Article

Model-Based State-Of-Charge and State-Of-Health Estimation Algorithms Utilizing a New Free Lithium-Ion Battery Cell Dataset for Benchmarking Purposes

Steven Neupert [1,†] and Julia Kowal [2]

[1] Technische Universität Berlin, Department Electrical Energy Storage Technology; s.neupert@tu-berlin.de
[2] Technische Universität Berlin, Department Electrical Energy Storage Technology; julia.kowal@tu-berlin.de
* Correspondence: s.neupert@tu-berlin.de
† Current address: Technische Universität Berlin, Department Electrical Energy Storage Technology; Einsteinufer 11; 10587 Berlin; Germany

Abstract: The state estimation for lithium-ion battery cells has been the topic of many publications concerning the different states of a battery cell. They often focus on a battery cell's state of charge (SOC) or state of health (SOH). Therefore this paper introduces a, on one hand, a new lithium-ion battery data set with dynamic validation data over degradation and on the other hand a model-based SOC and SOH estimation based on this dataset as a reference. An unscented Kalman filter-based approach was used for SOC estimation and extended with a holistic ageing model to handle the SOH estimation. The paper describes the dataset, the models, the parameterisation, the implementation of the state estimations, and their validation using parts of the dataset resulting in a SOC and SOH estimation over battery life. The results show that the dataset can be used to extract parameters, design models based on it and validate with dynamically degraded battery cells.

Keywords: SOC; SOH; Dataset; Ageing; Model; Estimation

1. Introduction

For lithium-ion batteries in applications, it must be ensured that they are operated in a safe operating area (SOA). The SOA includes boundaries for the voltage, the temperature, and the current. [1,2] To support that, different battery states must be tracked during operation. In addition, the battery management system should guarantee a reliable and efficient operation to ensure a safe operation. Therefore, the BMS tasks include collecting measurements of voltages, temperatures, and the system current to ensure the operation in the SOA and the estimation of the battery states that support the task for a safe and efficient operation. [3] The typical states the BMS estimates for lithium- batteries include the state of charge (SOC), the state of function (SOF), state-of-health (SOH) and the remaining useful life (RUL). Additional states are sometimes mentioned in the literature, such as the state of balance or the state of temperature. This paper considers two of these battery cell states because of their importance in a BMS. The state of charge (SOC) generally describes the charge available in the battery cell compared to a fully charged cell. Therefore, the SOC describes how long a battery cell may last with this charge and thus reflects in an electric vehicle the remaining range similar to a fuel gauge in an internal combustion engine vehicle. Electrochemically it describes the average lithium concentration of intercalated ions in the negative electrode. This is why the SOC cannot be directly measured in an application. The cell voltage includes the SOC to some extent, mainly the relation of OCV and SOC, but the surface concentration of lithium ions influences it in contrast to the SOC. [1–6] Since the SOC is essential for improving efficiency and influences battery cell ageing, it has to be estimated. A further hindrance is that it highly depends on the actual capacity of the battery cell, which is affected by the temperature, the state of ageing, and the current rate.

 doi:10.20944/preprints202305.0621.v1

[1] In addition to the general importance of the SOC, it is the foundation of other states. Its mathematical description is as follows.

$$\mathrm{SOC} = \frac{Q_{remain}}{C_{est}}, \tag{1}$$

where Q_{remain} describes the concentration of the intercalated lithium ions in the structure of the negative electrode, and C_{Est} is the current maximum capacity (estimated or measured) based on the temperature and the ageing state.

The SOH describes the degradation of the battery cell. It describes the battery cell as 100% SOH when it is considered new. In most cases, the parameters used to describe the SOH are the capacity and the internal battery cell internal resistance (IR). The initial quantities of the capacity and IR can be either the nominal values gathered from the datasheet or the measured parameters during the first check-up. When using the nominal capacities, the SOH can be over 100% because of the production tolerances of the single battery cells. A classic definition of the SOH is based on the ratio of the estimated and the nominal or initial capacity

$$\mathrm{SOH_C} = \frac{C_{est}}{\mathrm{C_{nom}}}. \tag{2}$$

This definition is specifically useful when the application is mainly interested in the energy capability of the battery cell. But it could be extended with the SOH that bases on the battery cell's internal resistance defined by

$$\mathrm{SOH_R} = \left| \frac{R_{est} - \mathrm{R_{EOL}}}{\mathrm{R_{nom}} - \mathrm{R_{EOL}}} \right|. \tag{3}$$

Depending on the scaling and combination of the definitions, the SOH describing the end-of-life could either be 80 % or 0 %. In this paper, it is 80 %. This work aims to show a new dataset that is freely available. It introduces and describes the measurements conducted (section 2) and the development steps of a state estimation approach for SOC (section 4) and SOH (section 5) used in multiple applications based on the measurements. The parameterisation of the model used for the model-based approach for SOC and SOH estimation is described (section 3) and will be validated on dynamically aged battery cell data (section 7). Therefore the paper sums up the approach for developing state estimation algorithms from the absolute beginning of the measurements to the modelling, algorithms, and validation at the end. It delivers the possibility to benchmark other state estimation algorithms in comparison to a common approach and allows for showing the advantages or disadvantages of developed approaches. Furthermore, the dataset could be used to develop new algorithms. That means the paper includes only a few new or improved steps throughout the developed state estimation but in a highly condensed manner. A literature review of the different SOC and SOH algorithms shows different categories of algorithms. The categories derived are direct measurements, model-based approaches, data-driven approaches, and hybrids of two or more categories.

1.1. Direct measurements for SOC and SOH estimation

Direct measurement approaches in battery cell state estimation describe measurements that can be directly used to estimate a state. That could be specific parameters, just the voltage, temperature, or current at specified points in time, depending on the state to be estimated.

One typical measurement is the measurement of the OCV to use the relation of the SOC and the OCV. The relation is measured before application for the battery cell and used during runtime to estimate the SOC. [7] The OCV can be measured after hours without a current, therefore in an application, the assumption, which is similar to the assumption for doing the short-term incremental OCV measurement and the low current measurement, that is, that the overpotential either after a short resting phase or during a low current section is nearly equal to the true OCV is applied, because in many cases there are no long

 doi:10.20944/preprints202305.0621.v1

resting phases. Often there are still small currents to supply the control units. This kind of SOC estimation is often used for calibration, to augment other algorithms, or under laboratory conditions.[3] An analysis based on the relation of the OCV and the SOC that slightly changes over degradation is used to estimate the SOH. The analysis is called ICA and describes the differentiation of the charge over the voltage, leading to peaks in the region of phase changes. That ICA is calculated from current measurements at constant currents for a full or a partial discharge/charge cycle. The battery cells' SOH can be estimated by tracking the position, amplitude, and enveloped area. [7,8] Of course, the change of these peaks has to be investigated before the application to know which change belongs to which SOH. However, still, these measurements are only applicable in some applications.

There is the well-known coulomb counting method. The coulomb counting approach integrates the charged or discharged charge over time. That integrated value is divided by the current capacity to estimate the SOC. In this case, the starting point, the starting SOC, of the coulomb counting approach has to be known. The coulomb counting method depends on the capacity, therefore on the SOH, and is influenced by the temperature and the current rate. This method is prone to measurement errors since they are accumulated over time. [3,7] Due to this disadvantage, it is often combined with the OCV measurement, where it is possible to recalibrate the counter. Based on the coulomb counter, a cycle counter could be established that counts the equivalent full cycles during usage. Therefore it is extended such that it counts the charge in the charge direction or the absolute value of the current in both the charging and discharging direction and is divided by either the capacity or twice the capacity. To be able to use this information correctly, many measurements have to be taken out to get the relation between the number of equivalent full cycles and the SOH because the trajectory of the degradation depends on the stress factors like the current rate, the temperature, the ΔDOD and the mean voltage. The measurement of impedance or internal cell resistance is another method to estimate a state—especially the SOH correlations with the increase of the internal cell resistance. Different methods are used to measure the internal resistance. Sometimes the internal resistance is tried to be identified at current changes during the application and filtering the results afterwards. This approach is often not measuring the actual internal resistance because the sample rate of the current and voltage measurement is too slow to get the internal resistance isolated. So what is measured is a combination of the internal resistance and the impedance of active electrochemical processes. More common is applying a current pulse to identify the internal resistance [8,9]. Same here, the measured resistance might not be the internal resistance, but it is part of it. The EIS measurement with a single frequency is possible to measure the true internal resistance. More frequencies are applied to gain more information using, for example, a multi-sine approach or just different frequencies as in a typical EIS measurement [3,8]. Furthermore, typical EIS in applications is researched to get more information on the battery cell impedance. Different electrochemical mechanisms can be analysed without opening the cell by investigating the EIS. Therefore, it is very interesting to use EIS, since it is a tool that could make a good diagnostic of ageing processes possible. [3,8,9] The application of EIS measurement is very complex because it is susceptible to changes in the measured system, which can be changes in the connection, temperature changes, and other disturbing influences. Another measurement technique that might be used is based on the joule effect. This approach analyses the generated heat and the rise of the temperature of the battery cell to identify the internal resistance. However, a calorimeter is needed to track temperature change and heat generation. [8,9] Other measurement equipment might be used as well [8]. Of course, using a calorimeter as an online measurement tool is unsuitable.

