhenius relationship for temperatures below 25$^{\circ}C$, where decreasing temperature leads to increased degradation (Waldmann et al., 2014).

Therefore, in the following an analytic, field data-based, empirical degradation model is developed, that is valid under real operation and especially temperature conditions. The aim is to find the relationship between usage conditions and the resulting SOH_C based on the field data at hand by using statistical

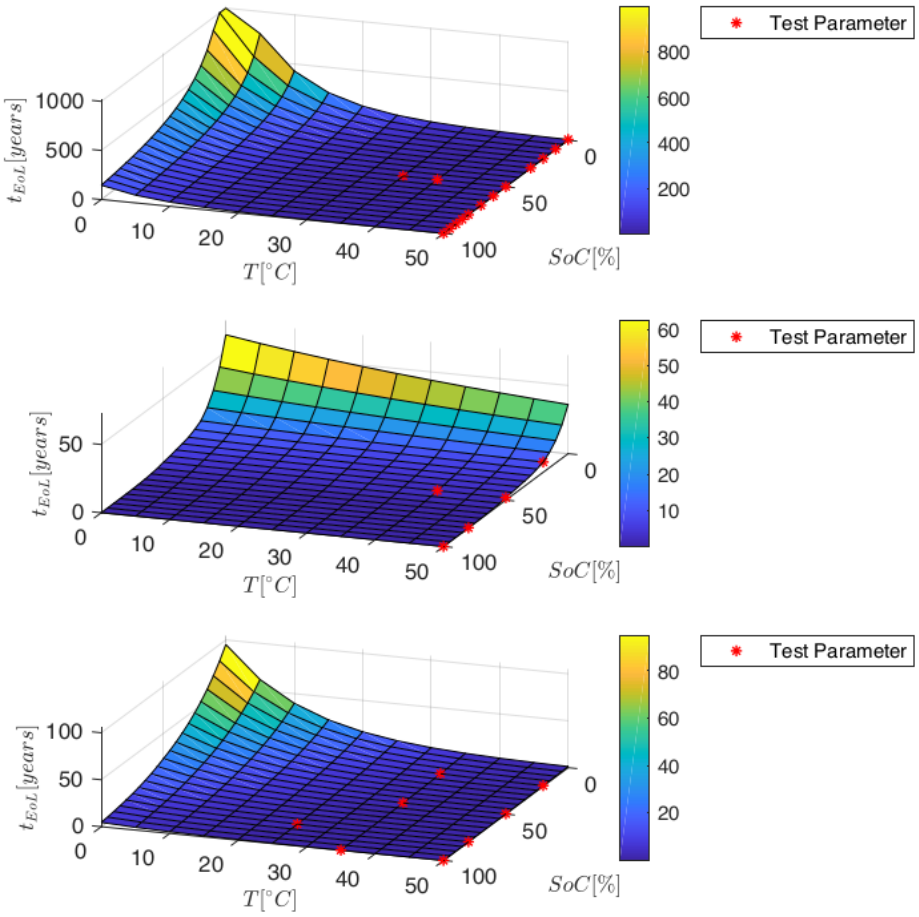

Figure 4.6: Time to EoL for the calendaric term presented by Schmalstieg et al. (2014); Marongiu et al. (2015); Ecker et al. (2012) (from top to bottom.

learning and predictive analytics.

4.2 Models

I perform statistical learning based on the field data of approximately 5000 EVs. The dependent variable–SOH_C–is estimated on-board the vehicle by the battery management system. Independent variables–features–either result from on-board measurements or estimations and will be described in more detail in the corresponding model paragraph.

Before the data can be used for prediction purposes, it is preprocessed and

cleaned by several plausibility steps. After preprocessing, the dataset is composed of approximately 5000 vehicles and 180,000 readouts. Throughout the time horizon of 2013 to 2017 vehicles have been put into operation at different times and total vehicle operating time ranges between 1 month and 2.2 years. In the following model flexibility is considered as the number of coefficients in the model.

4.2.1 Variables

In this dataset several variables related to battery degradation are available as potential predictors. This includes time t and charge throughput Q, which are given in days and Ampere-hours at the point of readout, respectively. Temperature, SOC, charge and discharge current are used in aggregated form by means of histograms including the history since the previous readout. In a preprocessing step, each histogram is probability normalized, leading to an empirical density function of temperature, SOC as well as battery current I for charging and discharging. Each bin of a histogram is used as a variable in the modeling of the following fitting functions. The advantage of this approach is the equal sampling of each variable, while allowing to incorporate the history of non-monotonic variables. Moreover, any linear or non-linear functional dependency can be modeled by learning a separate coefficient β for each histogram bin. Degradation relevant variables DOD and $\overline{SOC}$ are represented by scalars, and derived from the median of the history of DOD and $\overline{SOC}$ between two successive telematic readouts.

Apart from the main degradation relevant variables, the dataset provides access to other variables that might indirectly influence degradation. This includes the number of trips N_{trip}, the mileage MIL, a histogram of the charging duration $\mathbf{t_{chg}}$, the temperature before and after charging T^0_{chg} and T^1_{chg}, the relative amount of time the vehicle spends charging or driving (not in idle mode) t_{ON} and the dichotomous variable EV, which indicates whether the vehicle has an additional internal combustion engine (PHEV). In summary, this leads to a number of 50 dependent variables and one independent variable, the SOH_C.

Box-Cox Transformation of SOH_C

Whereas the representation of the variables T, SOC, I and charging duration in terms of empirical density functions allows to consider the variables in any functional relationship, also time and charge throughput are typically considered non-linearly. However, literature does not agree on a functional relationship. Exploiting the dataset with a tremendously large number of observations, I use Box-Cox transformation of the dependent variable SOH_C to systematically find the most suitable power transform.

For time, the functional relationship reported in literature is t^{k_1}, with $k_1 = 0.5$ (Marongiu et al., 2015; Sarasketa-Zabala et al., 2016; Ecker et al., 2012, 2014; Kaebitz et al., 2013; Wang et al., 2014), or $k_1 = 0.75$ (Schmalstieg et al., 2014). While charge throughput can be modeled implicitly by the number of equivalent full cycles (Marongiu et al., 2015), fewer literature explicitly includes a power transform Q^{k_2}. However, k_2 ranges between 0.5 (Schmalstieg et al., 2014) and $k_2 = 0.87$ (Sarasketa-Zabala et al., 2016).

Following the findings from literature, I systematically search for the power transform of SOH_C that allows to apply linear regression, i.e. that leads to a linear regression model with normally distributed errors, for both time and charge throughput, respectively. I define the searching space for λ_1 and $\lambda_2 \in [0,2]$, that transforms SOH_C into SOH'_C.

$$SOH'_C = \begin{cases} \frac{SOH_C^{\lambda}-1}{\lambda} & (\lambda \neq 0) \\ log(SOH_C) & (\lambda = 0) \end{cases} \tag{4.1}$$

Based on the empirical observations I find $\lambda_1 = 2$ and $\lambda_2 = 2$, which leads to the relationship in Equations 4.2 and 4.3. Whereas the transformed variables are used in a linear regression model, I can focus on the power transform and drop the linear transformations resulting from Box-Cox. Applying simple math leads to a transformation of the independent variable instead of SOH_C, as indicated in Equations 4.2 and 4.3, with each a scalar coefficient β.

$$\frac{SOH_C^2 - 1}{2} = \beta_i t \qquad \rightarrow \qquad SOH_C \propto \beta_j t^{1/2} \tag{4.2}$$

$$\frac{SOH_C^2 - 1}{2} = \beta_l Q \qquad \rightarrow \qquad SOH_C \propto \beta_k Q^{1/2} \tag{4.3}$$

4.2.2 Intercept Model

In the following Sections, different regression models are compared by their predictive accuracy on unseen samples and their flexibility in terms of number of coefficients. The most inflexible regression model is the intercept model, which is represented by the simple observation mean of SOH_C. The intercept model is introduced for comparison. The only coefficient β_0 of this model is fitted using ordinary least squares (OLS) and the test Mean Square Error (MSE) is derived by 10-fold CV.

R^2 is an in-sample, training-based metric representing the data variance explained by the model and ranges between 0 (no variance explained) and 1 (all variance explained). The training Root Mean Square Error (RMSE) is the square root of MSE based on observations in the training sample. Test RMSE, in contrast is the square root of the test MSE that needs to be minimized to find the best model. Both error measures, training and test RMSE are given in the unit of Ah and in case of the intercept model, test as well as training RMSE are equal to 2.2 Ah.

The intercept model is presented as a benchmark or lower bound with minimal predictive accuracy, since it does not include any of the degradation relevant variables, but the intercept β_0.

$$\widehat{SOH_c} = \beta_0 \tag{4.4}$$

Calendaric Model

As summarized in Section 1, degradation models are typically divided into a component of calendaric and cyclic aging. Therefore, a prediction model is presented that includes variables that have been identified to affect the calendaric term, i.e. temperature (T), SOC and time.

Considering calendaric aging terms for the parameters (Spotnitz, 2003; Kaebitz et al., 2013; Schmalstieg et al., 2014), the calendaric model is depicted in Equation 4.5. It includes time, T and SOC as linear variables, as well as interaction terms between time and T and SOC. β_1, β_2 and t are scalar per observation and **T** and **SOC** represent vectors who's elements include each respective histogram bin. Therefore, $\boldsymbol{\beta_3}$ to $\boldsymbol{\beta_6}$ are vectors containing a coefficient for each bin in the histogram.

$$\widehat{SOH_c} = \beta_1 + \beta_2 \cdot t^{0.5} + \boldsymbol{\beta_3} \cdot \mathbf{T} + \boldsymbol{\beta_4} \cdot \mathbf{SOC} + \boldsymbol{\beta_5} \cdot \mathbf{T} \cdot t^{0.5} + \boldsymbol{\beta_6} \cdot \mathbf{SOC} \cdot t^{0.5} \quad (4.5)$$

Compared to the intercept model, the calendaric model leads to considerable improvement in all metrics (Table 4.1). With a number of 40 variables in the model, the decrease of training RMSE from 2.2 to 1.24 Ah is not surprising. However, also test RMSE can be remarkably decreased from 2.2 Ah to 1.24 Ah. Even though the number of variables is quite high, the similar range of training- and test RMSE indicates high generalizability of the model, and high predictive accuracy.

Model	$\mathbf{R^2}$	**Training RMSE [Ah]**	**Test RMSE [Ah]**	**Number of Coefficients**
Calendaric	0.67	1.24	1.24	40
Cyclic	0.43	1.64	1.64	20
Calendaric+Cyclic	0.70	1.19	1.18	59
Linear	0.63	1.32	1.31	44
Lasso (Linear)	0.63	1.32	1.31	42
Interactions	0.77	1.03	240.72	947
Lasso (Interactions)	0.77	1.04	1.04	582
Quadratic	0.70	1.20	1.18	84
Lasso (Quadratic)	0.70	1.20	1.18	73

Table 4.1: Accuracy of Degradation Models

Cyclic Model

After analyzing the predictive accuracy of the calendaric model, the cyclic model is evaluated. Following previously reported degradation models (Jossen and Weydanz, 2006; Bashash et al., 2011; Linden and Reddy, 2011; Kaebitz et al., 2013; Barré et al., 2013; Ramadass et al., 2003), the variables considered in this work include charge throughput Q, DOD, $\overline{SOC}$ and the current I.

The coefficients β_7 to β_{10}, β_{12} and β_{13} in Equation 4.6 are scalar and $\boldsymbol{\beta_{11}}$ as well as $\boldsymbol{\beta_{14}}$ are represented by vectors of coefficients according to the dimension of the current histogram.

$$\begin{aligned}\widehat{SOH_c} &= \beta_7 + \beta_8 \cdot Q^{0.5} + \beta_9 \cdot DOD + \beta_{10} \cdot \overline{SOC} + \boldsymbol{\beta_{11}} \cdot \mathbf{I} + \\ &\quad \beta_{12} \cdot DOD \cdot Q^{0.5} + \beta_{13} \cdot \overline{SOC} \cdot Q^{0.5} + \boldsymbol{\beta_{14}} \cdot \mathbf{I} \cdot Q^{0.5}\end{aligned} \tag{4.6}$$

Comparing the results depicted in Table 4.1, the cyclic model with a test RMSE of 1.64 Ah does not outperform the calendaric model with a test RMSE of 1.24 Ah.

Therefore, both the calendaric and cyclic components are evaluated in a combined model.

Calendaric and Cyclic Model

The calendaric and cyclic model combines both above described terms additively. This leads to the scalar coefficients β_{15}, β_{16}, β_{21} to β_{23}, β_{25} and β_{26} as well as the vectors of coefficients $\boldsymbol{\beta_{17}}$ to $\boldsymbol{\beta_{20}}$, $\boldsymbol{\beta_{24}}$ and $\boldsymbol{\beta_{27}}$ in Equation 4.7. This sums up to a total of 59 coefficients that need to be learned from the data.

$$\begin{aligned}\widehat{SOH_c} &= \beta_{15} + \beta_{16} \cdot t^{0.5} + \boldsymbol{\beta_{17}} \cdot \mathbf{T} + \boldsymbol{\beta_{18}} \cdot \mathbf{SOC} + \boldsymbol{\beta_{19}} \cdot \mathbf{T} \cdot t^{0.5} + \boldsymbol{\beta_{20}} \cdot \boldsymbol{SOC} \cdot t + \\ &\quad \beta_{21} \cdot Q^{0.5} + \beta_{22} \cdot DOD + \beta_{23} \cdot \overline{SOC} + \boldsymbol{\beta_{24}} \cdot \mathbf{I} + \\ &\quad \beta_{25} \cdot DOD \cdot Q^{0.5} + \beta_{26} \cdot \overline{SOC} \cdot Q^{0.5} + \boldsymbol{\beta_{27}} \cdot \mathbf{I} \cdot Q^{0.5}\end{aligned} \tag{4.7}$$

Derived from Table 4.1, the combined model that uses information from both calendaric and cyclic variables, the test RMSE can be further reduced to 1.18 Ah compared to previously introduced models. Compared to the battery capacity of 60 Ah at begin of life, this corresponds to an error of less than 2%.