1.2. Model-based SOC and SOH estimation

In general, model-based state estimation approaches depend on the model; that could be that the approach only uses a model or combines an algorithm like a filter or observer with a model. Using models directly to estimate the battery cell state is primarily

 doi:10.20944/preprints202305.0621.v1

not used for SOC estimation. Suppose coulomb counting or the OCV relation to the SOC is considered model-based, then they are the only models for model-based SOC estimation. It is another case for SOH or RUL estimation. The basic models used for direct estimation are semi-empiric, empiric, and electrochemical ageing models used to estimate the SOH. Empiric and semi-empiric models are developed based on ageing measurements considering different stress factors. The model is fitted to the change of a parameter over time or full cycle equivalents, i.e., the internal resistance or the capacity, and is used to calculate the current SOH or RUL considering the cycles or the time at a specific stress factor. A weighted coulomb counter based cycle counter can be used to track the influence of the stress factors. The same approach can be used if the empiric model is replaced with the degradation data saved in a characteristic map. [7,8,10] By using a more sophisticated model like an electrochemical pseudo-two-dimensional model and using different ageing improvements, the extrapolation ability and, therefore, the prediction ability is improved for unseen data. In addition, the model can diagnose the ageing process. However, due to their complexity and the involvement of many parameters and multiple partial differential equations, they are unsuitable for online usage. Therefore the research focuses on simplifications and parameterisation. [7,8] Every other model-based state estimation approach includes a filter or observer. Filter and observer share that they use the same general framework. They both rely on the model-based prediction (time update) of the battery cell state and use the measurement to adjust the prediction (measurement update) to the current behaviour of the battery cell. Filters that are often used belong to the Kalman filter family, namely the extended Kalman filter or sigma-point Kalman filter, of which the unscented Kalman filter is a part of. They use the general approach of making model-based predictions and updating the prediction by measurement using the so-called Kalman gain. The different modifications of the Kalman filter try to linearise the system around the current working point using the derivative or an approximation. Known observers include the Luenberger and the sliding mode observer.

1.3. Data-driven SOC and SOH estimation

The data-driven approaches contain machine learning approaches that describe a group of algorithms capable of inferring a battery's behaviour from raw data to build a model that can predict the output or a state depending on the data without directly programming it for this specific behaviour. The methods used for state estimation can be categorised into methods that directly predict the output based on data and the methods that iteratively predict the output, where the output could be a state, parameters, or the voltage of the battery cell. Whether the prediction structure is direct or iterative depends on the task and algorithm. An algorithm that makes point predictions can have a direct structure, whereas an algorithm that predicts sequences or simulates something will have an iterative structure. Depending on the output, they can be further distinguished into algorithms that produce outputs with a measure of certainty/probability or not, probabilistic or non-probabilistic methods. Machine learning approaches include neural networks, autoregressive-moving average models and support vector machines. [7,11] The algorithms used in this paper are model-based approaches often used in applications. Most machine learning approaches are not applied in real applications. So far, only startups use machine learning approaches for SOH estimation. To parameterise the ageing of the battery cells, the degradation is interrupted by check-ups that consist of a capacity test, an OCV measurement using an incremental approach with a short relaxation time instead of a low current measurement and pulse tests for the parameterisation of the dynamic parts of the battery cell model. In general, the concept for the state estimation consists of a dynamic model describing the overpotentials and OCV, a model describing the ageing of the parameters of the battery cell, an unscented Kalman filter to derive the SOC, and an algorithm to derive the SOH from the ageing model.

 doi:10.20944/preprints202305.0621.v1

2. Measurements and Dataset

The dataset consists of 30 cells called HE4 from LG Chem aged in a hybrid ageing manner, so no differentiation between calendar or cycle ageing was made. Still, the battery cells were aged under different stress factors considering two temperatures (room temperature (23 °C, 25 °C and 45 °C), three different mean SOCs and three different DODs. Different currents are not considered during constant current ageing. In addition, it includes two cells aged under dynamic conditions at room temperature (23 °C). The following table 1 sums up the different stress factors and displays the stress factor matrix.

Table 1. Table of the different stress factors considered in the ageing measurements.

Temperature [°C]	ΔDOD	SOC_{mean}	Number of Cells	Cyclic Ageing
25	1	0.5	4	Constant current
25	0.3	0.35	4	Constant current
25	0.3	0.65	4	Constant current
45	1	0.5	4	Constant current
45	0.3	0.35	4	Constant current
45	0.3	0.65	4	Constant current
23	0.7	0.5	2	Constant current
23	0.3	0.5	2	Constant current
23	1	0.5	2	Dynamic profile

The ageing is frequently interrupted by a battery cell capacity test, every ten full cycle equivalents, and less frequently by a full characterisation. During the ageing process, the battery cells are cycled using constant current for charging and discharging, except for the two dynamically aged cells.

The characterisation is taken out at the specific ageing temperature since the outer mean SOC and DOD combinations are conducted at the two temperatures. The check-up characterisations consist of a capacity test, a hybrid of an incremental open circuit voltage, and a slightly adjusted hybrid pulse power characterisation (HPPC). Pre-tests have shown that the influence of the HPPC on the OCV is neglectable. Furthermore, the incremental method is an exact and time-efficient type of measuring the OCV and being able to do additional analysis, namely the differential voltage analysis [12]. The increments in OCV are based on every percent of the nominal capacity and the HPPC at every ten percent. The measurement equipment used consists of a Neware battery cycler with a voltage range of 10 V and ±10 A and either a Memmert oven for the 45 °C or a temperature chamber from Binder for the 25 °C measurements. The check-up is extended for the two dynamically aged cells with a validation part, including dynamic discharges and constant current constant voltage (CCCV) charging procedures. Figure 1 displays the used check-up and validation extension.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

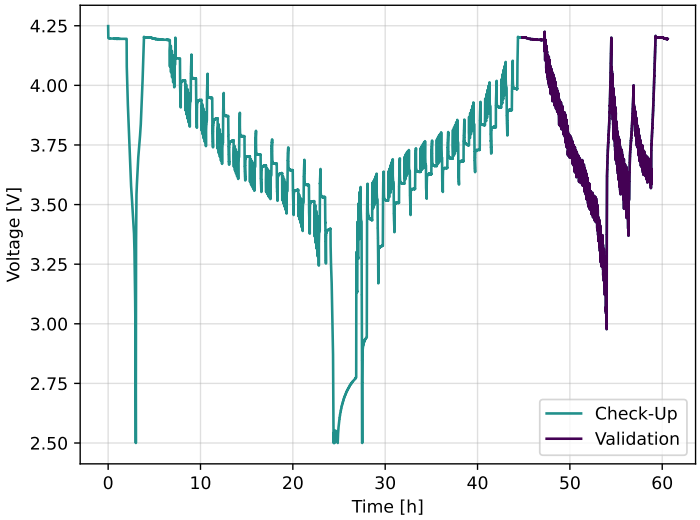

Figure 1. Voltage diagram of check-up procedure, including capacity test, hybrid OCV and HPPC and the validation.

For the dynamic ageing, the Federal Urban Driving Schedule was applied to a battery cell of the same type, so the cell was CCCV charged and then discharged to 2.5V using the FUDS repeatedly. This measurement was used to design dynamic parts of the ageing. To generate these parts, the so-called Random Pulse Method was used to create profiles of defined length, as explained in [13]. The general process comprises segmenting the measured current profiles in charging and discharging pulses and saving them to a database. Afterwards, the pulses are selected randomly and added to the profile. Even without actual driving data, a dynamic profile with random pulse sequences is generated that is ready to be used during the cyclic ageing of the cells. The cycling includes short cycles of 25%, 30%, 80%, and 95% DOD. These cycles are concatenated to sum up to about ten full cycle equivalents followed by a capacity test (see Figure 2).

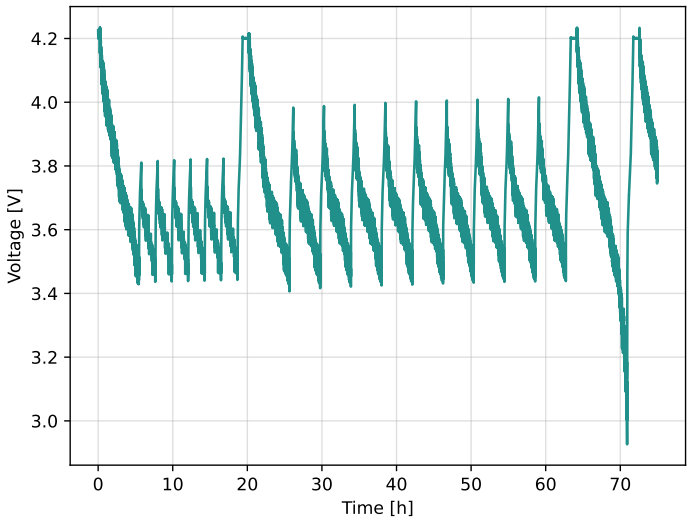

Figure 2. Voltage diagram of dynamic cycling procedure, including a capacity test.

 doi:10.20944/preprints202305.0621.v1

3. Dynamic Model Description

The modelling of a system has multiple purposes. It supports the design, analysis, verification, or validation of a system. In the case of BMS, it serves the purpose of analysing the system in terms of diagnosing and prognosing the actual state of the battery system. Different modelling approaches can be physio-chemical, equivalent circuits, or mathematical models. The data that can be measured is essential, and on the other hand, how efficient is the computation of the model?

Depending on the purpose, the model differs by the information needed to build the model. They are often divided into black box and white box models. Black box models do not need any system knowledge to model the behaviour. Empirical or mathematical models often represent them. In contrast, white box models are based on describing the processes that lead to specific behaviour of the system. These models could consist of multiple differential equations. In other words, white box models are based on knowledge, whereas black box models are based on data to represent the system's behaviour. The hybrid of these models is called the grey box model and uses both the knowledge and data-based approach to model the system's behaviour. This approach helps model a system where some parts of the internal processes are known and others are not. It is also used to simplify a model to reduce its complexity and focus on parts where a diagnosis should be possible.

In terms of battery models, the different categories of models lead to electrochemical models (white box), data-driven models like neural networks (black box), or equivalent circuit models with additional empirical models to reproduce the behaviour of battery cells.

The base of most electrochemical battery models is formed by the work by Doyle, Fuller, and Newman described in [14]. They describe the model as having a one-dimensional transport from the anode to the cathode, passing the polymer separator. Film formation at the lithium/polymer interface is discarded to keep the complexity manageable. At the same time, the transport is modelled via the concentrated solution theory leading to a binary electrolyte and a polymer solvent with one phase. Therefore the transport in the electrolyte is defined by the electrical conductivity, the transference number, and the diffusion coefficient. These parameters depend on the concentration and could lead to different physical properties. By doing so, the transport phenomena could be handled by that approach.[14] All in all, it is described such that it leads to a pseudo-two-dimensional problem based on a one-dimensional problem on micro and macro scale [15]. Since the process is nonlinear and pseudo-tow-dimensional, it results in an issue that is only numerically solvable and, therefore, unsuitable for every task like parameter estimation [15].

For further usage, the Doyle Fuller Newman model is simplified to form the so-called single particle model, which assumes a uniform reaction rate across the electrodes and neglects the electrolyte dynamics [16]. By these assumptions, the electrodes are modelled as single particles [17]. Electrochemical models could be used in applications like SOH estimation based on that simplification. Further improvements were made in research to represent degradation based on solid electrolyte interface development, cracking, and handling electrolyte dynamics [17–19]. However, they are still represented by complex differential equations with many parameters.

Electrochemical models are mostly either single-particle models or pseudo-two-dimensio models mainly used for design and analysis because they need detailed information on the battery structure, the materials, and the electrolyte [20]. They were also used for state estimation in some cases [21,22]. Nevertheless, they are complex and lack usability when the information on the battery cell is limited.

Data-driven models include every kind of mathematical model that could be used to reproduce the behaviour of a battery cell. They could be built up from neural networks, support vector machines, fuzzy logic-based, or others. These models are often called black box models since they do not deliver an analytical insight. So their parameters are not relatable to the system that should be modelled. However, they could have a good performance computationally and error-wise. They could be used for monitoring,

 doi:10.20944/preprints202305.0621.v1

diagnostics, design, and understanding physical phenomena. Nevertheless, depending on the application task, the correct data to train the model is needed, which could be very time-consuming since the data-driven models tend to overfit the data and cannot extrapolate the data to states beyond the training data. [6,23,24] The widely used equivalent circuit models (ECM) can represent the dynamic behaviour resulting from the reaction's electrochemical processes during charge or discharge. Depending on the application, the ECM is employed as a straightforward model that is computationally efficient. [25,26] Two different approaches for ECM try to emulate the battery behaviour using mostly common electrical components such as resistors, capacitors, and voltage sources. Their aim differs to a small extent since one approach is completely empirical, and the other aims to model the behaviour with electrochemical relations—the difference in generating lies in the underlying measurements. ECM with electrochemical relations starts with electrochemical impedance spectroscopy (EIS) measurements, and most empirical ECM starts with pulse measurements.

The laboratory measurements during the testing included a capacity test, an OCV, and pulse measurements. For electrochemical models, invasive measurements are needed. Nevertheless, it is still possible to identify the parameters of an ECM and data-driven models.

Overall, the ECM promises good computability, interpretability, and efficiency with a reasonable amount of parameters, and furthermore, it is often used in combination with model-based state estimation approaches.

The boundary conditions must be considered when designing an ECM for a specific application. That means which measurements could be taken out, which data could be collected, where the model should be used, and what is expected of the model. Following the literature on the next steps of a BMS it is the online diagnostics of battery cells, identifying which processes lead to the current behaviour of the cell. Hence, the literature on modelling for BMS usage seeks for at least potent physics-based ECM [23] or even reduced order models of electrochemical models [17,21,22]. For most physics-based ECMs, at least EIS measurements are needed. Speaking of the boundary conditions in an application, the most important for the model design and the parameter estimation are the measurements, sampling rate, and resolution. Since no specific measurement schemes are taken out in an application, the boundary conditions reduce furthermore to the resolution and rate. The resolution in a BMS ranges from 8 bit to 16 bit, hence, theoretically for a measurement in the range of 0 V to 5 V from 9.8 mV to 38 µV accuracy, but considering the noise from 16 mV to 300 µV accuracy [27]. The sampling rate that is applied in electric vehicles depends on whether it is a high-power application where a faster rate is needed or a high-energy application where a slower rate is feasible. Nevertheless, the rates range from $1\,s^{-1}$ to $10\,s^{-1}$ [27]. Specifically looking at the sampling rate, the number of electrochemical processes that could be investigated shrinks. Taking account of the sampling rate, the following figure 3 displays the electrochemical processes that could be modelled by showing the EIS, which includes a small part of the charge transfer, the diffusion, and the change of crystalline structure. This figure does not take the Nyquist-Shannon sampling theorem into account.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

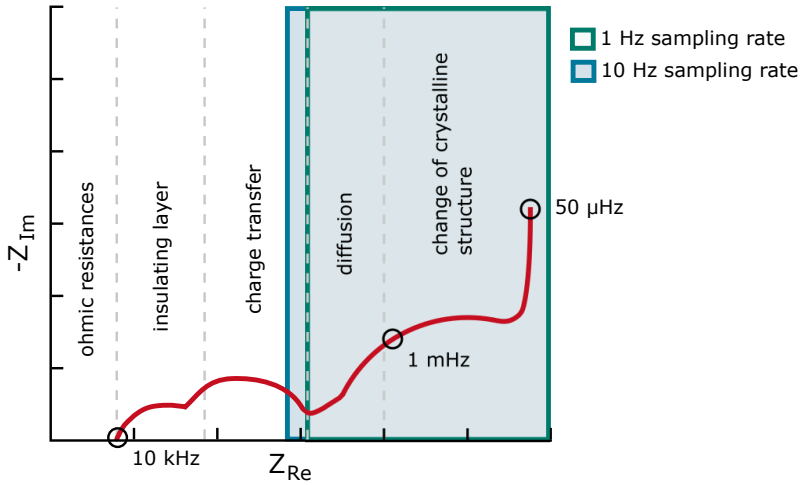

Figure 3. Electrochemical impedance spectra considering the sample time.

Different ECMs are still usable, considering the impedance spectra and sampling rate. They can be of fractional order or can include multiple RC elements. These models can be lumped. Since the main focus of the paper is not the modelling, the model used is a typical two-RC model with an OCV part. This type of model is an often used model when a model-based state estimation approach is followed [28,29] even of a lower order [30,31]. Overall the two RC model is able to model the EIS of a lithium-ion battery. Therefore, it is suitable to be used as a model for the SOC estimation and the representation of the battery cell in case of dynamic battery cell behaviour.

3.1. Dynamic Model Equations and Discretisation

This section states the model equations and their discretisation for the dynamic battery cell model. The model is displayed in figure 4. The differential equation describing the model could be established by using Kirchhof's and Ohm's laws and considering that $\tau = R \cdot C$.

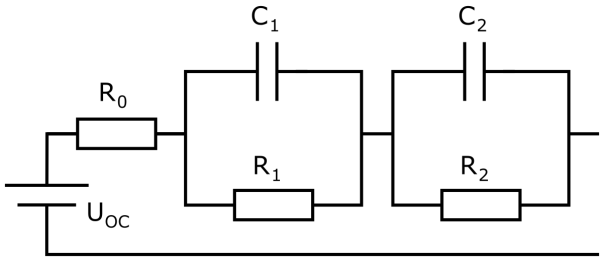

Figure 4. Battery model with two RC elements

$$U(t) = U_{OC}(t) + U_0(t) + U_1(t) + U_2(t) \tag{4}$$

$$U(t) = U_{OC}(t) + R_0 i(t) + R_1 i(t) - \tau_1 \frac{\mathrm{d}U_1(t)}{\mathrm{d}t} + R_2 i(t) - \tau_2 \frac{\mathrm{d}U_2(t)}{\mathrm{d}t} \tag{5}$$

For the next step to get to the equations needed, the differential equation is Laplace transformed using the law of linearity and leads to

$$U(s) = U_{OC}(s) + R_0 i(s) + \frac{R_1}{1+\tau_1 s} i(s) + \frac{R_2}{1+\tau_2 s} i(s). \tag{6}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

By applying the bilinear transformation

$$s = \frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} \tag{7}$$

and discretising every term on its own leads to the following results.

$$Z\{U_{OC}(s)\} = U_{OC}(z) \tag{8}$$

$$Z\{U_0(s)\} = Z\{R_0 i(s)\} = R_0 i(z) \tag{9}$$

$$\begin{aligned} Z\{U_1(s)\} &= Z\{\frac{R_1}{1+\tau_1 s} i(s)\} \\ &= \frac{R_1}{1+\frac{2}{T}\tau_1} i(z) + \frac{R_1}{1+\frac{2}{T}\tau_1} i(z) z^{-1} - \frac{1-\frac{2}{T}\tau_1}{1+\frac{2}{T}\tau_1} U_1(z) z^{-1}. \end{aligned} \tag{10}$$

$$\begin{aligned} Z\{U_2(s)\} &= Z\{\frac{R_2}{1+\tau_2 s} i(s)\} \\ &= \frac{R_2}{1+\frac{2}{T}\tau_2} i(z) + \frac{R_2}{1+\frac{2}{T}\tau_2} i(z) z^{-1} - \frac{1-\frac{2}{T}\tau_2}{1+\frac{2}{T}\tau_2} U_2(z) z^{-1} \end{aligned} \tag{11}$$

By using these equations and transforming them to the discrete-time domain and formulating them in state space leads to the model that state estimation algorithms could handle

$$\begin{bmatrix} SOC(k) \\ U_1(k) \\ U_2(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1-\frac{2}{T}\tau_1}{1+\frac{2}{T}\tau_1} & 0 \\ 0 & 0 & \frac{1-\frac{2}{T}\tau_2}{1+\frac{2}{T}\tau_2} \end{bmatrix} \begin{bmatrix} SOC(k-1) \\ U_1(k-1) \\ U_2(k-1) \end{bmatrix} + \begin{bmatrix} \frac{1}{C_N} & 0 \\ \frac{R_1}{1+\frac{2}{T}\tau_1} & \frac{R_1}{1+\frac{2}{T}\tau_1} \\ \frac{R_2}{1+\frac{2}{T}\tau_2} & \frac{R_2}{1+\frac{2}{T}\tau_2} \end{bmatrix} \tag{12}$$

$$U(k) = U_{OC}(SOC(k)) + U_1(k) + U_2(k) + R_0 i(k). \tag{13}$$

3.2. Parameterisation

The parameterisation is the process of extracting the model's parameters from the measurements. To be able to extract the parameters, every ageing cycle is frequently interrupted by a check-up. Every check-up is preprocessed to get the measurement's capacity, the OCV, and the pulses out. Before preprocessing, measurements are imported and formatted into a universal format because only the import function has to be changed if the saved measurement file changes. The capacity is calculated from the 1C discharging; a CCCV charge precedes that. Another CCCV charge follows this discharge and the hybrid test phase of the incremental OCV and the HPPC pulse test. The OCV in the charging and discharging direction can be estimated from the hybrid phase. See figure 1 for an example of the check-up. During the preprocessing, the check-up data in a standardized import data format is separated into different parts, starting from the OCV voltages extracted from the hybrid part from the resting phase of the incremental OCV to the relaxations of the HPPC pulses that are used to determine the dynamic parameters of the battery model, the battery internal resistance, the resistances of the RC elements and its time constants. By using the pulse relaxation, the change of OCV does not have to be included because there is none during the relaxation. Furthermore, the capacity is directly extracted from the 1C discharging, and the validation is extracted as well if present. The validation part is only present for the cells aged under dynamic regimes.