Full Models

In order to identify relevant variables, interactions and relationships systematically, all available variables are initially included in the full models. A linear, a quadratic and a model with interaction terms are considered, with a number of 44, 84 and 947 coefficients, respectively. The training- and test error can be derived from Table 4.1.

Comparing the in-sample measures R^2 and training RMSE, the model including interactions with the highest flexibility leads to the best results, as expected. However, the out-of-sample measure–the test RMSE–which needs to be minimized, is much larger than the training RMSE. Thus, it can be concluded that the interaction model with 947 coefficients leads to extreme overfitting and has high variance. It therefore has hardly any predictive accuracy nor generalizability. The interaction model, on the other hand, has a high training as well as test error and therefore high bias. Instead, the model that includes calendaric and cyclic terms as well as the quadratic model, leads to the smallest test error in terms of RMSE compared to the linear and interaction model, before the Lasso shrinkage is applied.

In order to systematically shrink models to a subset of coefficients that lead to the best predictive accuracy and to solve the trade-off between bias and variance, Lasso is applied to the full models, i.e. the linear, quadratic and interaction model. As depicted in Table 4.1, Lasso reduces the linear model to a subset of 42 of 44 variables, while the test RMSE remains constant at 1.31 Ah. The quadratic model can be reduced from 84 to 73 coefficients and the test error remains constant at 1.18 Ah. Considerable improvement can be achieved by applying Lasso to the interaction model. In this case the number of coefficients is reduced from 947 to 582. These 582 most relevant coefficients decrease the test error from a large number of more than 240 Ah to 1.04 Ah.

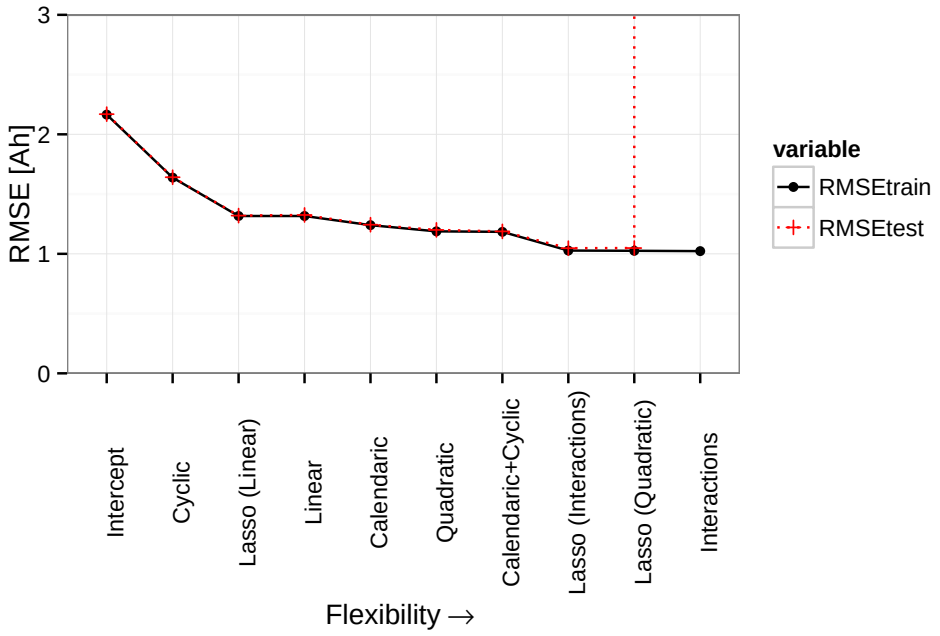

Figure 4.7: Test and train error for degradation models with increasing flexibility.

The interaction model, shrinked by Lasso, corresponds to the model with the smallest test error and solves the bias-variance trade-off. Therefore, it has the highest predictive accuracy and high generalizability.

Graphically, all considered degradation models are depicted in Figure 4.7. It can be observed that training and test error are in similar ranges. Only the interaction model differs, where the test error clearly overshoots the minimal training error, indicating clear overfitting of the training data.

I have found that literature-inspired models that under test, i.e. the calendaric, the cyclic and the combined calendaric and cyclic model, are outperformed in terms of test RMSE by the shrinked, but still more detailed interaction model. This indicates that the currently performed accelerated aging tests presented in literature, do not cover all degradation relevant effects and interactions.

Coefficient Interpretation

To simplify the interpretability of the coefficients, coefficient sizes of the linear model are analyzed for the histogram-based variables temperature, SOC and

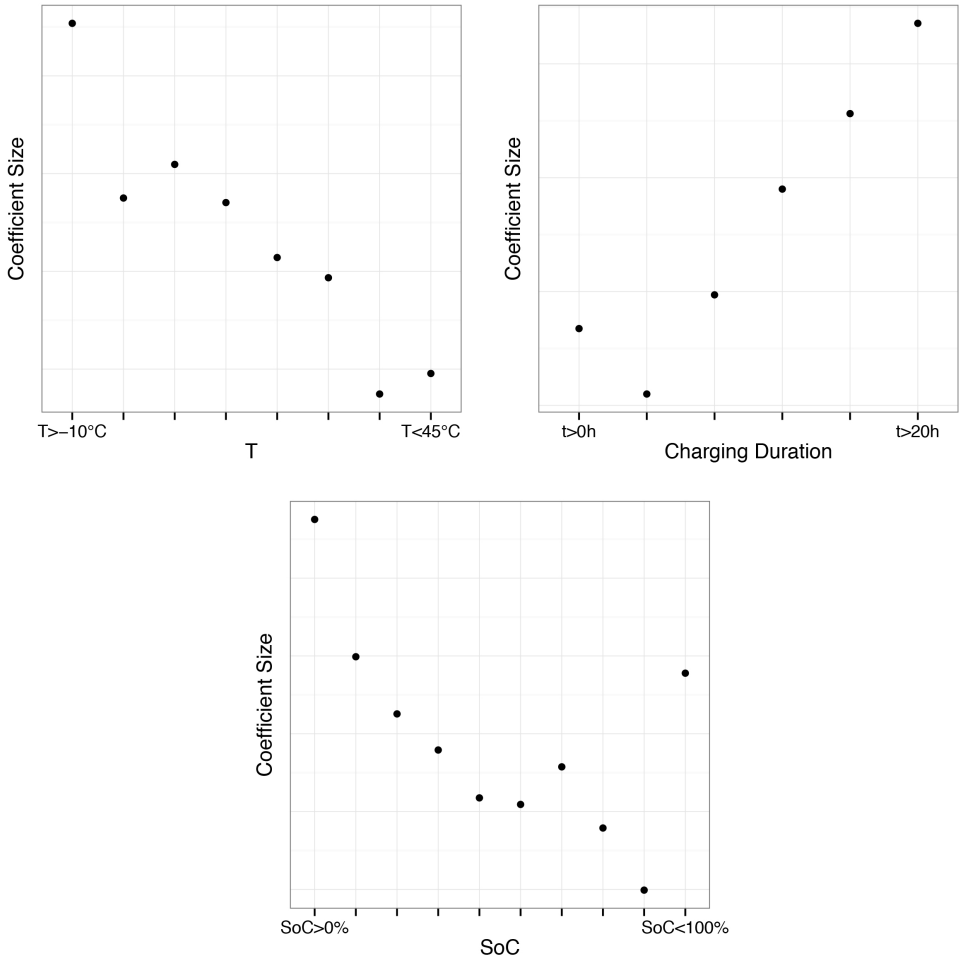

Figure 4.8: Coefficient size of temperature, SOC and charging duration (bins are not equidistantly spaced).

charging duration.

Figure 4.8 depicts the relative coefficient size for each bin of the temperature histogram. While temperatures between -10 to 45°C can be observed in the field data, the distribution of bins is non-equidistant and bin widths are not constant. However, larger relative coefficient sizes correspond to increased degradation for the respective bin. The overall tendency corresponds to the behavior reported in literature and degradation is increased with increasing temperature.

Figure 4.8 illustrates the relative coefficient sizes for the bins in the SOC histogram (bins are not equidistantly spaced). An increasing SOC leads to increased degradation up to 50% (fourth point from the right). Then, the observed relation-

ship is inverted before the degradation maximally increases with increasing SOC at an average SOC of 85% (second point from the right). However, a high SOC above 95% (rightmost point) leads to a rather low coefficient size and therefore reduced degradation, which seems counterintuitive as compared to the findings from literature. The observed behavior, however, is an artifact and results from the fact that in the SOC-histogram the nominal SOC is used (SOC relative to the nominal capacity). The rightmost SOC-bin is only filled for new batteries and therefore indicates low degradation.

The relative coefficient sizes for the histogram of charging duration are depicted in Figure 4.8. Apart from very low charging durations, the coefficient sizes clearly indicate decreasing degradation with increasing charging duration. Charging duration, however, depends on several factors. A low C-rate for example increases charging duration and in this case leads to reduced degradation. This corresponds to the findings from literature where high C-rates typically increase degradation.

	Degradation Relevant Variables							Other Variables					
Model	**SOC**	**I**	**T**	**DOD**	$\overline{\mathbf{SOC}}$	**t**	**Q**	$\mathbf{N_{trip}}$	**MIL**	$\mathbf{t_{chg}}$	$\mathbf{T_{chg}^{0/1}}$	$\mathbf{t_{ON}}$	EV
Calendaric	✓	x	✓	x	x	✓	x	x	x	x	x	x	x
Cyclic	x	✓	x	✓	✓	x	✓	x	x	x	x	x	x
Calendaric+Cyclic	✓	✓	✓	✓	✓	✓	✓	x	x	x	x	x	x
Linear	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lasso (Linear)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Interactions	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lasso (Interactions)	✓	✓	x	✓	✓	✓	✓	x	x	✓	x	x	✓
Quadratic	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lasso (Quadratic)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	x	✓

Table 4.2: Overview of degradation relevant variables. ✓: Variable is considered in the degradation model, x: Variable is not considered.

Table 4.2 depicts the variables included in the different empirical degradation models described in this article. Degradation relevant variables have been identified in literature as summarized in detail in Section 4.1.2. Other variables correspond to variables that are available in the underlying dataset and influence battery aging indirectly, such as the number of trips N_{trip}, the mileage MIL, the histogram of charging duration $\mathbf{t_{chg}}$, the temperature before and after charging

$T_{chg}^{0/1}$, the relative amount of time of vehicle operation t_{ON} and whether or not the vehicle is a pure EV or has an additional internal combustion enginge.

The calendaric, cyclic model include subsets of degradation relevant variables, and the combined model includes all degradation relevant variables. The linear, quadratic and the interactions model are composed of all variables relevant for degradation as well as other variables. Lasso is used to shrink the set of coefficients and hence the set of variables is reduced to the most relevant ones. However, Lasso reduces the linear model merely by some bins of the temperature and charging duration histograms. Similar effects can be observed when applying Lasso to the quadratic model. The resulting subset of variables no longer includes the relative operating time of the vehicle t_{ON}, as well as some bins of the histograms t_{chg}, I, SOC, T and the quadratic term of T_{chg}^{1}.

The best model in terms of test error has been identified as the interactions model after Lasso shrinkage. In this case, the resulting model does no longer include temperature histogram as a main effect. Compared to the findings from literature, this result is quite surprising. However, as shown in Section 4.1.2, the major part of observed temperatures lies in the range of 10 to 20$^{\circ}C$ (for example due to active cooling), such that in this dataset the influence of similar temperatures might be small and leads to the exclusion of the temperature histogram as a main effect. Thus, it cannot be generally concluded that the cell temperature does not influence battery degradation.

Considering other variables, merely t_{chg} and the dichotomous variable EV are considered in the shrinked interaction model. The charging duration is correlated with the mileage between charges and the energy throughput, both influencing the SOC and therefore the charging duration. Moreover, charging duration is influenced by the charging power. It is therefore strongly affected by user behavior and considered relevant for the model with the highest predictive accuracy.

Vehicles that are equipped with an additional internal combustion engine are typically operated in different SOC ranges. The internal combustion engine serves as a guarantee of range after the battery has been fully discharged. In contrast, the pure EV is recharged more frequently, as the vehicle cannot drive

any further after reaching the lower discharge bound. Hence, the variable EV can be interpreted as a representation of user behavior.

In summary it can be concluded that (i) all degradation relevant variables and their interactions need to be included in a complete, empirical degradation model (except temperature in this case), and (ii) the consideration of user behavior-related variables that influence the operating conditions of the battery, considerably improve the predictive accuracy of degradation models.

4.3 Discussion and Conclusion

In analysis a statistical approach is presented to develop an empirical battery degradation model based on field data from a large EV fleet operated under real environmental and usage conditions. As I do not base the model on physical hypothesis, the approach can be applied to any other EV fleet with different battery technology.

The empirical degradation model is based on on-board estimates of the SOH_C that serve as the ground truth. Whereas SOH_C on-board estimation is still a field of extensive research, such estimates exhibit some error. Thus, the accuracy of the empirical model is limited by the accuracy of on-board estimates.