3.2.1. Open circuit Voltage Modelling

The OCV is a central part of the model since it describes the midterm behaviour of a battery if the battery behaviour is segmented into short-term, dynamic behaviour, mid-term, OCV behaviour, and long-term, described by the ageing behaviour. There are different

 doi:10.20944/preprints202305.0621.v1

approaches to modelling the OCV in the literature, ranging from tables to empirical models like polynomials or more electrochemical-based models [32–35]. For more information on different models for OCV modelling, please refer, for example, to Pillai et al. [35]. To model the OCV, a table-based approach will be used, which can lead to a stable and easy method to represent the OCV when sampled correctly. The used approach is based on the inflection point method described by Sundaresan et al. [32]. A few changes were added to the sampling to improve the performance. The following steps were followed to sample the measured OCV curve:

1. Calculation of the signed curvature of the OCV according to Narula [36]

$$\kappa = \frac{x''y' - x'y''}{((x')^2 + (y')^2)^{3/2}} \tag{14}$$

2. Find roots of the curvature to segment the OCV. Therefore calculate

$$\kappa(x) = 0 \tag{15}$$

   to get the roots. Add the first measurement point and roots to the samples.

$$samples = [x(0), \kappa_{roots}, x(end)] \tag{16}$$

3. Delete inner samples if the section range is lower than 5% of the SOC for minimum section size.
4. Allocate samples equally to the sections.
5. If there are still samples left:
   (a) Calculate the error and curvature peaks per section
   (b) Add a sample to the section with the highest error
   (c) Distribute samples equidistantly in sections
   (d) Go to 5
6. Finished sampling the OCV

The main differences to Sundaresan et al. [32] is the formula for calculating the curvature and the steps to distribute the samples, despite the equal distribution. Because the approach described in the literature could lead to many samples in just two sections, even if other sections would benefit from additional samples, this distribution depends on the number of excess samples after the equal distribution. If there are no samples left, the step of distributing the excess samples is not conducted. The following figure 5 displays the result of the sampling compared to the measurement.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

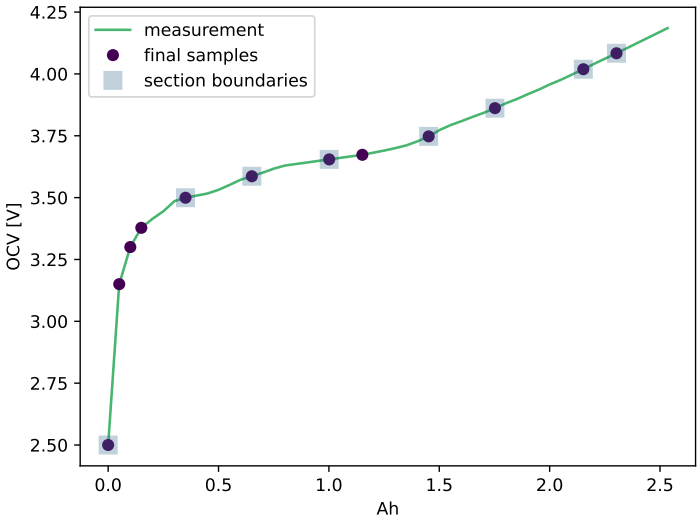

Figure 5. Sampling result of an OCV measurement.

3.2.2. Dynamic Parameter Identification

The battery model's dynamic parameters are the battery cells' ohmic resistance, the resistances belonging to the RC elements, and the time constants. The time constants instead of the RC capacities are used because they are more simple for identifying boundaries, and the distribution of relaxation times could be easily used for start parameter estimation. The relaxation of the HPPC part of the check-up is used to identify the parameters. Identifying the parameters is divided into the preprocessing of the relaxation, the start parameter estimation and the fitting itself. The relaxation consists of switching from the current pulse to the relaxation and the complete relaxation. The measurement is checked for doubled data points and NaNs during preprocessing. Afterwards, the relaxation is freed from the OCV by subtracting the last voltage of the relaxation. Because the relaxation is 1800 s long and sampled with 0.1 s, it consists of many samples, and many of these samples are part of the relaxed voltage. Many samples in this area work during the fitting process as a weighting since the faster processes should be fitted as well, and the relaxation is resampled. Logarithmic resampling was used to have more samples at the beginning of the relaxation and fewer at the end.

Estimating start parameters is essential; the better the start parameters, the better the fit. In this case, estimating the start parameters is relatively complex. The overall process consists of different steps starting with the estimation of the ohmic series resistance using

$$R_0 = \frac{U(t = 0.3\,\mathrm{s}) - U(t = 0\,\mathrm{s})}{I(t = 0.3\,\mathrm{s}) - I(t = 0\,\mathrm{s})}. \tag{17}$$

Then the next step includes estimating the resistances and time constants of the RC elements. The step includes calculating the so-called distribution of relaxation times (DRT) to get meaningful parameters. Since no electrochemical impedance spectroscopy was measured, the DRT is estimated by overfitting the relaxation with 5 RC elements, calculating the impedance spectra using the fitted parameters and the impedance equation, and calculating the DRT using the package pyDRTtools by Wan et al. [37]. The DRT is evaluated for local peaks, whether they lie in the range of the boundaries used for fitting the parameters later on. If they lie in the range of one of the ranges of the time constants,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

the time constant is used as starting parameter of the specific RC element time constant. The resistance corresponding to the respective time constant is calculated by

$$U_{RC,1} = U_{pp}(t = 0) - U_{pp}(t = 5\tau_{RC,1}) \tag{18}$$
$$R_{RC,1} = U_{RC,1} / I_{pulse,t=0} \tag{19}$$

for the first and the other RC elements

$$U_{RC,n} = U_{pp}(t = 5\tau_{RC,n-1}) - U_{pp}(t = 5\tau_{RC,n}) \tag{20}$$
$$R_{RC,1} = U_{RC,n} / I_{pulse}(t = 0). \tag{21}$$

where U_{pp} is the preprocessed measured relaxation voltage, $R_{RC,n}$ the resistance and $\tau_{RC,n}$ the time constant of the n-th RC element. At this point, the start parameter estimation is finished. For robustness, the start parameters for time constants are the mean of their boundaries if no peak in the region of the boundaries is found.

Now the data is fitted using the lmfit package by Newville et al. [38] applying the minimizer with the Nelder-Mead algorithm. Only the parameters of the RC elements are fitted, and the ohmic resistance is kept constant throughout the fitting process. Because the fitting is done on a relaxation process, the fitted model has to include the starting voltage of the RC elements because instead of having 0 V at the RC element, the capacity is pre-charged by the current pulse before. Therefore the discrete equation for a single RC element itself is as follows.

$$U[k]_{RC,n} = e^{\frac{-t[k]}{\tau_{RC,n}}} U[k-1]_{RC,n} + R_{RC,n} I[k] \left(1 - e^{\frac{-t[k]}{\tau_{RC,n}}}\right) \qquad , k = [1, size(t)] \tag{22}$$

$$U[k]_{RC,n} = R_{RC,n} I[k] (1 - e^{\frac{-t_{pulse}}{\tau_{RC,n}}}) \qquad , k = 0 \tag{23}$$

See figure 6 for a start parameter result and a fitting result.

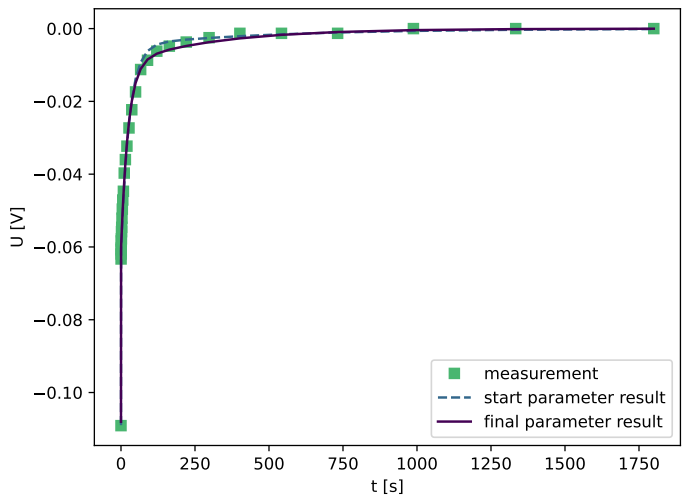

Figure 6. Result of the start parameter estimation and the fitting.

These processes are applied to every check-up and the extracted relaxations.

4. SOC Estimation

The Kalman filters are a widespread family of filters used for state estimation. The general Kalman filter is the simplest of the filters. It recursively uses the system described

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

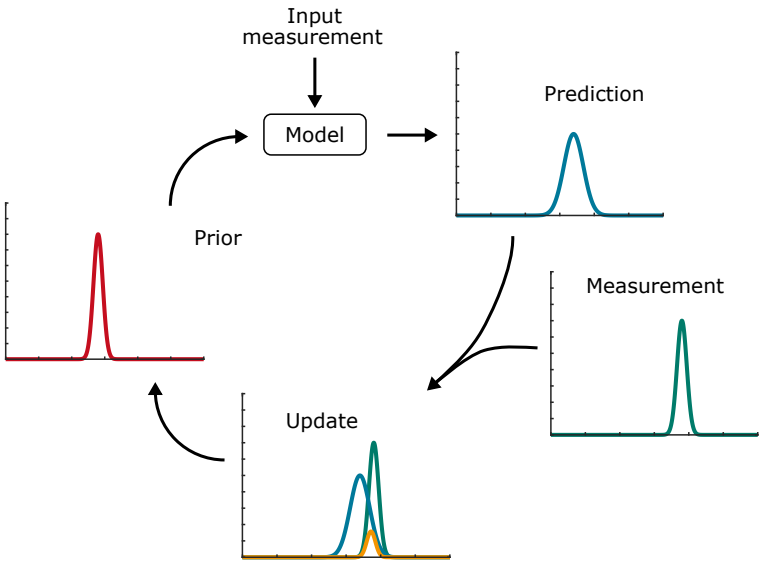

Figure 7. Steps of the Kalman Filter.

in state space and the measurements to estimate the state. By weighting the prediction and the measurement, the Kalman filter can account for noise, variations in the measurement, and inaccuracies of the used model. Since it is based on the Bayesian filter, it considers the state to be a normal Gaussian distribution described by its mean and the variance or covariance matrix for multiple dimensions. So, by using the model for the prediction, a Gaussian distribution is estimated based on its last state and the model's behaviour described by the state space equations. Uncertainties are added to account for model and measurement imperfections. The measurement represents another Gaussian distribution, and by multiplying the prediction with the measurement, a new distribution can be calculated. By extracting the so-called Kalman Gain, the new state could be estimated. Figure 7 summarises the mentioned steps. The Kalman filter is not feasible for battery state estimation since it only operates with a linear system, which is not the case for battery models and leads to errors during estimation. Different variants have been developed. [3,7–9]

The Extended Kalman filter (EKF) is one variant that linearizes the current mean and the covariance with a Taylor series approximation in every step to nullify the nonlinearities. In general, the EKF uses the same framework of time update and measurement update but calculates each time the Jacobian matrix of partial derivatives of the nonlinear state equation for the state and the process noise and Jacobian matrix of partial derivatives of the measurement equation for the state and the measurement noise. The problem with this kind of linearization is that the distributions of random variables are no longer normal after the nonlinear transformation. Often the EKF is used to estimate the SOC using an ECM or the SOH based on estimating the capacity and internal resistance. [9,31,39,40]

Another variant of approaches summarized under Sigma Point Kalman filter (SPKF) handles the nonlinear problem using a statistical approach. So, the SPKF uses multiple Sigma-Points depending on the dimensions of the state vector. For these points, the time and measurement update is calculated, which means they are all transformed, and afterwards, they are all used to estimate the current state vector. The advantage is that this filter does not calculate Jacobians or Hessians without losing information and precision. SPKF are differentiated based on the selection rule of the Sigma-Points in central difference Kalman filter, unscented Kalman filter or cubature Kalman filter. [3,9,23,31,40,41] In general, the UKF follows the same process as other filters, including the prediction and update step. Of course, these steps can be further divided into smaller substeps. Its name is derived from the unscented transform, an approach used for the statistical calculations of random variables that undergo a nonlinear transformation. Instead of calculating the

 doi:10.20944/preprints202305.0621.v1

Jacobian and Hessian matrices, the UKF uses discrete samples, called sigma points, that are projected through the transformation. By using weights, the mean and the covariance can be calculated. [42,43] The general process is described here:

1. Initialization of the mean and the covariance with expectations, where x is the state vector.

$$\hat{x}_0 = \mathbb{E}[x_0] \tag{24}$$
$$P_0 = \mathbb{E}[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] \tag{25}$$

2. Prediction of the current state, including sampling and weight calculation.
   (a) Calculate the sigma points based on the mean $\hat{x}_{k-1}$, depending on the dimension of the state vector and mean N using the composite scaling factor λ:

$$\mathcal{X}_{k-1,i} = \begin{cases} \hat{x}_{k-1} & ,i = 0 \\ \hat{x}_{k-1} + \sqrt{(N+\lambda)P_{k-1}} & ,i = 1,...N \\ \hat{x}_{k-1} - \sqrt{(N+\lambda)P_{k-1}} & ,i = N+1,...2N \end{cases} \tag{26}$$

   where the scaling factor is

$$\lambda = \alpha^{\beta}(N+\kappa) - N \tag{27}$$

   with α determining the spread of the samples, β depends on the expected type of distribution, where for Gaussian $\beta = 2$, and κ being the scaling factor that's usually equal to $3 - N$.
   (b) Transform the samples using the model system equation (F) and the input (u)

$$\mathcal{X}^*{}_k = F(\mathcal{X}_{k-1}, u_{k-1}). \tag{28}$$

   (c) Calculate the predicted mean $\hat{x}_k^-$ and covariance based on the samples by using the weights $W_{m,i}$ for the mean and $W_{c,i}$ for the covariance in addition to the process noise covariance R^p

$$\hat{x}_k^- = \sum_{i=0}^{2N} W_{m,i}\mathcal{X}^*{}_k \tag{29}$$
$$P_k^- = \sum_{i=0}^{2N} W_{c,i}(\mathcal{X}^*{}_k - \hat{x}_k^-)(\mathcal{X}^*{}_k - \hat{x}_k^-)^T + R^p \tag{30}$$

   and augment samples with additive noise

$$\mathcal{X}_{k,i} = \begin{cases} \mathcal{X}^*{}_{k,i} & ,i = 0, \\ \mathcal{X}^*{}_{k,i,0} + \sqrt{(N+\lambda)R^p} & ,i = 1,...N, \\ \mathcal{X}^*{}_{k,i,0} - \sqrt{(N+\lambda)R^p} & ,i = N+1,...2N. \end{cases} \tag{31}$$

   (d) Calculate the output with the output equation of the model (H) using the samples and calculate its mean with the weights:

$$\mathcal{Y}_k = H(\mathcal{X}_k), \tag{32}$$
$$\hat{y}_k^- = \sum_{i=0}^{2N} W_{m,i}\mathcal{Y}_{k,i} \tag{33}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

3. The Measurement update calculates the new state of the model using the prediction, the Kalman gain, and the measurement of the output. Furthermore, the measurement noise is added.
   (a) Calculate the Kalman gain:

$$P_{yy,k} = \sum_{i=0}^{2N} W_{c,i}(\mathcal{Y}_{k,i} - \hat{y}_k^-)(\mathcal{Y}_{k,i} - \hat{y}_k^-)^T + R^m \tag{34}$$

$$P_{xy,k} = \sum_{i=0}^{2N} W_{c,i}(\mathcal{X}_{k,i} - \hat{x}_k^-)(\mathcal{Y}_{k,i} - \hat{y}_k^-)^T \tag{35}$$

$$\mathcal{K} = P_{xy,k} P_{yy,k}^{-1}. \tag{36}$$

   (b) Calculate the state and covariance of the state of the model using the Kalman gain and the measurement of the output (y_k):

$$\hat{x}_k = \hat{x}_k^- + \mathcal{K}(y_k - \hat{y}_k^-) \tag{37}$$

$$P_k = P_k^- - \mathcal{K} P_{yy,k} \mathcal{K}_k^T. \tag{38}$$

   (c) After the calculation of the current state, it is shifted to be the old state ($\hat{x}_{k-1}$), and the same is done for the covariance (P_{k-1}). Now, start all over again with the prediction steps and so forth.

These process steps are applied to the model described in section 3 to estimate the internal cells of the model consisting of the SOC and the overpotentials of the two RC elements. Now the state estimation consists of the SOC estimation based on the model using the UKF. It is extended by a SOH estimation based on an ageing model.

5. SOH Estimation

As for the SOC estimation, the SOH estimation can be grouped into direct measurements, model-based and data-driven approaches. Direct measurements for the SOH estimation consider an analysis based on the relation of the OCV and the SOC, where the identification of slight changes over the degradation is used to estimate the SOH. The analysis is called incremental capacity analysis (ICA) and describes the differentiation of the charge over the voltage that leads to peaks in the region of phase changes. That ICA is calculated from current measurements at constant currents for a full or a partial discharge/charge cycle. The battery cells' SOH can be estimated by tracking the position, amplitude, and enveloped area. [7,8] Of course, the change of these peaks has to be investigated prior to the application to know which change belongs to which SOH. Another direct approach is the so-called cycle counter. Based on the coulomb counter, a cycle counter could be established that counts the equivalent full cycles during usage. Therefore it is extended such that it only counts the charge in the charge direction or the absolute value of the current in both the charging and discharging direction and is divided by either the capacity or two times the capacity. To be able to use this information correctly, many measurements have to be conducted to get the relation between a number of equivalent full cycles and the SOH, because the trajectory of the degradation highly depends on the stress factors like the current rate, the temperature, the ΔDOD, and the mean voltage. Another approach is to use the relation of the internal resistance with the ageing processes and derives the SOH based on the measurement of the internal resistance either from EIS measurements or pulses. The next group of methods includes the model-based approach, divided into direct and indirect approaches, where direct models that are used for the direct estimation are, for example, semi-empiric, empiric, and electrochemical ageing models used to estimate the SOH. Empiric and semi-empiric models are developed based on ageing measurements that consider different stress factors. The model is fitted to the change of a parameter over time, or full cycle equivalents, i.e., the internal resistance or the capacity, and is used to

 doi:10.20944/preprints202305.0621.v1

calculate the current SOH or RUL considering the cycles or the time at a specific stress factor. A weighted coulomb counter-based cycle counter can be used to track the influence of stress factors. The same approach can be used if the empiric model is replaced with the degradation data saved in a characteristic map. [7,8,10] By using a more sophisticated model like an electrochemical pseudo-two-dimensional model and using different ageing improvements, the extrapolation ability and, therefore, the prediction ability is improved for unseen data. In addition, the model could be able to diagnose the ageing process. However, due to their complexity and the involvement of many parameters and multiple partial differential equations, they are unsuitable for online usage. Therefore the research focuses on simplifications and parameterisation. [7,8] Model-based approaches that incorporate models are, in most cases, observers and filters. The model is used to predict the measurement, where the model's state consists of its parameters and not its SOC or overpotential. [6,6,9,31,39] In the group of data-driven approaches, some parts of the direct models can be used since they rely highly on the data or machine learning approaches like the support vector machine or the neural networks. These algorithms can be used to directly estimate the SOH based on measurements like EIS or charge profiles that are measured during the application.[3,7,8,44] A much more specialised approach for sequential data is the recursive neural network (RNN). One central part of an RNN is parameter sharing, meaning that parameters are shared between multiple parts of the model. The second central part is that the network has connections to results in the past. Therefore, the present values of a variable can influence future values of its own. Different approaches are considered to be RNNs. [45] These models are often used for RUL prediction and capacity degradation prediction in general.