The analyses reveal that test parameter combinations presented in recent literature on empirical degradation modeling–typically conducted under laboratory conditions–notably deviate from real environmental and usage conditions, especially in terms of temperature. Furthermore, data fitting is performed typically in-sample, which may lead to overfitting. Smart charging strategies, however, require generalizable degradation models in order to optimize the charging policy also with respect to degradation. I therefore present an empirical degradation model with high predictive accuracy, valid under real environmental and usage conditions and validated using cross-validation based on unseen data subsets.

Comparing empirical degradation models presented in literature, no consistent functional relationship can be found. While authors agree on the monotonous relationship between capacity fade and time as well as charge throughput, different exponential factors are reported. Typically, the best fit is

identified by using a set of some predefined functions that have been reported in literature previously. To overcome this ambiguity, I present a systematic approach to identify the functional relationship using Box-Cox-Transformation. I find that degradation is following a square root function of both time and charge throughput.

In recent literature degradation models with subsets of degradation relevant variables are presented. In this work, all relevant variables and usage related variables are included into the presented models. Furthermore, variable shrinkage is applied in terms of the Lasso in order to select variables with high predictive power. As I have demonstrated, this approach allows to transform the model that includes interaction terms from a highly overfitted model into a more compact model with minimum test error. Furthermore, the analysis reveals that the shrinked interaction model outperforms the literature-based model that includes variables relevant for calendaric and cyclic aging. Therefore, degradation models currently presented in literature do not consider all degradation relevant variables and interactions. Whereas in recent literature, subsets of relevant variables are included in empirical degradation models, the results show that all relevant variables as well as their interactions need to be included. For the dataset considered in this article, temperature mostly lies in a similar range (for example due to active cooling), such that the main effect of temperature related variables are no longer part of the best, shrinked interactions model. Furthermore, the consideration of user behavior in terms of the histogram of charging duration and the binary variable EV which indicates whether or not the vehicle is a pure EV, lead to a model with minimum test error and therefore maximum predictive accuracy.

Based on a detailed degradation model, smart charging strategies can be developed and allow for maximum battery life by degradation optimal charging decisions.

Part III

Prescriptive Modeling

Chapter 5

Degradation Optimal Charging

Whereas users typically maximize the available range by immediate and full recharging, battery degradation is promoted by such behavior. Therefore, in the following I develop a smart charging strategy aimed at minimizing battery degradation given that users' mobility requirements are met.

As described in detail in Part I, it is known that the major drivers for battery aging are operating conditions related to time, energy throughput, SOC and temperature (Jossen, 2006). In order to propose a strategy for battery life optimal behavior, the consideration of all these factors on battery degradation is crucial. In the following, a detailed battery degradation model is highlighted in order to serve as the objective function for an optimization model. Following that, degradation optimal charging is compared to naive as fast as possible charging as well as the contrary strategy of as late as possible charging. Considering that deterministic knowledge of the upcoming driving profile is accessible accurately, a more convenient charging heuristic is investigated and the trade-off between flexibility and battery degradation is solved.

5.1 Optimization Model

In Section 2.3.1 an overview of usage related degradation drivers is given and degradation models in literature are presented. However, in literature no model exists that includes all degradation relevant factors. In order to present a reproducible evaluation the objective function, i.e. a degradation model is chosen that

is available in literature.

This work is focused on the degradation induced by charging decisions in terms of timing and amount of energy charged. Thus, all degradation factors that have been identified so far are relevant. Hardly any model presented in Table 2.1 considers C-rate, except for Wang et al. (2014). However, the model including most relevant factors has been presented by Schmalstieg et al. (2014). It is furthermore based on the testing of 37 parameter variations compared to 15 and 17 variations that are considered in the models by Marongiu et al. (2015) and Sarasketa-Zabala et al. (2016), respectively. Therefore, this model provides the most reliable fit. For these reasons, in the following battery degradation is considered according to the model introduced by Schmalstieg et al. (2014). The main characteristics of this model are described briefly in the following.

Schmalstieg et al. (2014) derive an aging model from the results of accelerated aging tests of commercial Sanyo UR18650E Li-NMC 18650 cells with a capacity of 2.15 Ah. In this degradation model, instead of SOC the electric potential v is considered. The relationship between SOC and v is determined by the OCV curve. The OCV curve of the cell considered in Schmalstieg et al. (2014) can be found in Ecker et al. (2014). In this article the relationship is assumed to be linear, for the sake of simplicity. The SOC corresponds to the electric potential v measured in *Volt* $[V]$. The upper and lower bound on cell level is derived from manufacturer cell specifications as $v_{ub} = 4.1V$ and $v_{lb} = 3.32V$.

The model estimates the current SOH_c from a calendaric and a cyclic aging term as shown in (2.8), which is constructed according to (5.1). All cyclic aging tests have been conducted at $35°C$.

$$C = 1 - \alpha_{cal}(T, v) \cdot t^{0.75} - \beta_{cyc}(\overline{v}, DOD) \cdot \sqrt{Q} \tag{5.1}$$

The calendaric term α_{cal} is shown in (5.2). The term results from a data fitting procedure and includes an Arrhenius dependency for temperature T in Kelvin and is linearly depending on v. Time t has been found to impact capacity in a monotonically decreasing fashion with a function proportional to $t^{0.75}$ (with t in

days).

$$\alpha_{cal}(T,v) = (7.543 \cdot v - 23.75) \cdot 10^6 e^{-\frac{6976K}{T}} \tag{5.2}$$

The cyclic aging term β_{cyc} is shown in (5.3). It is composed of a quadratic component considering $\overline{SOC}$ (as the average cycling voltage $\overline{v}$), a constant term, and a third component assuming a linear impact of DOD. Obviously, the cyclic term is proportional to the square root of charge throughput (Q [Ah]), corresponding to a monotonically increasing amount of degradation.

$$\beta_{cyc}(\overline{v},DOD) = 7.348 \cdot 10^{-3}(\overline{v} - 3.667)^2 + \quad 7.6 \cdot 10^{-4} + 4.081 \cdot 10^{-3} DOD \tag{5.3}$$

To account for the unconsidered C-rate, in the following slow charging with low charging currents are assumed.

Although currently available EVs easily cover average daily range needs, recent studies about EV user behavior show that range anxiety and habituality lead to frequent recharging and the retention of mostly unnecessary range buffers (Franke and Krems, 2013a). Practically, this behavior leads to a maximization of available range, i.e. maximization of SOC. In contrast, the calendaric term suggests a minimization of SOC to optimize battery life.

As Figure 5.1 shows, the calendaric term is especially important in high temperature cases. The figure depicts that the ratio of calendaric and cyclic aging develops non-linearly with t (left-hand side graph) and equivalent full cycles[1] Q (right-hand side graph). This requires successively adopted charging policies. The figure shows the ratio for different sample points of operation over a time horizon of ten years, with operating points $\overline{SOC} \in \{25\%, 50\%, 75\%\}$ and temperature $T \in \{20°C, 35°C\}$. Comparing the ratio at 20 and $35°C$ results in substantially different ratios and ratio developments.

In all $35°C$ cases, the calendaric term quickly gains importance and even surpasses the cyclic term (ratio > 1). The calendaric aging, however, is minimal with low voltage levels (SOC) approaching zero. This suggests to charge as late

[1] One equivalent full cycle corresponds to a charge throughput twice the battery capacity.

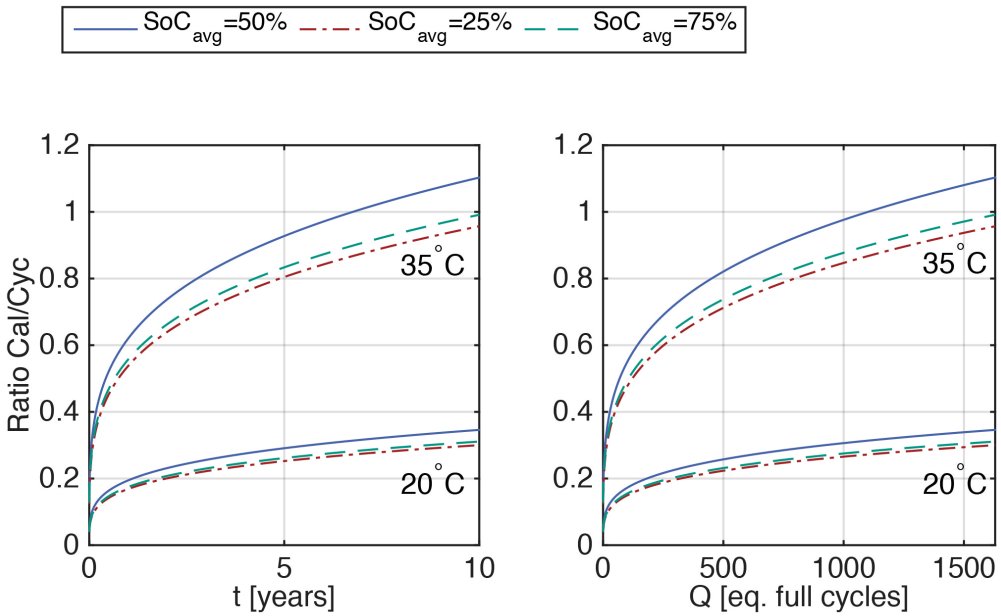

Figure 5.1: Ratio of calendaric over cyclic term for different $\overline{SOC}$ and T.

as possible and only as much as required to meet the next mobility requirement, i.e., minimizing SOC. This strategy obviously conflicts with current user preferences, where SOC is typically maximized.

For a lower temperature of $20°C$, the ratio persistently lies below 0.4, indicating the greater influence of the cyclic over the calendaric term. Cyclic aging is minimal when recharging happens around the degradation minimizing $\overline{SOC}(\overline{v})$ of 50% (3.66 V corresponds to an $\overline{SOC}$ of approximately 50% when a linear relationship between SOC and the cell voltage $\in [3.32, 4.1]V$ is assumed) in combination with a small DOD.

As a consequence, I expect charging recommendations according to the strategies I present in the following to offer considerable potential to extend battery lifetime.

I define a charging strategy as a set of rules indicating when and how much to charge. Two simple benchmarking strategies, that I consider in this simulation-based analysis, are As-Fast-As-Possible Charging (AFAP) charging and As-Late-As-Possible Charging (ALAP) charging. While AFAP reflects the typical charging behavior found today, ALAP is a hypothetical strategy assuming full infor-

Indices	Domain	Description
t	$\mathbb{N}$	Time slot (ts)
Parameters		
$\mathcal{T}$		Time horizon (with $t = 1, ..., \mathcal{T}$)
L_t	$\{0,1\}$	Location vector ($L_t = 1$ charging location)
ξ_t	$\{0,1\}$	Indicator function ($\xi_t = 1$ driving)
Q_t	$\mathbb{R}^+$	Charge throughput in Ah
V^+	$\mathbb{R}^+$	Voltage increase per ts (charge)
V_t^-	$\mathbb{R}^+$	Voltage decrease per ts (discharge)
Variables		
v_t	$[v_{lb}, v_{ub}]$	State of charge at t
$\overline{v}_t$	$[v_{lb}, v_{ub}]$	$\overline{SOC}$ (within cycle)
v_t^{min}	$[v_{lb}, v_{ub}]$	Minimum SOC (within cycle)
v_t^{max}	$[v_{lb}, v_{ub}]$	Maximum SOC (within cycle)
DOD_t	$[0, v_{ub} - v_{lb}]$	Depth of discharge (within cycle)
Φ_t	$[0,1]$	Decision Variable: Charging amount in t

Table 5.1: Nomenclature

mation on future trips (Flath et al., 2014). Both strategies will serve as benchmark strategies for Degradation Optimal Charging (OPT), the strategy to estimate the maximum battery life proposed in this work. OPT is formulated as a quadratic continuous optimization model that serves to calculate the optimal charging decisions.

The nomenclature used throughout this work is summarized in Table 5.1. Time slots are indicated by the subscript t and the considered time horizon is represented by $\mathcal{T}$. The availability of a charging location is indicated by the binary parameter L and driving is depicted by the indicator variable ξ. Furthermore, Q corresponds to the monotonically increasing charging throughput in Ampere-hours. The parameters V^+ and V^- depict the increase and decrease of voltage per time slot, respectively. The variables v, $\overline{v}$, v^{min} and v^{max} are given in Volts and represent the SOC, the average SOC as well as the minimum and maximum SOC within a cycle, respectively. DOD corresponds to the delta between v^{max} and v^{min} and finally the decision variable Φ, that is determined by the optimization, indicates the charging amount per time slot.

I will now formally introduce the three charging strategies considered in this

work, namely OPT, AFAP and ALAP.

5.1.1 Optimal Charging

The objective of OPT is to minimize battery degradation. In order to calculate battery degradation the calendaric and the cyclic aging component are summed in each time slot throughout an optimization horizon $\mathcal{T}$ as shown in Equation (5.4). Both, the cyclic and calendaric component directly depend on charging decisions.

$$\min_{\Phi} \quad \sum_{t=1}^{\mathcal{T}-1} \left(\underbrace{\alpha_{cal} \cdot (t^{0.75} - (t-1)^{0.75})}_{\text{calendaric component}} + \underbrace{\beta_{cyc} \cdot (\sqrt{Q_t} - \sqrt{Q_{t-1}})}_{\text{cyclic component}} \right) \tag{5.4}$$

OPT minimizes the sum of calendaric and the cyclic battery degradation per time slot. Therefore, the goal is to minimize the SOC (v) and the DOD in combination with a $\overline{SOC}$ ($\overline{v}$) around 50% (3.66 V). However, the ratio between the calendaric and cyclic term is time dynamic and also depends on the accumulated charge throughput, demanding for adoption of OPT over time. I compare OPT with AFAP and ALAP charging that represent two polar strategies that either only account for uncertainty in the mobility requirements or only address the calendaric aging factor, respectively.