For this work, the decision falls to the direct model-based approach, where an ageing model is designed based on the measurements made and a cycle counter to access the model. This approach is common and sometimes extended with a filter or observer. The following sections describe the processing of the raw ageing data to combined data sets and, from there, the ageing model and its usage in combination with the cycle counter.

5.1. Aging data processing

The fitting of the relaxations of each check-up leads to the identified dynamic parameters for various SOCs at different ageing states of each cell. For every stress factor combination, at least two cells were used. For every check-up, parameters for charging and discharging direction are collected from the HPPC parts' relaxations during the discharging and charging measurement of the incremental OCV. These data sets of each cell are combined as if it is one cell's data. At first, the charging and discharging values of the parameters are combined and sampled to a fixed grid of SOCs. This is conducted in two steps, at first, the data is fitted using a polynomial of the fifth degree, and then the model is sampled with the original SOCs to be interpolated/extrapolated to the fixed grid of SOCs. The fit of a polynomial is useful because the SOCs of the parameters in the charging and discharging direction differ, so by fitting a polynomial instead of calculating the mean, the step of bringing both to a fixed grid at this point is skipped, and a kind of trend line is estimated. The measurements represent the parameters for charging and discharging directions, and the colour changes with the SOC. However, the polynomial of a relatively high degree compared to the number of samples leads to problems resembling the behaviour outside the given SOCs. By combining the fit and the extrapolation, the estimate outside the given SOCs is more conservative. Another solution to the extrapolation can be the clipping of the characteristic map. An example of the result of this process is displayed in figure 8.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

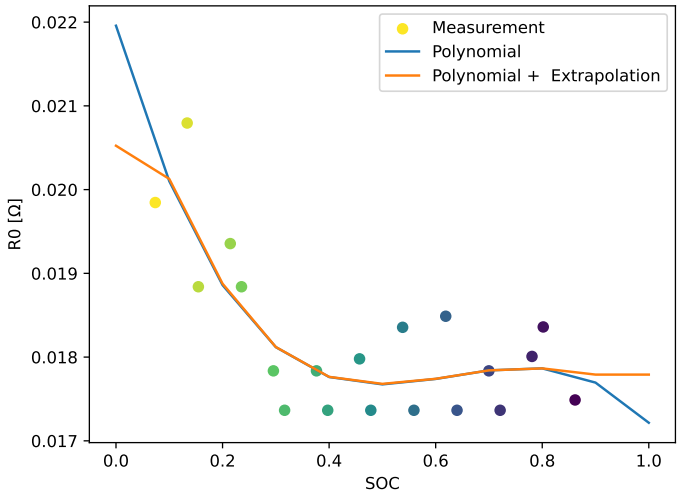

Figure 8. Result of the combination of charging and discharging parameter values using fitting and extrapolation. Different colours of measurements represent the SOC the parameter is estimated at. The points belong to parameters estimated in charging and discharging direction.

A similar approach combines the capacities of the cells aged under the same stress factors over the full cycle equivalents (FCE). After combining the parameters for charging and discharging direction, the parameters can be displayed as a 3D map over SOC and SOH. The following steps combine the maps of the different cells aged with the same stress factors, resulting in maps with a fixed SOC and SOH, which makes it easier for later usage. During this process, the maps of the group of battery cells are fitted using the lmfit package by defining a multidimensional polynomial. As before, this process eliminates the step of bringing all maps to the same fixed SOH values (see figure 9).

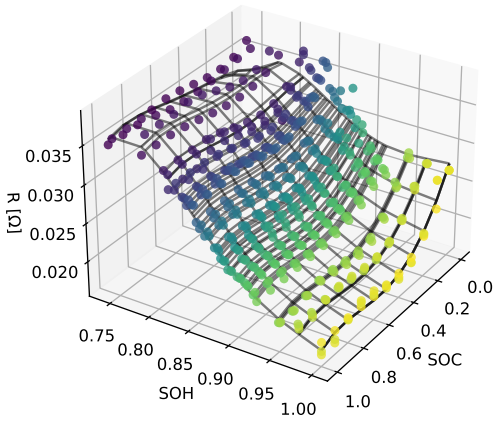

Figure 9. Result of the combination of different cells using a multidimensional polynomial fit.

At this point, the data of the cells that share the same stress factors are combined into a single data set. The following steps include the design of the ageing model based on the data and how the degradation-based change is combined with the battery model.

5.2. Aging model

The ageing model, in general, is kept very simple. A simple way to implement a model is to create characteristic maps and access these maps with the extracted stress factors to get the parameters' change. The previously preprocessed data is further processed to get to the characteristic maps. For the characteristic map of the capacity, the data of the capacities are normalised to the first value of each dataset for each stress factor combination so that the graph corresponds to the SOH over the FCE as shown in figure 10.

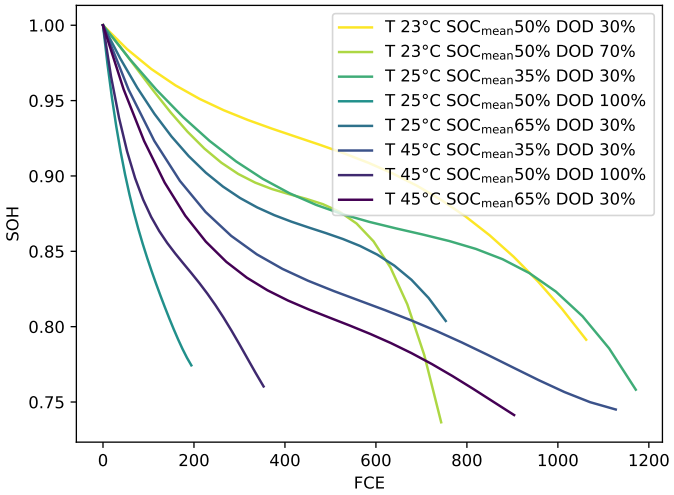

Figure 10. Resulting characteristic map for the capacities.

A similar approach is used for the other parameters, where at each ageing state, the mean is calculated and normalised to its initial value. The parameters, consisting of RC resistances and time constants τ instead of the capacities, besides the battery cells' internal resistance, do not follow a specific trend. This is why the other parameters' change is often neglected, and the modelling focuses on changing the internal resistance and the battery cell capacity.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

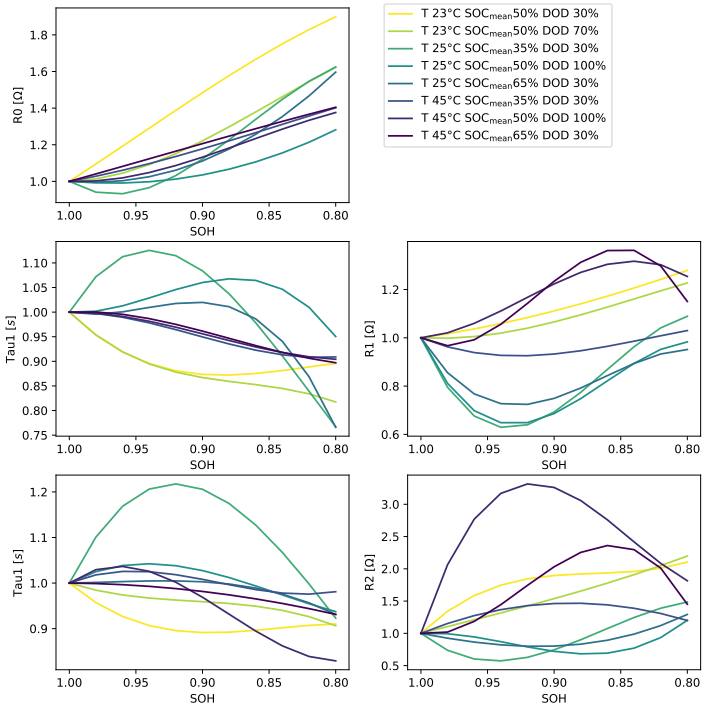

Figure 11. Resulting characteristic map for the dynamic parameters.

These characteristic maps are used to calculate the parameters' change during degradation for the aggregated stress factors. These maps are combined with a simple method to extract the stress factors during runtime to calculate the change and add it to the change of the specific parameter. The feature extraction is done by calculating the minimum and maximum SOC for the DOD, and the mean temperature and by summing up the SOC and calculating the mean for a specified window based on the calculation of FCE. Wrapping everything up for the ageing model consists of the continuous calculation of the FCE and stress factors that are fed into the characteristic map of the SOH over the FCE to get the change of SOH and the current SOH as follows

$$\Delta SOH = SOH(SOC_{mean}, T, DOD, FCE[k]) - SOH(SOC_{mean}, T, DOD, FCE[k-1]) \quad (39)$$

$$SOH[k] = SOH[k-1] + \Delta SOH. \quad (40)$$

The calculated SOH[k] and the SOH[k-1] of the last update are used to get the change of the dynamic parameters P

$$\Delta P = P(SOC_{mean}, T, DOD, SOH[k]) - P(SOC_{mean}, T, DOD, SOH[k-1]) \quad (41)$$

$$P[k] = P[k-1] + \Delta P. \quad (42)$$

By normalising the maps and calculating the change in the parameters, the battery model can be initialised with parameter maps that describe the behaviour of the parameters for different temperatures and SOCs. These initial maps are not normalised for the basic parameterisation. Therefore, the degradation of the parameters is included as a factor of the initial maps.

 doi:10.20944/preprints202305.0621.v1

6. System Overview

This section summarises the different pieces used to predict the SOC and the SOH. The system starts with the measurements, where the current is used to predict the states of the system by using the battery model and the voltage to update the states inside the UKF. The battery model depends on the different parameters of the dynamic model; these parameters depend on the SOH, the temperature and the SOC. Therefore, to get to the SOH, the measurements and the SOC are analysed to get the features that describe the stress factors during ageing in the feature extractor. It calculates the full cycle equivalents based on the current. The condensed stress factors are forwarded into the ageing model, which works on one side as a parameter model, describing the SOC and temperature-dependent behaviour. On the other hand, the ageing model describes the SOH based on stress factors. Updating the parameters is done in every iteration for SOC and temperature dependencies and when it is issued by the feature extractor for the SOH dependence. By running iteratively through the measurements of the SOC, the features and the SOH are estimated for every sample. Figure 12 displays a schematic of the system.