5.1.2 As Fast as Possible Charging

With AFAP the vehicle is charged immediately whenever possible, for instance, directly after arriving at home. This strategy has frequently been mentioned in the literature, for instance inSchuller (2015); Flath et al. (2014) and is the most convenient for the user. The term 'naive charging' is often used synonymously for AFAP charging. It requires the least organizational effort and minimizes the risk of not being able to realize trips or having a breakdown in case of spontaneous or unplanned trips. Applying AFAP, time slots in which the battery is charged are calculated by maximizing the vehicle's SOC as shown in (5.5).

$$\max_{\Phi} \quad \sum_{t \in \mathcal{T}} v_t \tag{5.5}$$

This strategy is well in line with the empirically observed driver behaviour with respect to the range anxiety phenomenon (Franke and Krems, 2013a) as the potential driving distance is maximized.

5.1.3 As Late as Possible Charging

Compared to AFAP the ALAP charging strategy requires substantially more organizational effort. The amount of energy charged corresponds to the required energy such that the next trip is feasible. Therefore, charging occurs only if necessary–as late as possible–which may lead to a delayed start of charging and the battery typically is not fully charged. ALAP aims at charging only the amount of energy required between two subsequent charging opportunities. Therefore, the SOC in terms of v_t is minimized as shown in (5.6).

$$\min_{\Phi} \quad \sum_{t \in \mathcal{T}} v_t \tag{5.6}$$

For end users it might by inconvenient to apply this strategy as precise information on the next trip(s) would be mandatory. However, both AFAP and ALAP strategies correspond to extreme behavior. Therefore, they serve as benchmark scenarios for OPT. To account for the user inconvenience arising from ALAP charging, I consider two different approaches to include a safety buffer of range. This range buffer corresponds to a minimum of range held available combined with ALAP charging.

I differentiate between a strict and a soft version of ALAP charging combined with a range buffer, that holds a certain constant range available. The strict approach assumes the problem to be solved according to the objective function in (5.6), combined with an adjustment of the lower voltage bound v_{lb}. While the soft approach does not modify voltage bounds, voltage levels below the range buffer threshold $\underline{v}_{rb}$ are hit only if necessary to complete the current trip. The soft approach–ALAP with uncertainty buffer $ALAP_b$–is solved according to the

objective function in (5.7). This will force v_t to be as close to $\underline{v}_{rb}$, while–within the voltage bounds–over- and undershooting is allowed.

$$\min_{\Phi} \quad \sum_{t \in \mathcal{T}} (v_t - \underline{v}_{rb})^2 \tag{5.7}$$

5.1.4 Constraints

I will now introduce a set of constraints that the charging strategies are subject to. The vector $\boldsymbol{L} = \langle L_0, ...L_{\mathcal{T}} \rangle$ defines whether an EV is located at a charging location ($L_t = 1$) or charging is not possible ($L_t = 0$) at time t. The vector directly results from the driving profile and is assumed to be known up to the optimization horizon. The charging decision is represented in the vector $\boldsymbol{\Phi} = \langle \Phi_0, ...\Phi_{\mathcal{T}} \rangle$. No charging is expressed by $\Phi_t = 0$ and, for example, $\Phi_t = 0.5$ corresponds to charging with 50% of the maximum power. Constraint (5.8) indicates that charging is only allowed for a charging location in t.

$$\Phi_t \leq L_t \tag{5.8}$$

The battery's cell voltage, given by the vector $\boldsymbol{v} = \langle v_0, ...v_{\mathcal{T}} \rangle$ is bounded by v_{lb} for a completely discharged battery and v_{ub} corresponding to a fully charged battery.

During driving, voltage (i.e., SOC) is reduced according to the driving profile, indicated by the vector $\boldsymbol{V^-} = \langle V_0^-, ...V_{\mathcal{T}}^- \rangle$. Based on the driving profile, which includes information about the distance $\boldsymbol{s} = \langle s_0, ...s_{\mathcal{T}} \rangle$ travelled per time slot, the consumption is transformed by Equation 5.9 from meters into volts. Therefore, I assume an average consumption per kilometer γ in Wh/m and a linear relation between cell voltage and nominal capacity C_{nom} in Watt hours, which corresponds to the initially available capacity before degradation.

$$V_t^- = \frac{v_{ub} - v_{lb}}{C_{nom}} \cdot s_t \cdot \gamma \tag{5.9}$$

The maximum charging energy, on the other hand, is given by $V^+ \in \mathbb{R}^+$. Consequently, the cell voltage in t, v_t, depends on the cell voltage in $t-1$, v_{t-1}, the

charging decision vector $\mathbf{\Phi}$ and driving behaviour $\boldsymbol{V_t^-}$ (5.10).

$$v_t = v_{t-1} + \Phi_t \cdot V^+ - V_t^- \tag{5.10}$$

The definition of DOD and $\overline{SOC}$ is straightforward in accelerated aging tests, where the SOC at start of discharge and end of charge is typically equal. However, under dynamic real-world driving and charging profiles, the SOC at start of discharge and end of charge may differ. Assuming that discharged energy will be recharged with certainty, I assume DOD and $\overline{SOC}$ to be defined per trip. Therefore. $\overline{SOC}$ corresponds to the average SOC within a trip. The vector $DOD = \langle DOD_0, ...DOD_{\mathcal{T}} \rangle$ is determined by the difference of the maximum cell voltage before the trip v_t^{max} and the minimum voltage v_t^{min} after the trip as shown in (5.11). In the preprocessing procedure of the driving profile, trips that spread over several time slots are summarized to one time slot to simplify calculations.

$$DOD_t = \frac{v_t^{max} - v_t^{min}}{v_{ub} - v_{lb}} \tag{5.11}$$

The vector $\boldsymbol{\overline{SOC}}$ ($\overline{v} = \langle v_0, ...v_{\mathcal{T}} \rangle$) is calculated correspondingly based on DOD as indicated by Equation (5.12).

$$\overline{v}_t = v_t^{min} + \frac{1}{2}(v_t^{max} - v_t^{min}) \tag{5.12}$$

I define a binary indicator vector $\boldsymbol{\xi} = \langle \xi_0, ...\xi_{\mathcal{T}} \rangle$, where $\xi_t = 1$ indicates the occurrence of a trip.

$$v_t^{max} = v_t \cdot (1 - \xi_t) + v_{t-1}^{max} \cdot \xi_t \tag{5.13}$$

$$v_t^{min} = v_{t-1} \cdot (1 - \xi_t) + v_t \cdot \xi_t \tag{5.14}$$

The required unit for Q is given in *Ampere-hours* [Ah], therefore the amount of discharge (V^-) given in *Volt* must be converted to Ah using a linear assignment and the proportion of nominal cell capacity C_{nom} in Ampere-hours and maximum DOD in volts ($DOD_{max} = v_{ub} - v_{lb}$) as indicated in Equation (5.15).

$$Q_t = Q_{t-1} + \frac{C_{nom}}{DOD_{max}} \cdot V_t^- \tag{5.15}$$

Finally, I introduce an additional constraint that allows for comparability of different charging approaches. I therefore assume the battery to be fully charged at the start and the end of the simulation run as shown in Equation (5.16).

$$v_1 = v_{\mathcal{T}} = v_{ub} \tag{5.16}$$

The complete resulting optimization model is summarized in Equation 5.17 and (5.18). The alternatively used objective functions are presented in Equation 5.17 to determine the battery degradation minimizing charging strategy for OPT, AFAP, ALAP and $ALAP_b$. The constraints in (5.18) apply for each of the four objective functions.

$$
\begin{array}{llll}
OPT & AFAP & ALAP & ALAP_b \\
\min\limits_{\Phi} \sum\limits_{t=1}^{\mathcal{T}-1} \alpha_{cal,t} \cdot \left(t^{\frac{3}{4}} - (t-1)^{\frac{3}{4}} \right) + \beta_{cyc,t} \cdot \left(\sqrt{Q_t} - \sqrt{Q_{t-1}} \right) & \max\limits_{\Phi} \sum\limits_{t=1}^{\mathcal{T}-1} v_t & \min\limits_{\Phi} \sum\limits_{t=1}^{\mathcal{T}-1} v_t & \min\limits_{\Phi} \sum\limits_{t=1}^{\mathcal{T}-1} (v_t - \underline{v}_{rb})^2
\end{array} \tag{5.17}
$$

$$
\begin{aligned}
s.t. \quad \Phi_t &\leq L_t && \forall t \in \mathcal{T} \\
v_t &= v_{t-1} + \Phi_t \cdot V^+ - V_t^- && \forall t \in \mathcal{T} \backslash \{1\} \\
DOD_t &= \frac{v_t^{max} - v_t^{min}}{v_{ub} - v_{lb}} && \forall t \in \mathcal{T} \\
\overline{v}_t &= v_t^{min} + \frac{1}{2}(v_t^{max} - v_t^{min}) && \forall t \in \mathcal{T} \\
v_t^{max} &= v_t \cdot (1 - \xi_t) + v_{t-1}^{max} \cdot \xi_t && \forall t \in \mathcal{T} \\
v_t^{min} &= v_{t-1} \cdot (1 - \xi_t) + v_t \cdot \xi_t && \forall t \in \mathcal{T} \\
v_1 &= v_{ub} \\
v_{\mathcal{T}} &= v_{ub} \\
v_t &\geq v_{lb} && \forall t \in \mathcal{T} \\
v_t &\leq v_{ub} && \forall t \in \mathcal{T} \\
\overline{v}_t &\geq v_{lb} && \forall t \in \mathcal{T} \\
\overline{v}_t &\leq v_{ub} && \forall t \in \mathcal{T} \\
v_t^{min} &\geq v_{lb} && \forall t \in \mathcal{T} \\
v_t^{min} &\leq v_{ub} && \forall t \in \mathcal{T} \\
v_t^{max} &\geq v_{lb} && \forall t \in \mathcal{T} \\
v_t^{max} &\leq v_{ub} && \forall t \in \mathcal{T} \\
DOD_t &\geq 0 && \forall t \in \mathcal{T} \\
DOD_t &\leq v_{ub} - v_{lb} && \forall t \in \mathcal{T} \\
\Phi_t &\geq 0 && \forall t \in \mathcal{T} \\
\Phi_t &\leq 1 && \forall t \in \mathcal{T}
\end{aligned} \tag{5.18}
$$

5.2 Simulation Design

This Section describes the simulation design employed to determine the potentials of battery degradation reduction with OPT as compared to ALAP and AFAP. Subsequently, I will analyze the trade-off between range flexibility and battery life with the charging heuristic $ALAP_b$. I build on real-world vehicle parameters and empirical mobility requirements.

OEMs typically define the EOL criterion to be fulfilled when the battery reaches 80% of the initially available capacity (Jossen and Weydanz, 2006; Spotnitz, 2003). Although operation is further possible (Saxena et al., 2015), a battery then shows unstable behavior discouraging a further usage of the battery in EVs.

In order to analyze the differences in operating time, I simulate OPT, AFAP, ALAP and $ALAP_b$ charging strategies for 1596 vehicles with individual empirical driving profiles until the 80% EOL criterion is reached. To determine time to EOL, the dynamics in calendaric and cyclic aging over longer periods need to be considered as the absolute and relative impact of both terms changes over time (see Section 5.1).

Figure 5.2 depicts the simulation process. For each vehicle, each strategy is calculated over the time horizon of one year. In case the EOL criterion has not yet been reached, the optimization is continued for the subsequent year. Otherwise, if the maximum number of 40 years is reached, the procedure continues with the next vehicle. This upper bound for the maximum number of years is set to keep the simulation effort tractable and also due to the rapid development of batteries that limits the relevance of EOL calculations beyond such a large time period. The simulation is implemented in Python 2.7.6 using the Gurobi 6.5.0 MIP solver.