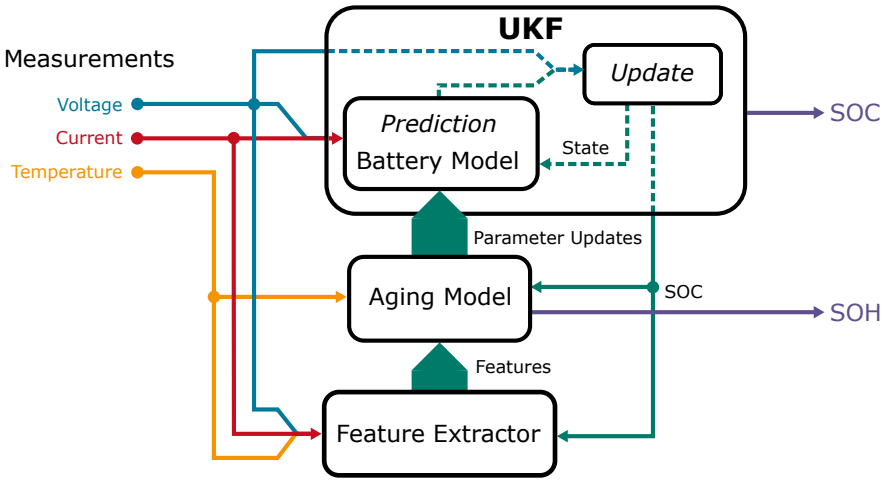

Figure 12. Schematic overview of the system depicting the UKF, the ageing model and the feature extractor.

7. Results and Discussion

The validation of the whole system consisting of implemented models and algorithms is realised by iterating the data of the dynamically cycled cells. That means starting with the first check-up and alternating through check-up and cycling data. To do so without taking too much time, the data was resampled with a sampling rate of 1 s and the updating of the feature extractor was set to every two full cycle equivalents. The extraction result is displayed in 13.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

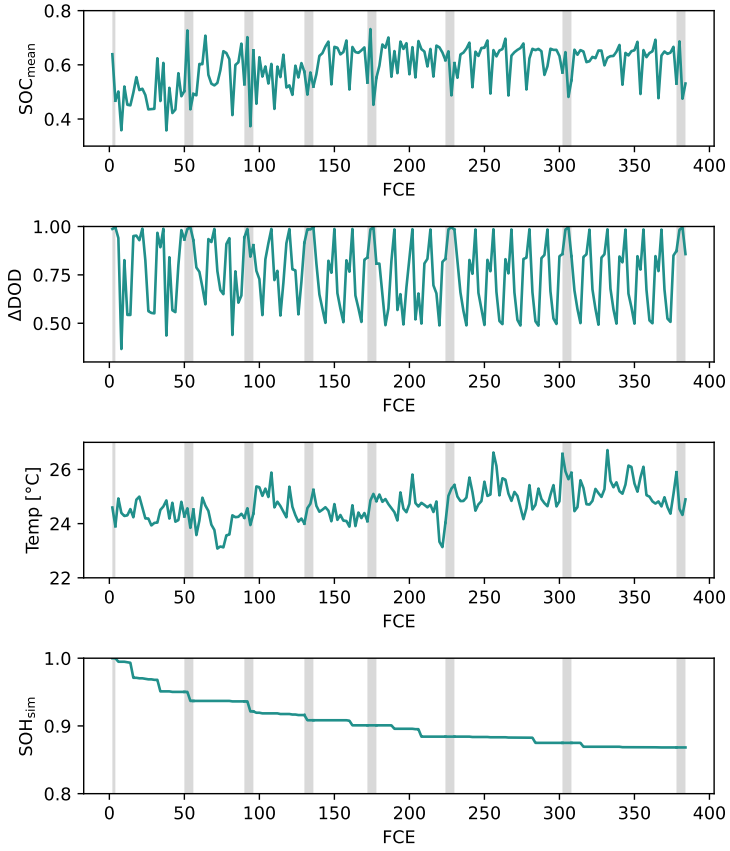

Figure 13. Features extracted over laboratory battery life. Check-ups are marked with an underlying grey area.

The figure displays the degradation process of the battery cell. Check-ups are marked with an underlying grey area. During the degradation, the mean of the SOC is around 50 % and transitions to about 65 %. Whereas the ΔDOD switches between 50 %, 80 % and full cycles. Smaller cycles are included, but since the feature extractor updates every two full cycle, equivalents the system is experiencing in this duration of bigger cycles, the small cycles are neglected. The temperature is in general, about 25 °C and increases minimally over degradation. During the degradation process, the SOH estimation error increases over time. That is an expected behaviour since the ageing model does not represent dynamic ageing, and the feature extractor does not capture the stress factors of smaller stress cycles. The absolute error in the SOH rises to 1.2 % at the end. The 1C capacity measured during the check-ups is used as a reference for the SOH. The evolution of the SOH in comparison to the reference and its absolute error is displayed in 14.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

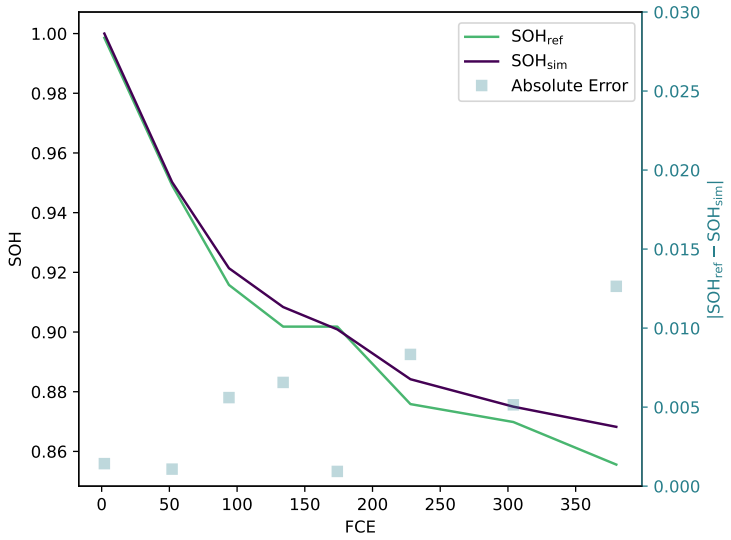

Figure 14. Comparison of the estimated SOH and the reference SOH

The error of the ageing model is included in the SOC estimation since the model for the internal state prediction of UKF uses the capacity to calculate the current state vector, including the SOC. This circumstance leads to an increasing error in the SOC estimation of the degradation of the battery cell. Therefore, the SOC and SOH errors correlate as the Pearson correlation coefficient confirms with 0.76 (see figure 15). The RMSE of the SOC is calculated for the dynamic sections of check-ups because a SOC reference based on the measured capacity of the corresponding check-up can be used.

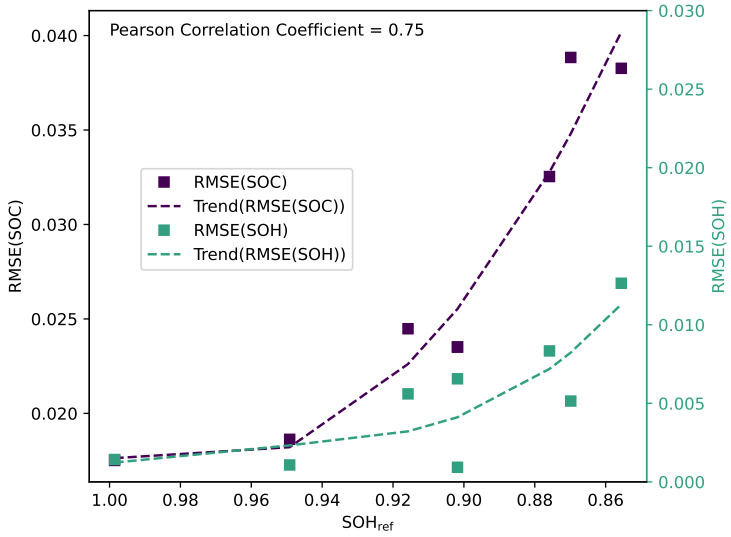

Figure 15. RMSE of the estimated SOC and SOH with their trendline and Pearson correlation coefficient

Interpreting the diagram leads to the conclusion that the error will rise faster with further degradation. It may then reach the limit of the implemented UKF. Nevertheless, the error of the SOC estimation is minimal, with about 4 % at its maximum. A regression plot is generated to analyse the performance of the SOC estimation. The regression plot displays the performance of the estimated SOC (SOC_{sim}) in contrast to the reference SOC (see figure 16). When a marker is above the diagonal line, the SOC is overestimated. Otherwise, it is underestimated. For visibility reasons, the SOC was resampled, only plotting every 300th estimate.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2023 doi:10.20944/preprints202305.0621.v1

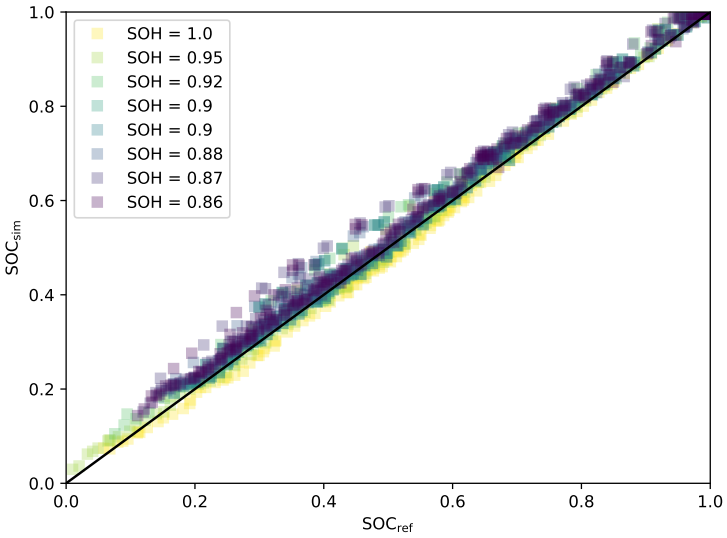

Figure 16. Regression plot of the SOC for different ageing states

The diagram shows that over ageing, the SOC estimation is increasing, but especially in the range of the 20 % to 60 % SOC, the estimation has more outliers. Outliers might result from the change of the OCV over ageing, which is not part of the model because it was neglected during the system development. In addition to more outliers, estimation overestimates the SOC with increasing degradation. The overestimation results from underestimating the degradation, leading to a higher SOH than the reference. Since the measurement segments start fully charged, the cell discharges slower than the reference suggests resulting in the overestimation of the SOC.