5.2.1 Vehicle Parameters

For the parametrization a battery aging model of industry standard Li-NMC 18650 battery cells is assumed. A vehicle with an exemplary battery capacity of 20 kWh is considered, which is similar to available compact class EVs. For in-

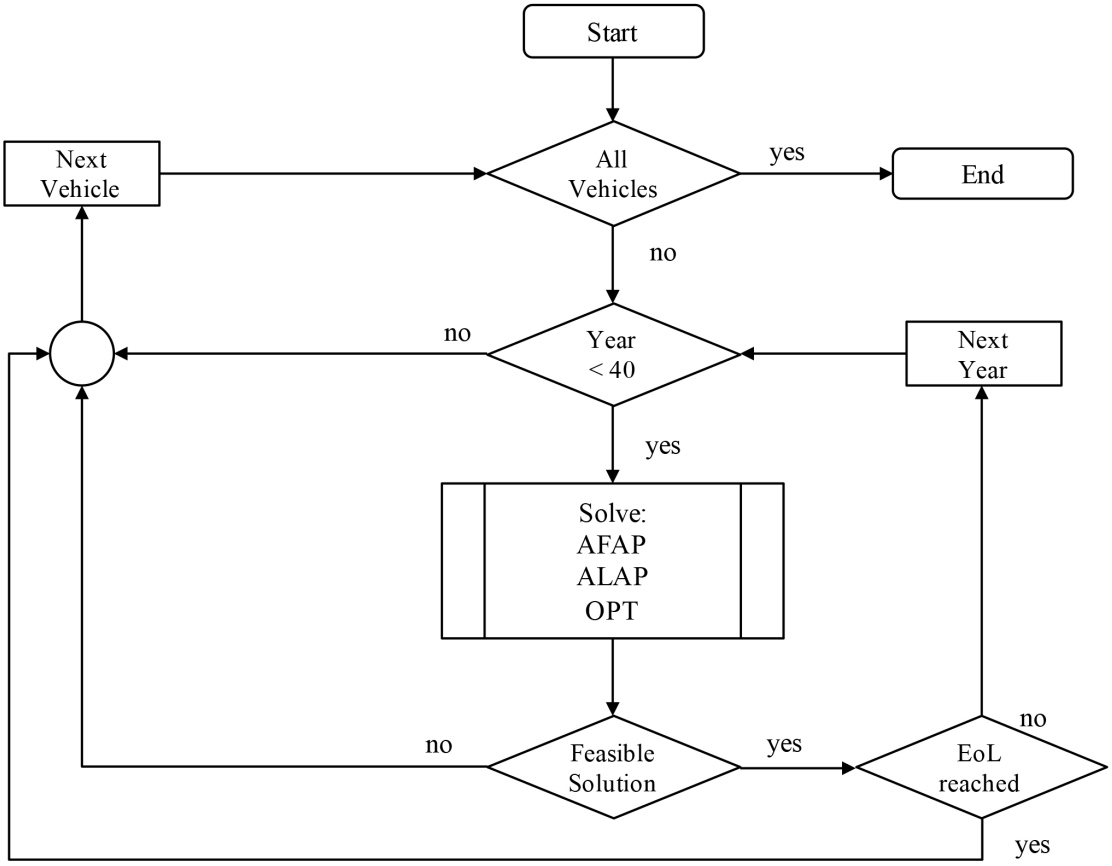

Figure 5.2: Program chart of the simulation environment.

stance, a VW e-Golf offers a capacity of 24.2 kWh.[2], the initial BMW i3 has 18.8 kWh[3] and the Mitsubishi i Miev provides 16 kWh of operational capacity[4]. The 20 kWh capacity can be obtained by combining 2500 Li-NMC cells as specified above in a battery pack[5]. This approach is consistent with manufacturing approaches: for instance Tesla Roadster battery pack is composed by 6801 of 18650 Li-ion cells (Linden and Reddy, 2011).

To consider SOC variation based on the distance passed in a driving profile, a linear mapping between SOC (or cell voltage) and the maximum available cell

[2]http://emobility.volkswagen.de/int/en/private/cars/eGolf.html
[3]www.bmw.de
[4]www.mitsubishi-motors.de
[5]Derived by dividing 20 kWh by the nominal capacity of one cell in Ah multiplied with its nominal voltage of 3.65 V

capacity is assumed. This allows to convert the charge throughput per time slot (in kWh) to V_t^- by assuming a constant consumption of $\gamma = 0.2$ Wh/m.

The cell temperature is assumed to correspond to a constant ambient temperature. The original cycling analyses of Schmalstieg et al. (2014) have been conducted under a constant temperature of $35°C$, which I will also assume in the simulations. As the average temperature in many EV relevant climatic zones, such as Europe, is much lower than $35°C$, lower temperatures of $20°C$ and $10°C$ are considered. This analysis is based on a charging power of 3.6 kW, which corresponds to standard home sockets in Germany and a C-rate well below 1C (in particular a C-rate of 0.18 is used), such that aging-effects due to high C-rates (fast charging) can be neglected.

5.2.2 Empirical Mobility Data

Mobility requirements are derived from the MOP, a continuous representative panel that is being recorded since 1994 (BMVBS, 2008). The mobility panel contains all types of trips for a participating household during the observation period of one week. From this data set I focus on the group of full-time employees, that cover the largest (cumulative and absolute) mileage in the set and extract both the distance (in kilometers) and location per time slot (a location is for example *driving*, *home*, *work*, *shopping*, etc.).

As the trips have been covered with conventional vehicles, I apply a preprocessing step to filter profiles with trips longer than the EV's battery capacity, i.e. the maximum range the EV would cover. After preprocessing I receive a set of 1717 feasible driving profiles with an average distance of 24.23 km covered per day.

As driving profiles from the MOP, with a duration of one week, are reported in discrete time steps of 15 minutes, I consider a time horizon for the optimization of $\mathcal{T} = 4 * 24 * 7 * 52 = 34944$ time slots for this simulation. This represents one year sampled in 15 minute time slots. Charging is assumed to be only possible while the vehicle is parked at home, cf. Constraint (5.8).

To allow for comparability of different charging strategies, charge throughput

is held constant by setting the cell voltage to v_{ub} the first and last time slot within the optimization horizon (Equation 5.16). This requirement cannot be met in case the last time slot of a profile contains a driving event or the number of slots that allow for charging is not sufficient to recharge the battery to v_{ub}. Hence, this analysis is conducted with the remaining set of 1596 feasible profiles fulfilling all of the requirements.

5.3 Results and Discussion

The time span to EOL of the different charging strategies is compared under different temperatures. I quantify the potential of OPT to reduce battery degradation and increase the EOL time frame compared to AFAP and ALAP. Based thereon, a discussion on the implications and limitations of OPT when applied in practice follows. OPT requires known, fully deterministic range requirements that cannot be expected to be accurately available in real life settings. Therefore, the sensitivity of time to EOL is studied on range buffer sizes introduced to deal with uncertainty and propose a battery degradation aware charging heuristic $ALAP_b$.

5.3.1 Time Span to End-of-Life

The distributions of time in years until approaching EOL (80% of the original capacity) of AFAP, ALAP and OPT charging for temperatures of 35, 20 and 10°C is depicted in Figure 5.3. The figure shows boxplots over the driving profiles considered in the simulation. As mentioned in Section 5.2, the maximum number of years considered in the simulation is limited to 40, indicated by the upper whisker for the cases of 20 and especially 10°C.

Mean and median values of the time to EOL are shown in Table 5.2. The average time to EOL is increasing by a factor of approximately four, from a mean time to EOL of 2.03 to 8.52 years, when using OPT instead of a naive AFAP charging strategy for the case of 35°C. For a temperature of 35°C, however, no significant differences between OPT and ALAP charging are found.

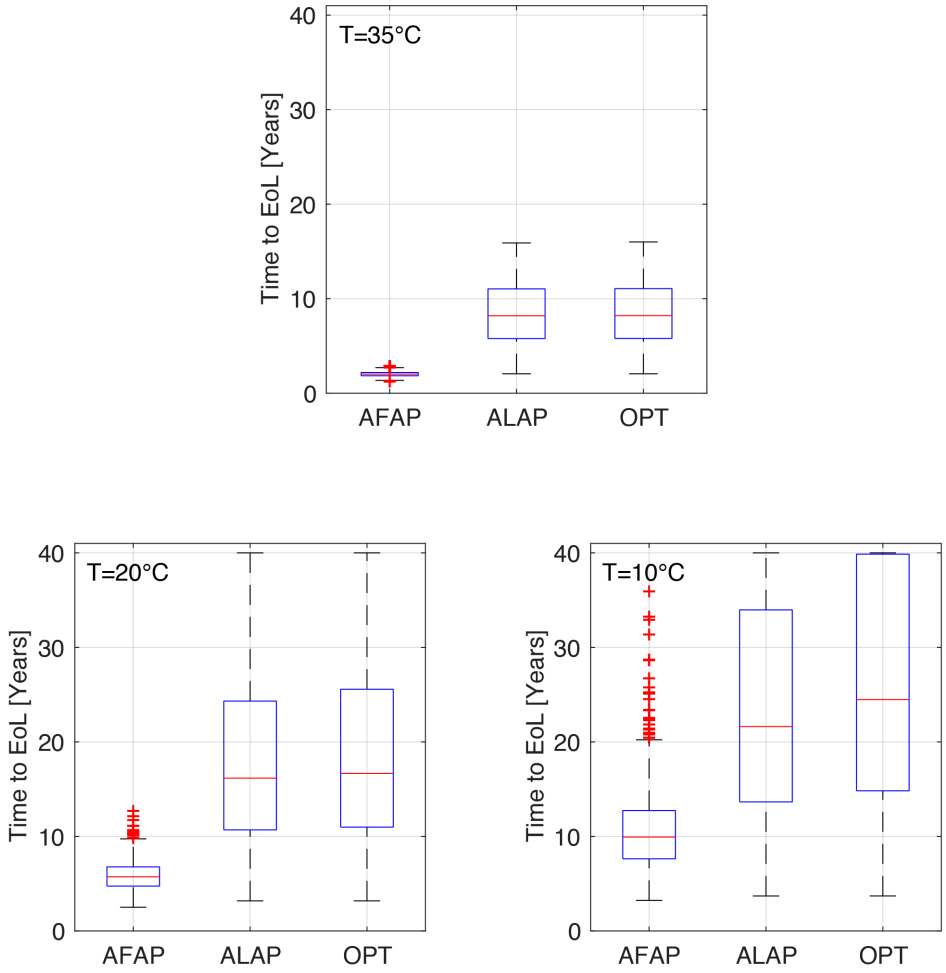

Figure 5.3: Boxplots of time to EOL in years for AFAP, ALAP and OPT charging strategy for 35, 20 and 10°C.

Assuming a temperature of 20°C, OPT charging on average increases battery lifetime by factor 3.17 compared to AFAP from 5.9 to 18.73 years. Comparing OPT to ALAP, the similarity is persistent, similar to the case of 35°C, however, the lifetime can be increased on average from 18.23 to 18.73 years by half a year. The analysis in Section 5.1, has shown the gain of importance of the cyclic term with decreasing temperature, reflected by the results for 10°C, where the average time to EOL increases from 10.85 to 23.3 and 25.06 years when applying AFAP, ALAP and OPT charging, respectively.

Significant differences can be found between AFAP and ALAP, as well as AFAP and OPT for any temperature. However, differences between OPT and

	35°C			20°C			10°C		
	AFAP	ALAP	OPT	AFAP	ALAP	OPT	AFAP	ALAP	OPT
Mean	2.03	8.50	8.52	5.90	18.23	18.73	10.85	23.3	25.06
Median	2.02	8.19	8.21	5.73	16.17	16.67	9.93	21.62	24.48
p-values									
AFAP	1	<0.001	<0.001	1	<0.001	<0.001	1	<0.001	<0.001
ALAP	<0.001	1	0.89	<0.001	1	0.32	<0.001	1	0.004
OPT	<0.001	0.89	1	<0.001	0.32	1	<0.001	0.004	1

Table 5.2: Time to EOL in years for AFAP, ALAP and OPT charging strategy for 35, 20 and 10 °C, considering *home* charging. P-values are given according to Wilcoxon rank sum test.

ALAP are only significant for a temperature of 10°C (p=0.004), as depicted in Table 5.2.

In particular under a high temperature of 35 °C OPT is very similar to ALAP. Therefore, the ratio between the calendaric and cyclic degradation term resulting from the simulation in Figure 5.4 is analyzed. Similar to Figure 5.1, valid in selected, exemplary operating points, Figure 5.4 depicts the ratio of the calendaric and cyclic aging component of the OPT charging strategy as an average of all driving profiles. The plot is differentiated by temperature throughout the first year and the corresponding charge throughput in equivalent full cycles. Here, the importance gain of the calendaric compared to the cyclic component over time can clearly be seen, especially with higher temperatures.

In the first year of operation, the ratio is persistently below one, indicating the higher weighted cyclic term. This finding is in conflict with the results presented previously, which indicate the predominance of the calendaric term, in particular under high temperatures, due to the similarity between OPT and ALAP. However, the contradiction can be explained by considering the time of operation. By definition, calendaric aging arises at any time, while cyclic aging merely occurs while driving (and charging). In the data set considered for the simulation, driving represents on average approximately 6% of the time of the day. Therefore, I conclude that the cyclic term outweighs the calendaric term, but due to the in-

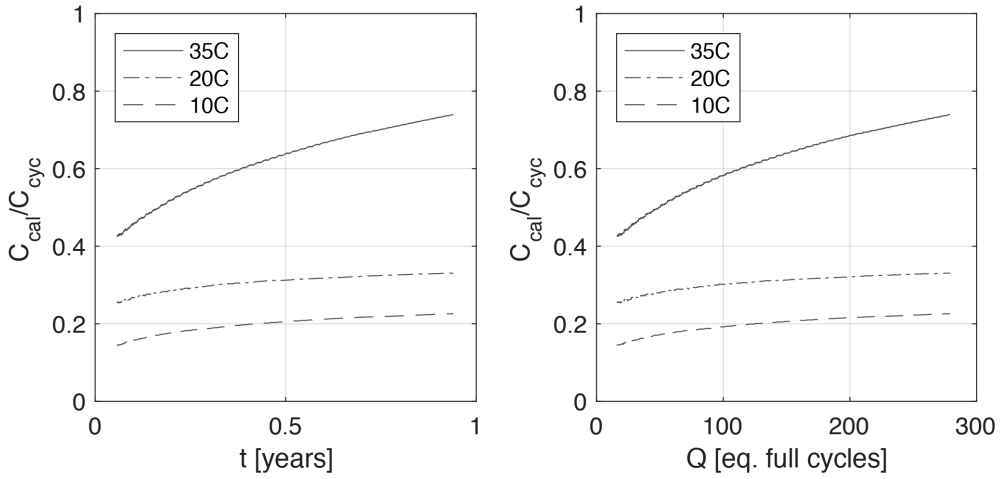

Figure 5.4: Average ratio between the calendaric and cyclic aging component for the first year of all driving profiles for OPT charging.

frequent occurrence of driving, OPT charging closely corresponds to ALAP, in particular under high temperatures.

With decreasing temperature the relative enhancement of battery life is decreasing from an approximate factor of four to a factor of three and two between OPT and AFAP charging. I thus conclude that AFAP charging is especially harmful under high temperature settings. Even under low temperature (10°C) conditions AFAP charging reduces battery life by 50% as compared to OPT.

Using OPT instead of ALAP charging leads to a lifetime extension that slightly increases with decreasing temperature. Under high temperatures OPT closely corresponds to ALAP, and ALAP leads to a nearly identical battery lifetime. Therefore, ALAP can be interpreted as a charging heuristic for OPT charging assuming precise forecasts of the upcoming range requirements. However, ALAP corresponds to the strategy that restricts flexibility most and does not allow for unplanned trips or inaccurate range predictions and is therefore inconvenient for the user in real-world applications. This trade-off is evaluated by introducing $ALAP_b$ charging strategy under consideration of additional range buffers.

Analysis of OPT Charging Strategy

OPT charging strategy nearly corresponds to ALAP. This observation is depicted for one week in an exemplary driving profile in Figure 5.5. Obviously, ALAP and AFAP are constant throughout time, OPT however changes with both time and cycles as shown in Section 5.1 and Figure 5.1. The average SOC of OPT rapidly decreases from the first year to year 5, due to the increasing importance of the calendaric term. This exemplary profile reaches EOL after 12.4 years using OPT and a temperature of 20°C. As indicated by Figure 5.5 OPT nearly corresponds to ALAP, especially with increasing age. At any age, OPT closely corresponds to ALAP, due to the minimization of the calendaric term. Consequently, recharging occurs just before the next trip, similarly to ALAP. Applying AFAP, charging starts at the beginning of every possible charging interval independent of the next trip. This is illustrated in Figure 5.5.

The correlation coefficient for an exemplary driving profile between OPT and ALAP as well as OPT and AFAP is depicted in Figure 5.7 for the corresponding SOC profile per cumulated distance per year. Considering OPT and AFAP, it becomes evident that both charging strategies lead to a highly correlated SOC profile. Throughout the first year the correlation is already high with $r_{OPT,ALAP} \approx 0.94$. Thus, with increasing time and distance travelled, OPT and ALAP are becoming nearly identical with a correlation coefficient $r_{OPT,ALAP} > 0.99$. Comparing OPT and AFAP, it can be concluded that hardly any correlation is prevailed, with a correlation coefficient converging $r \approx -0.25$. Considering all driving profiles and all years, the average correlation coefficient between OPT and AFAP is $\bar{r}_{OPT,AFAP} = -0.28$ for 20°C. Thus, I conclude that for all driving profiles hardly any correlation exists between OPT and AFAP charging. However, for OPT and ALAP, the average correlation coefficient is $\bar{r}_{OPT,ALAP} = 0.96$, corresponding to a high correlation between both charging strategies throughout the lifetime.

The correlation coefficient between OPT and ALAP indicates that OPT converges to ALAP, i.e. the correlation coefficient converges to 1. In order to analyze the correlation for all driving profiles, the functional relationship depicted in Equation 5.19 is fit to the data.

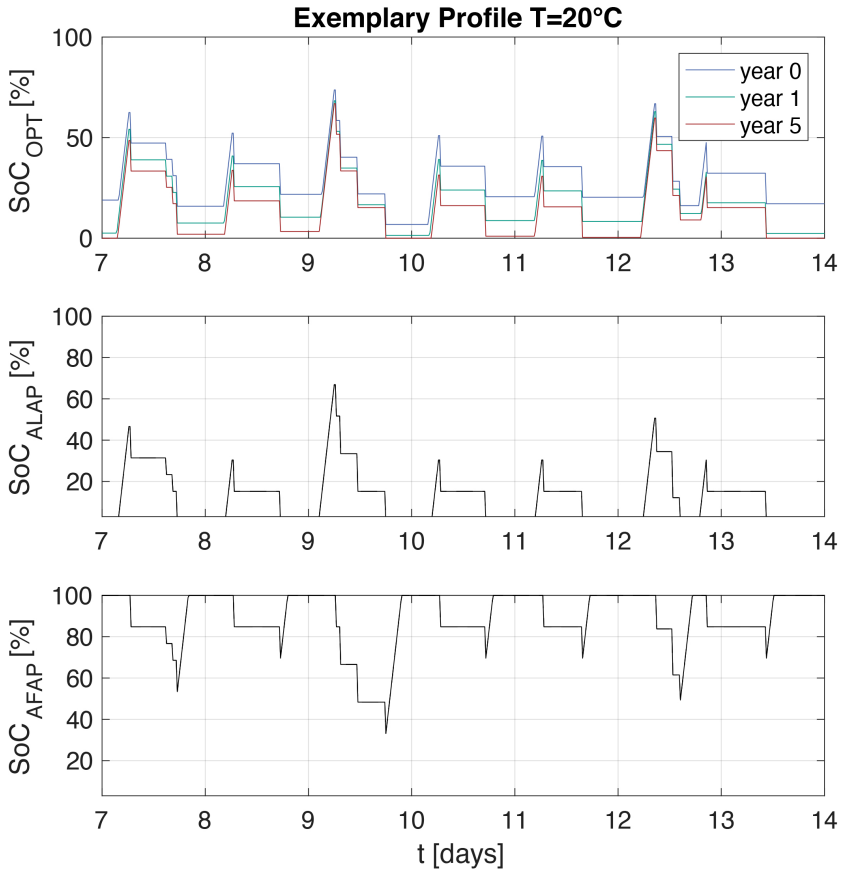

Figure 5.5: Exemplary SOC profile for the first two years and year 5 of OPT, ALAP and AFAP for 20°C.

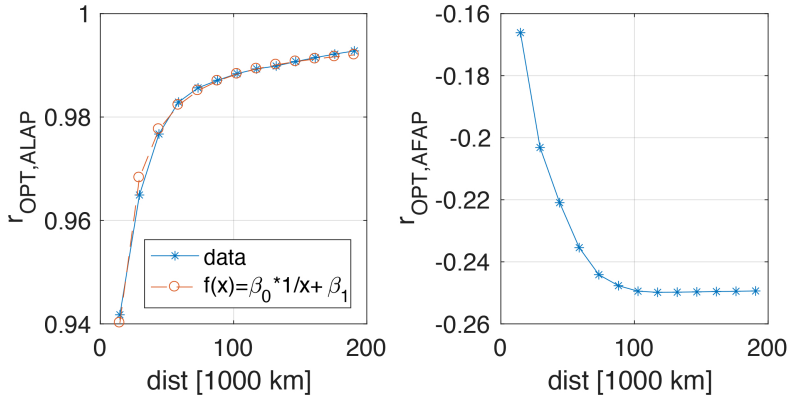

Figure 5.6: Correlation coefficient between OPT and ALAP as well as OPT and ALAP per year of an exemplary SOC profile and 20°C.

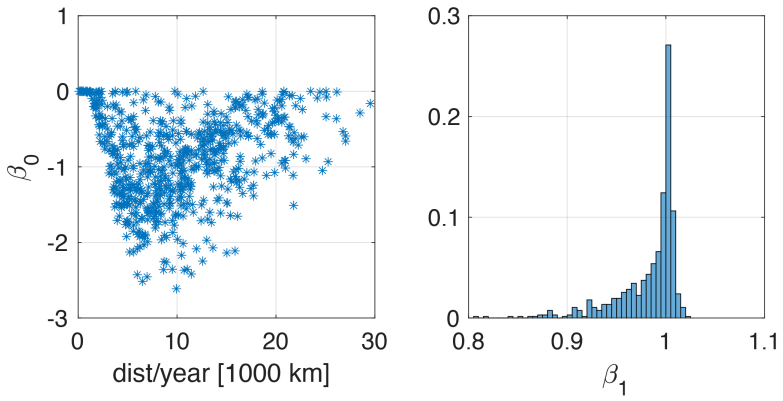

Figure 5.7: Regression coefficients for the correlation coefficient between OPT and ALAP *SOC* profiles.

$$f(x) = \beta_0 \cdot 1/x + \beta_1 \tag{5.19}$$

The resulting coefficients are compared depending on the type of travel, characterized by the distance travelled. Figure 5.7 on the right depicts a histogram of β_1 for each driving profile. β_1 represents the intercept and it is frequently close to 1, which represents perfect correlation between OPT and ALAP SOC profiles. Therefore, I conclude that for each considered driving profile OPT converges to ALAP. The coefficient β_0 is plotted over the type of travel, i.e. the distance per year, on the left in Figure 5.7. It is valid that the closer β_0 is to zero, the faster OPT converges to ALAP and vice versa. Performing another linear regression, reveals a negative slope, indicating that higher distances travelled per year lead to faster convergence of OPT towards ALAP. However, these results are not significant ($R^2 = 0.02$), such that no clear relationship can be identified.

5.3.2 Degradation Aware Charging Heuristics

A range buffer corresponds to a certain amount of range that is constantly held available. Franke and Krems (2013b,c) find typical range buffers between 10 to approximately 25 km for the user comfortable range. Thus, users prefer to retain at least 10 km of range e.g. for unexpected or emergency trips. In this setting, this finding is translated to capacity buffers between 5 and 60% of capacity, covering

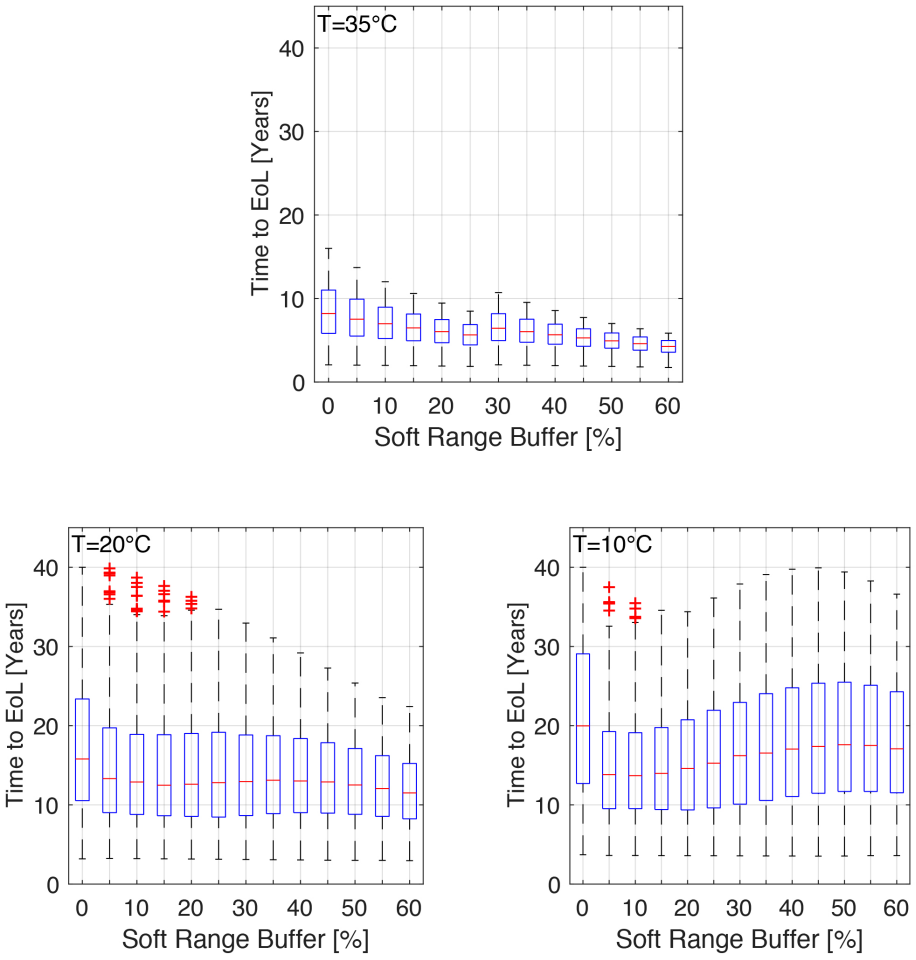

Figure 5.8: Time to EOL for ALAP charging in combination with different range buffers and different temperatures.

the preferred range buffers observed by (Franke and Krems, 2013b,c) as well as the optimal $\overline{SOC}$ around 50%. Increasing range buffer size allows for increased flexibility in terms of unplanned trips as well as range prediction errors.

The consideration of a range buffer trades-off flexibility and battery life. While a range buffer of 0% entails the least amount of flexibility but maximum battery life (representing ALAP), a buffer of 60% increases range flexibility but reduces battery life.

The boxplots in Figure 5.8 show the distributions of the time to EOL over all driving profiles per flexible range buffer (from 0 (ALAP) to 60 %) for the three different temperatures considered. The next sections address these results in

detail.

Evaluation of $ALAP_b$ at 35°C

For the case of 35°C, mean and median time to EOL decrease with range buffer levels up to a buffer of 25% (Table 5.3 and Figure 5.8). Interestingly, EOL then increases with a buffer of 30% (local maximum), before it again declines strictly with buffer levels above 30%.

Whereas the optimal operating point for the cyclic degradation term for $\overline{SOC}$ is at 50%, the optimal point of the calendaric term has a SOC value of 0%. The calendaric term outweighs the cyclic term in high temperature settings (cf. Section 5.3.1). ALAP however, is not an optional charging strategy as a range buffer above 10% is intended by the user. In the case of 35°C, a range buffer of 30% leads to both, low calendaric as well as cyclic degradation. For smaller range buffers, such as 25% the sum of calendaric and cyclic degradation is increased (in terms of median time to EOL depicted in Table 5.3).

This observation is quantified in Table 5.3. On average ALAP leads to a battery life of 8.5 years. The introduction of range buffers of increasing size decreases battery life to 4.2 years for the range buffer of 60%. This result is not surprising, since ALAP is close to OPT for 35°C and the introduction of a range buffer inhibits SOC minimization as a certain amount of range is constantly held available. For a range buffer of 30%, the average battery life is increased compared to a buffer of 25% from 5.61 years to 6.54 years.

In summary, range buffers reduce time to EOL compared to ALAP for a temperature of 35°C. However, as a range buffer above 10% is assumed to be intended by the user according to (Franke and Krems, 2013b,c), the results recommend to even increase buffer-levels (and flexibility) to 30% in order to better trade-off flexibility and battery life.

Compared to naive AFAP charging the introduction of any of the presented range buffers is beneficial, as it increases the average battery life from 2.03 to 4.2 years even for the maximum range buffer considered of 60%.

	OPT	AFAP	ALAP 0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%
35°C															
Mean	8.52	2.03	8.50	7.72	7.07	6.52	6.04	5.61	6.54	6.08	5.65	5.26	4.88	4.53	4.20
Median	8.21	2.02	8.19	7.52	6.98	6.48	6.03	5.64	6.44	6.03	5.66	5.28	4.93	4.59	4.26
20°C															
Mean	18.73	5.90	18.23	15.04	14.54	14.34	14.30	14.28	14.23	14.09	13.84	13.48	13.01	12.44	11.79
Median	16.67	5.73	16.17	13.31	12.89	12.48	12.61	12.81	12.94	13.11	13.01	12.90	12.51	12.06	11.51
10°C															
Mean	25.06	10.85	23.30	14.80	14.71	14.97	15.57	16.37	17.22	17.99	18.60	18.97	19.05	18.81	18.27
Median	24.48	9.93	21.62	13.82	13.69	13.97	14.60	15.26	16.22	16.54	17.05	17.37	17.59	17.5	17.08

Table 5.3: Time to EOL for OPT, ALAP and $ALAP_b$ charging with different soft range buffers.

Evaluation of $ALAP_b$ at 20°C

For a temperature of 20°C, Figure 5.8 shows a monotonically decreasing median of time to EOL up to a range buffer of 30% (numbers are given in Table 5.3). A local peak of the median is exhibited at a range buffer of 35%, similar to the temperature of 35°C. However, the range buffer of 35%, that trades-off battery life and flexibility, is closer to 50%–the optimum of the cyclic degradation term– due to the lower temperature. The mean of time to EOL reveals a monotonically decreasing development with increasing range buffer, as depicted in Table 5.3. Also in this case the recommendation is to use the buffer of 35% and meet the comfortable range for the user and allows for a long battery life.

Evaluation of $ALAP_b$ at 10°C

A similar recommendation results from cases with temperatures around 10°C. Assuming a range buffer higher than 10%, the reduced impact of the calendaric term leads to a global maximum median time to EOL of 19.05 years with a range buffer of 50% (Table 5.3). Thus, the optimal operating point of 50% for $\overline{SOC}$ for the cyclic degradation term is met. Assuming that at least small buffers are mandatory, it is beneficial to consider large buffers at around 50% to maximize time to EOL. Hence, in this case the trade-off between flexibility and time to EOL vanishes to a large extend.

In summary, the trade-off between battery life and flexibility is obvious as none of the considered $ALAP_b$ charging strategies can approach the time to EOL achieved when ALAP is applied. Maximum battery life comes at the cost of minimum flexibility and increased range flexibility leads to a decrease in battery life. However, users aim at retaining a range buffer above 10% as indicated by Franke and Krems (2013b,c). Based on the results it is recommend to apply a range buffer of 30%, 35% and 50% for temperatures of 35, 20 and 10°C, respectively. Using a Wilcoxon rank sum test, differences between the recommended range buffer and a buffer of 10% are significant with $p < 0.001$ for temperatures of 10 and 35°C. No significant difference can be found for a temperature of 20°C.

The trade-off between range flexibility and battery life vanishes with decreasing temperature, due to the decreased weight of the calendaric term under lower temperatures. As a consequence, the OPT charging strategy deviates from SOC minimization and therefore intrinsically allows for more flexibility. Therefore, in climate zones with lower average temperatures or under active battery cooling, it is beneficial to apply the $ALAP_b$ charging heuristic with a large range buffer of 50% to increase battery life, flexibility and applicability.

5.4 Conclusions and Limitations

While recent studies on EV user behavior indicate that users prefer frequent and full recharging (AFAP), changing this charging behavior can tremendously extend battery life. Based on simulation results build on a comprehensive battery cell aging model and empirical mobility data, I show that a battery degradation minimal (optimal) charging strategy (OPT) extends battery life by a factor of two or higher. AFAP is especially harmful in cases of higher average operating temperatures. OPT is close to as-late-as-possible (ALAP) charging at high temperatures of $35°C$.

However, ALAP and OPT require full information about the next-range requirements, that cannot be expected to be available precisely in real-life settings. I therefore investigated the trade-off between flexibility and battery life by introducing flexible range buffers between 5 to 60% to ALAP, i.e. $ALAP_b$.

I find that a lower range buffer of 30% is beneficial for high temperatures (35°C). For decreasing temperatures the trade-off between battery life and flexibility is solved with increased range buffer, i.e. 35 and 50% for a temperature of 20 and 10°C. For low temperatures, which can be achieved for example using battery cooling systems, $ALAP_b$ charging with a range buffer of 50% can be applied as an easy-to-use charging heuristic and allows for both battery life extension, flexibility and therefore user convenience.

In summary, while none of the presented $ALAP_b$ strategies including range buffers perform close to OPT, the harm of range buffers reduces with decreasing temperature such that this trade-off is less pronounced in climate zones with average (operational) temperatures around 10°C or with active battery cooling systems that enable a performance in such a temperature range. However, $ALAP_b$ charging can be implemented as an easy-to-use smart charging heuristic, that leads to considerable battery life extension compared to the currently applied, naive AFAP charging.

There are several approaches to expand the presented model. First, the model objective function is based on the degradation model developed by Schmalstieg et al. (2014). As I have carried out in Section 2.3.1, the detailed comparison of different degradation models does not indicate major differences in the results, due to the similar structure of the calendaric as well as the cyclic aging components. However, comparing the results of different degradation models reveals that authors of the corresponding literature typically perform extensive testing but with different combinations of variables relevant to degradation. Therefore, this simulation is based on the most detailed model presented by Schmalstieg et al. (2014). However, generalized test procedures and combinations of tested variables in degradation tests would improve comparability and generalizability of such models.

Part IV

Finale

Chapter 6

Conclusions and Outlook

ELECTRIC vehicles are a necessary means in individual mobility to reduce greenhouse gas emissions and to preserve the earth for future generations. However, the widespread acceptance of EVs is delayed mainly by issues resulting from the high voltage battery. This includes the high price of an EV, as well as the limited range and long recharging times. Another difficulty comes with battery degradation. It is driven by both time and cycling of the battery and leads to a loss of capacity, i.e. a reduction of the available range.

In practice, these issues result in a phenomenon called range anxiety, which causes often unnecessary, frequent and full recharging by the user. Whereas the complex process of battery degradation also highly depends on usage conditions, high states of charge typically increase aging. Consequently, there is a huge potential to optimize charging decisions in order to maximize battery life as well as the cumulated range of an EV.

In this work a mathematical optimization model is developed to determine charging recommendations, i.e. when and how much to charge, to minimize battery degradation. Based on this optimal charging strategy, a charging heuristic is derived, that can be applied easily by the user. On the one hand, the understanding of battery degradation, and the recommendation of battery-life-maximizing user behavior is crucial for an OEMs guarantee design, the prediction of failure rates and predictive maintenance. On the other hand, degradation optimal behavior allows the user to take advantage of the highest possible range at full charge throughout the lifetime of the EV.

To develop a prescriptive analytics model and to derive a degradation optimal charging strategy a degradation model, that is valid under realistic environmental and usage conditions is crucial. As of now, degradation models are typically developed based on accelerated aging tests of cells performed under laboratory conditions. Therefore, recently available field-data is a source of tremendously increased information, as it results from highly dynamic, realistic environmental and usage conditions. This new approach, however, requires proper data selection, reduction and sampling techniques. These are prerequisites to meet the restrictions of limited memory of on-board vehicle ECUs and bandwidth for telematic data transmission of potentially available signals.

6.1 Contribution

In this work, a battery degradation minimizing charging strategy is developed using prescriptive analytics by introducing a continuous quadratic programing model. Based thereon, a convenient charging heuristic is derived that can be recommended to and easily applied by the user.

However, as a prerequisite to derive degradation optimal charging, thorough analyses on data representation and prediction of battery degradation are performed intially. To summarize the contributions of this thesis, individual contributions are presented in the following based on the research questions depicted in Part I.

Data Representation: By now, it is unclear how to handle the vast amount of battery degradation related data that is potentially available by EVs currently on the streets. Data storage restrictions that arise from embedded devices used on-board of vehicles as well as restrictions of the bandwidth of telematic transmission, require data reduction and selection techniques. In a simulation based analysis, considering degradation relevant variables as described in the literature, I was able to show, that a reduced set of features allows for an accurate prediction of battery degradation in EVs based on standard equipment. Comparable predictive accuracy to the full model of 40 features can be achieved by a

subset of 32 and 33 variables using Lasso and Elastic Net for variable selection.

Descriptive analytics on the simulated data indicate that the socio-demographic background of drivers and the resulting driving profiles, for example of full-time employees and retired persons considerably influence the time to EoL. I was able to show, that for full-time employees it is beneficial to prefer an EV for purchase that guarantees a certain battery lifetime instead of a distance covered. For retired persons, the contrary applies. From an OEMs point of view, this finding needs to be considered in the guarantee design of the high voltage battery.

Predictive Degradation Modeling: Battery degradation models in literature are typically based on accelerated aging tests, performed under laboratory conditions. Firstly, such tests allow for limited selection and combination of parameters and test conditions. Secondly, this work revealed that tested and real-world conditions considerably differ in terms of temperature. Thirdly, the performance of such models is typically measured in-sample, and potential overfitting cannot be excluded. In order to solve these issues, a detailed degradation model is derived from field-data of more than 5000 EVs operated under real-world usage conditions.

Literature on empirical degradation models does not agree on a functional relationship between capacity fade and time as well as charge throughput. In this work, the relationship is analyzed systematically using Box-Cox-Transformation. Hence, I found a square root relationship between both time and charge throughput and degradation, i.e. capacity fade. Different degradation models have been tested on predictive accuracy and generalizability using cross validation. In literature, degradation models are presented with subsets of relevant variables. In this work, I contribute to the literature by including all variables relevant for aging as well as usage related variables from field-data. Applying the Lasso for shrinkage allows to select variables with high predictive power and to derive a generalizable model. This approach allows to transform the highly overfitted model including interaction terms into a compact model with minimum test error. Moreover, my analysis revealed that the best, shrinked

interaction terms model outperforms literature inspired models. Thus, I conclude that in literature, not all relevant variables and interactions are considered to accurately predict battery degradation.

Battery Life Optimal Charging: Currently, EV users prefer frequent and full recharging. Whereas battery degradation is–amongst other factors–increased under high states of charge, there is a huge potential to increase battery life by recommending optimal charging behavior.

After analyzing the data representation of the potential huge amount of data available in the field and the development of a battery degradation model, prescriptive analytics were applied to derive battery life optimal charging (OPT). Compared to naive as-fast-as-possible (AFAP) charging the developed OPT charging strategy extends battery life by a factor of two or higher, depending on the temperature considered. Thus, battery life is extended approximately by factor 4, 3 and 2 for temperatures of 35, 20 and $10°C$, respectively. AFAP charging is found to be especially harmful under high temperatures. Overall, OPT is similar to ALAP charging, such that for any temperature, charging occurs right before the following trip instead of at the arrival at the charging location.

OPT charging strategy requires full information on the future driving profile without flexibility for unexpected or emergency trips or inaccurate range predictions. Therefore, a charging heuristic was developed based on ALAP charging that considers the trade-off between flexibility and battery life by introducing range buffers between 5 and 60%.

The trade-off between battery life and flexibility is solved for buffer sizes of 30, 35 and 50% for temperatures of 35, 20 and 10°C, respectively. The results indicate, that if a range buffer is intended, higher range buffers are beneficial. In summary, if an active battery cooling system is available, it is recommended to apply $ALAP_b$ charging combined with a range buffer of 50%. This increases both battery life and flexibility and allows for high user convenience.

6.2 Future Work

In this work, the development of a degradation minimizing charging strategy and a charging heuristic with high applicability, has been shown to dramatically increase battery life by accounting for range flexibility. Based on these insights, several directions for future work can be identified.

Firstly, optimization of charging decisions in the area of V2G need to be evaluated with respect to efficiency and monetary benefit, when battery degradation is considered.

Secondly, when thinking of a charging recommendation system that provides online information about the degradation optimal behavior, reliable predictions of the next trips are required. To install such a recommendation system, a simplified charging strategy, which I have presented in terms of $ALAP_b$ charging, combined with statistical assistance systems that help the user to estimate the next range requirements can be applied.

Thirdly, proper incentives and design of information systems need to be found to convince the user to deviate from AFAP charging. Such incentives might be monetary or environmental, comparably to the design of eco-driving assistance systems (Dogan et al., 2014). On the one hand, the decision to deviate from ALAP charging depends on the knowledge of upcoming trips and abilities to charge. On the other hand, this decision might be influenced by cognitive biases that foster range anxiety. Thus, range preferences are typically higher than range needs (Franke and Krems, 2013b). This research question is open for future work.

References

Aggarwal, C. C. (2013). *Managing and Mining Sensor Data*. Springer Science & Business Media.

Alhonsuo, M., L. Virtanen, J. Rantakari, A. Colley, T. Koivum, et al. (2016). My-data approach for personal health–a service design case for young athletes. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pp. 3493–3502. IEEE.

Almuhtady, A., S. Lee, E. Romeijn, M. Wynblatt, and J. Ni (2014). A degradation-informed battery-swapping policy for fleets of electric or hybrid-electric vehicles. *Transportation Science 48*(4), 609–618.

Barré, A., B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, and D. Riu (2013). A review on lithium-ion battery aging mechanisms and estimations for automotive applications. *Jounal of Power Sources 241*, 680–689.

Barré, A., F. Suard, M. Gérard, and D. Riu (2014). A real-time data-driven method for battery health prognostics in electric vehicle use. In *Proceedings of the 2nd European Conference of the Prognostics and Health Management Society*.

Bashash, S., S. J. Moura, J. C. Forman, and H. K. Fathy (2011). Plug-In Hybrid Electric Vehicle Charge Pattern Optimization for Energy Cost and Battery Longevity. *Journal of Power Sources 196*(1), 541 – 549.

BMU (2012). Klima und energie - energiewende.

BMVBS (2008). German mobility panel (deutsches mobilitätspanel), panelauswertung 2007. *Deutsches Bundesministerium für Verkehr, Bau und Stadtentwicklung* ([Online]. Available: http://mobili- taetspanel.ifv.uni-karlsruhe.de).

Corchero, C., S. González-Villafranca, and M. Sanmartí (2014). European electric vehicle fleet: driving and charging data analysis. In *Electric Vehicle Conference (IEVC), 2014 IEEE International*, pp. 1–6. IEEE.

Department of Energy (2017). Costs of lithium-ion batteries for vehicles.

Deutsches Institut fuer Normung, D. (2015). Din en 13306:2015-09 maintenance - maintenance terminology. Technical report.

Devie, A., E. Vinot, S. Pelissier, and P. Venet (2012). Real-world battery duty profile of a neighbourhood electric vehicle. *Transportation Research Part C: Emerging Technologies 25*, 122–133.

Dietze, M. (2015). *Entwicklung optimierter Betriebs- und Ladestrategien für Fahrzeuge mit Vehicle-2-Grid-Funktionalität*. Ph. D. thesis, Karlsruhe Institute of Technology (KIT).

Dogan, E., J. W. Bolderdijk, and L. Steg (2014). Making small numbers count: environmental and financial feedback in promoting eco-driving behaviours. *Journal of Consumer Policy 37*(3), 413–422.

Eberle, U. and R. von Helmolt (2010). Sustainable transportation based on electric vehicle concepts: a brief overview. *Energy & Environmental Science 3*(6), 689–699.

ECF (2010). Roadmap 2050 - a practical guide to a prosperous, low-carbon future - technical analysis executive summary.

Ecker, M., J. B. Gerschler, J. Vogel, S. Käbitz, F. Hust, P. Dechent, and D. U. Sauer (2012). Development of a lifetime prediction model for lithium-ion batteries based on extended accelerated aging test data. *Journal of Power Sources 215*, 248–257.

Ecker, M., N. Nieto, S. Käbitz, J. Schmalstieg, H. Blanke, A. Warnecke, and D. U. Sauer (2014). Calendar and cycle life study of li (nimnco) o 2-based 18650 lithium-ion batteries. *Journal of Power Sources 248*, 839–851.

EU (2007). Richtlinie 2007/46/eg.

Flath, C., J. Ilg, S. Gottwalt, H. Schmeck, and C. Weinhardt (2014). Improving electric vehicle charging coordination through area pricing. *Transportation Science 48*(4), 619–634.

Franke, T. and J. Krems (2013a). Understanding charging behaviour of electric vehicle users. *Transportation Research Part F: Traffic Psychology and Behavior 21*, 75–89.

Franke, T. and J. Krems (2013b). What drives range preferences in electric vehicle users? *Transport Policy 40*, 56–62.

Franke, T. and J. F. Krems (2013c). Interacting with limited mobility resources: Psychological range levels in electric vehicle use. *Transportation Research Part A: Policy and Practice 48*, 109–122.

Franke, T., I. Neumann, F. Bühler, P. Cocron, and J. F. Krems (2012). Experiencing range in an electric vehicle: Understanding psychological barriers. *Applied Psychology 61*(3), 368–391.

Inc., U. T. (2013). Uber gps analysis.

Jabeen, F., D. Olaru, B. Smith, T. Braunl, and S. Speidel (2013). Electric vehicle battery charging behaviour: findings from a driver survey. In *Australasian Transport Research Forum (ATRF), 36th, 2013, Brisbane, Queensland, Australia*.

Jossen, A. (2006). Fundamentals of battery dynamics. *Journal of Power Sources 154*(2), 530–538.

Jossen, A. and W. Weydanz (2006). *Moderne Akkumulatoren richtig einsetzten*. Reichardt Verlag.

Kaebitz, S., J. B. Gerschler, M. Ecker, Y. Yurdagel, B. Emmermacher, D. Andre, T. Mitsch, and D. U. Sauer (2013). Cycle and calendar life study of graphite linmncoo li-ion high energy system. part a: Full cell characterization. *Journal of Power Sources*.

Kempton, W. and J. Tomic (2005). Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *Journal of power sources 144*(1), 268–279.

Lam, L., P. Bauer, and E. Kelder (2011). A practical circuit-based model for li-ion battery cells in electric vehicle applications. In *2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC)*, pp. 1–9. IEEE.

Larminie, J. and J. Lowry (2004). *Electric vehicle technology explained*. John Wiley & Sons.

Linden, D. and T. B. Reddy (2011). Handbook of batteries.

Liu, D., J. Pang, J. Zhou, Y. Peng, and M. Pecht (2013). Prognostics for state of health estimation of lithium-ion batteries based on combination gaussian process functional regression. *Microelectronics Reliability 53*(6), 832–839.

Marongiu, A., M. Roscher, and D. U. Sauer (2015). Influence of the vehicle-to-grid strategy on the aging behavior of lithium battery electric vehicles. *Applied Energy 137*, 899–912.

Neubauer, J., A. Brooker, and E. Wood (2012). Sensitivity of battery electric vehicle economics to drive patterns, vehicle range and charge strategies. *Journal of Power Sources 209*, 269–277.

Neubauer, J. and E. Wood (2014). The impact of range anxiety and home, workplace, and public charging infrastructure on simulated battery electric vehicle lifetime untility. *Journal of power sources 257*, 12–20.

NPE (2011). Zweiter Bericht der Nationalen Plattform Elektromobilität. Technical report, Nationale Plattform Elektromobilitaet (NPE).

Nykvist, B. and M. Nilsson (2015). Rapidly falling costs of battery packs for electric vehicles. *Nature Climate Change 5*(4), 329–332.

Pearre, N. S., W. Kempton, R. L. Guensler, and V. V. Elango (2011). Electric Vehicles: How Much Range is Required for a Day's Driving? *Transportation Research Part C: Emerging Technologies*, 1–14.

Peças Lopes, J. A., F. J. Soares, and P. M. Rocha Almeida (2009). Identifying Management Procedures to Deal with Connection of Electric Vehicles in the Grid. In *Proceedings of the 2009 IEEE Bucharest Power Tech Conference*, pp. 1–8.

Peterson, S. B., J. Apt, and J. Whitacre (2010a). Lithium-Ion Battery Cell Degradation Resulting from Realistic Vehicle and Vehicle-to-Grid Utilization. *Journal of Power Sources 195*(8), 2385–2392.

Peterson, S. B., J. Apt, and J. Whitacre (2010b). Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization. *Journal of Power Sources 195*, 2385–2392.

Prytz, R. (2014). Machine learning methods for vehicle predictive maintenance using off-board and on-board data.

Ramadass, P., B. Haran, R. White, and B. N. Popov (2003). Mathematical modeling of the capacity fade of li-ion cells. *Journal of Power Sources 123*, 230–240.

Rodgers, L., S. Zoepf, and J. Prenninger (2014). Analysing the energy consumption of the bmw activee field trial vehicles with application to distance to empty algorithms. *Transportation Research Procedia 4*, 42–54.

Rolim, C. C., G. N. Gonçalves, T. L. Farias, and Ó. Rodrigues (2012). Impacts of electric vehicle adoption on driver behavior and environmental performance. *Procedia-Social and Behavioral Sciences 54*, 706–715.

Sarasketa-Zabala, E., E. Martinez-Laserna, M. Berecibar, I. Gandiaga, L. Rodriguez-Martinez, and I. Villarreal (2016). Realistic lifetime prediction approach for li-ion batteries. *Applied Energy 162*, 839–852.

Saxena, S., C. Le Floch, J. MacDonald, and S. Moura (2015). Quantifying ev battery end-of-life through analysis of travel needs with vehicle powertrain models. *Journal of Power Sources 282*, 265–276.

Schmalstieg, J., S. Käbitz, M. Ecker, and D. U. Sauer (2014). A holistic aging model for li (nimnco) o 2 based 18650 lithium-ion batteries. *Journal of Power Sources 257*, 325–334.

Schoch, J. (2016). Modeling of battery life optimal charging strategies based on empirical mobility data. *it - Information Technology 58*(1), 22–28.

Schoch, J., J. Gaerttner, A. Schuller, and T. Setzer (2018). Enhancing electric vehicle sustainability through battery life optimal charging. Transporation Research Part B (forthcoming).

Schoch, J., E. Knobbe, B. Schweiger, and T. Setzer (2018). Field data-based predictive modeling of electric vehicle battery degradation. Working Paper.

Schoch, J., P. Staudt, and T. Setzer (2017). Smart data selection and reduction for electric vehicle service analytics. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Schuller, A. (2015). *Charging Coordination Paradigms of Electric Vehicles*. Power Systems. Springer Singapore.

Schuller, A., B. Dietz, C. Flath, and C. Weinhardt (2014, Sept). Charging Strategies for Battery Electric Vehicles: Economic Benchmark and V2G Potential. *Power Systems, IEEE Transactions on 29*(5), 2014–2022.

Smart, J. and S. Schey (2012). Battery electric vehicle driving and charging behavior observed early in the ev project. *SAE International Journal of Alternative Powertrains 1*(2012-01-0199), 27–33.

Soimar, M. and M. Kluger (2000). The challenge of cvts in current heavy-duty powertrains. *Diesel Progress North American Edition 66*(4), 68–69.

Spotnitz, R. (2003). Simulation of capacity fade in li-ion batteries. *Jounal of Power Sources*, 72–80.

Sun, X.-H., T. Yamamoto, and T. Morikawa (2015). Charge timing choice behavior of battery electric vehicle users. *Transportation Research Part D: Transport and Environment 37*, 97–107.

Sung, W. and C. B. Shin (2015). Electrochemical model of a lithium-ion battery implemented into an automotive battery management system. *Computers & Chemical Engineering 76*, 87–97.

Sweda, T. M., I. S. Dolinskaya, and D. Klabjan (2016). Optimal recharging policies for electric vehicles. *Transportation Science*.

TuTiempo.net (2014). El tiempo en madrid.

UBA (2012). Emissionsstandards - eu verordnung nr. 510/2011.

Underground, W. (2014). Weather history for kphx.

Valentine, K. F., W. G. Temple, and K. M. Zhang (2012). Electric Vehicle Charging and Wind Power Integration: Coupled or Decoupled Electricity Market Resources? In *Proceedings of the IEEE 2012 PES General Meeting*. IEEE PES.

Waldmann, T., M. Wilka, M. Kasper, M. Fleischhammer, and M. Wohlfahrt-Mehrens (2014). Temperature dependent ageing mechanisms in lithium-ion batteries–a post-mortem study. *Journal of Power Sources 262*, 129–135.

Wang, D., J. Coignard, T. Zeng, C. Zhang, and S. Saxena (2016). Quantifying electric vehicle battery degradation from driving vs. vehicle-to-grid services. *Journal of Power Sources 332*, 193 – 203.

Wang, J., J. Purewal, P. Liu, J. Hicks-Garner, S. Soukazian, E. Sherman, A. Sorenson, L. Vu, H. Tataria, and M. W. Verbrugge (2014). Degradation of lithium ion batteries employing graphite negatives and nickel–cobalt–manganese oxide+ spinel manganese oxide positives: Part 1, aging mechanisms and life estimation. *Journal of Power Sources 269*, 937–948.

Wei, L. and Y. Guan (2014). Optimal control of plug-in hybrid electric vehicles with market impact and risk attitude. *Transportation Science 48*(4), 467–482.

Wetterdienst, D. (2014). Historische stündliche lufttemperatur station id 3379.

Zhang, Y., G. W. Gantt, M. J. Rychlinski, R. M. Edwards, J. J. Correia, and C. E. Wolf (2009). Connected vehicle diagnostics and prognostics, concept, and initial practice. *IEEE Transactions on Reliability 2*(58), 286–294.

Zoepf, S., D. MacKenzie, D. Keith, and W. Chernicoff (2013). Charging choices and fuel displacement in a large-scale demonstration of plug-in hybrid electric

vehicles. *Transportation Research Record: Journal of the Transportation Research Board* (2385), 1–10.