Compared to other algorithms used in the literature, the system performs comparably to other implementations. In general, the RMSE lies in the range of about 1 % to 2 % SOC in literature. The paper of Yang et al. [46] includes a survey for multiple approaches. However, many approaches are tested in relatively simple circumstances where the cell is cycled with constant currents or schedules like the Urban Dynamic Driving Schedule that consists of multiple constant current pulses. There is only sparse literature where the SOC estimation is applied to the whole battery cell life and a cell life with dynamic profiles.

8. Conclusions

Overall the paper includes all the steps to set up a general common approach for SOC and SOH estimation with a few additions like another discretisation of the model, an enhanced OCV sampling method, a specific parameterisation of the relaxation of the pulses and a simple feature extraction. In contrast to the literature, this paper includes all the steps from measurements to modelling, parameterisation, ageing modelling and bringing it together with the feature extraction to run through a whole battery degradation process. The degradation consists of dynamic cyclic ageing and the check-ups in between, leading to the opportunity to analyse the influence of the SOH estimation on the SOC estimation and the performance of the approaches considering dynamic load profiles.

Author Contributions: S. Neupert contributed to the tasks of conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, visualization; J. Kowal contributed to supervision and reviewing; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The data set is available under https://depositonce.tu-berlin.de/handle/11303/18795 or you might navigate to https://depositonce.tu-berlin.de and search for "Lithium-Ion Battery Cell Degradation Data Set".

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BMS	Battery Management System
CCCV	Constant Current Constant Voltage
DOD	Depth of Discharge
DRT	Distribution of Relaxation Times
ECM	Equivalent Circuit Model
EIS	Electrochemical Impedance Spectroscopy
EKF	Extended Kalman Filter
FCE	Full Cycle Equivalent
FUDS	Federal Urban Driving Schedule
HPPC	Hybrid Pulse Power Characterisation
ICA	Incremental Capacity Analysis
IR	Internal Resistance
OCV	Open Circuit Voltage
RMSE	Root Mean Square Error
RNN	Recursive Neural Network
RUL	Remaining Useful Life
SOC	State of Charge
SOH	State of Health
SOF	State of Function
SPKF	Sigma Point Kalman Filter
UKF	Unscented Kalman Filter

References

1. Park, S.; et. al.; Review of state-of-the-art battery state estimation technologies for battery management systems of stationary energy storage systems *JPE* **2020**, *20*, 1526 - 1540.
2. Ungurean, L.; et. al.; Battery state of health estimation: a structured review of models, methods and commercial devices *Int. J. Energy Res.* **2017**, *41*, 151–181.
3. Ali, M. U.; et al.; Towards a Smarter Battery Management System for Electric Vehicle Applications: A Critical Review of Lithium-Ion Battery State of Charge Estimation *Energies* **2019**, *12*, 446.
4. Kirchev, A.; et al.; Battery Management and Battery Diagnostics *Electrochemical Energy Storage for Renewable Sources and Grid Balancing: Elsevier* **2015**, 411–435.
5. Vezzini, A.; et al.; Lithium-Ion Battery Management *Lithium-Ion Batteries: Elsevier* **2014**, 345 - 360.
6. Wang, Y.; et al.; A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems.*Renewable and Sustainable Energy Reviews* **2020**, *131*, 110015.
7. Hu, X.; et al.; State estimation for advanced battery management: Key challenges and future trends *Renewable and Sustainable Energy Reviews* **2019**, *114*, 109344.
8. Ge, M.-F.; et al.; A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries *Measurement* **2021**, *174*, 109057.
9. Berecibar, M.; et al.; Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renewable and Sustainable Energy Reviews* **2016**, *56*, 572–587.
10. Baghdadi, I.; et al.; Lithium battery aging model based on Dakin's degradation approach. *Journal of Power Sources* **2016**, *325*, 273–285.
11. Li, Y.; et al.; Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and Sustainable Energy Reviews* **2019**, 109254.
12. Petzl, M.; et al.; Advancements in OCV Measurement and Analysis for Lithium-Ion Batteries*IEEE Transactions on Energy Conversion* **2013**, *28*, 675.
13. Kellner, Q.; et al.; Battery cycle life test development for high-performance electric vehicle applications*Journal of Energy Storage* **2018**, *15*, 228.
14. Doyle, M.; et al.; Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Inserti Cell.*J. Electrochem. Soc.* **1993**, *140*, 1526–1533.

15. Richardson, M.; et al.; Generalised single particle models for high-rate operation of graded lithium-ion electrodes: Systematic derivation and validation.*Electrochimica Acta* **2020**, *339*, 135862.
16. Howey, D. A.; et al., UKACC 12th International Conference on Control (CONTROL), Piscataway, 5-7 Sept. 2018.
17. Li, J.; et al.; A Single Particle Model for Lithium-Ion Batteries with Electrolyte and Stress-Enhanced Diffusion Physics.*J. Electrochem. Soc.* **2017**, *164*, A874-A883.
18. Pang, H.; et al.; Parameter identification and systematic validation of an enhanced single-particle model with aging degradation physics for Li-ion batteries.*Electrochimica Acta* **2019**, *307*, 474–487.
19. Lotfi, N.; et al., American Control Conference, Seattle, USA, 24-26 May 2017.
20. Laue, V.; et al.; Practical identifiability of electrochemical P2D models for lithium-ion batteries.*Electrochimica Acta* **2021**, *51*, 1253–1265.
21. Moura, S. J.; et al.; Battery State Estimation for a Single Particle Model With Electrolyte Dynamics.*IEEE Trans. Contr. Syst. Technol.* **2017**, *25*, 453–468.
22. Moura, S. J.; et al.; Battery Adaptive Observer for a Single-Particle Model With Intercalation-Induced Stress.*IEEE Trans. Contr. Syst. Technol.* **2020**, *28*, 1363–1377.
23. Krewer, U.; et al.; Review—Dynamic Models of Li-Ion Batteries for Diagnosis and Operation: A Review and Perspective.*J. Electrochem. Soc.* **2018**, *165*, A3656-A3673.
24. Saidani, F.; et al.; Lithium-ion battery models: a comparative study and a model-based powerline communication.*Adv. Radio Sci.* **2017**, *15*, 83–91.
25. Lai, X.; et al.; A comparative study of different equivalent circuit models for estimating state-of-charge of lithium-ion batteries.*Electrochimica Acta* **2018**, *259*, 566–577.
26. Madani, S. S.; et al.; A Review of Different Electric Equivalent Circuit Models and Parameter Identification Methods of Lithium-Ion Batteries.*ECS Trans.* **2018**, *87*, 23–37.
27. Andrea, D., *Lithium-ion batteries and applications. A practical and comprehensive guide to lithium-ion batteries and arrays, from toys to towns*; Artech House: Boston, USA, 2020.
28. Wu, J.; et al., *SOC Estimation of Li-ion Battery by Adaptive Dual Kalman Filter under Typical Working Conditions.*; 2019 IEEE 3rd International Electrical and Energy Conference (CIEEC), Beijing, China, 7-9 Sept. 2019.
29. How, Dickson N. T.; et al.; A State of Charge Estimation for Lithium-Ion Batteries Using Model-Based and Data-Driven Methods: A Review*IEEE Access* **2019** *7* 136116-136136
30. Fleischer, C.; et al.; On-line adaptive battery impedance parameter and state estimation considering physical principles in reduced order equivalent circuit battery models part 2. Parameter and state estimation.*Journal of Power Sources* **2014**, *262*, 457–482.
31. Klee Barillas, J.; et al.; A comparative study and validation of state estimation algorithms for Li-ion batteries in battery management systems.*Applied Energy* **2015**, *155*, 455–462.
32. Sundaresan, S.; et al.; Tabular Open Circuit Voltage Modelling of Li-Ion Batteries for Robust SOC Estimation*Energies* **2022**, *15*.
33. Lavigne, L.; et al.; Lithium-ion Open Circuit Voltage (OCV) curve modelling and its ageing adjustment*Journal of Power Sources* **2016**, *324* 694.
34. Yu, Q.-Q.; et al.; A Comparative Study on Open Circuit Voltage Models for Lithium-ion Batteries*CHINESE JOURNAL OF MECHANICAL ENGINEERING* **2018**, *31*.
35. Pillai, P.; et al.; Open-Circuit Voltage Models for Battery Management Systems: A Review*Energies* **2022**, *15* 6803.
36. Narula, M.; Curve Curvature in Python*url: https://www.delftstack.com/howto/numpy/curvature-formula-numpy/* **2022**.
37. Wan, T. H.; et al.; Influence of the discretization methods on the distribution of relaxation times deconvolution: implementing radial basis functions with DRTtools.*Electrochimica Acta* **2015**, *184* 483-499 .
38. Newville, M.; et al.; lmfit/lmfit-py: 1.1.0*Zenodo* **2022**, *version: 1.1.0* .
39. Welch, B; An Introduction to Kalman Filter*University of North Carolina* **2006**
40. Wang, Zuolu; et al.; A review on online state of charge and state of health estimation for lithium-ion batteries in electric vehicles*Energy Reports* **2021** *7* 5141-5161
41. Ren, Hongbin; et al.; A comparative study of lumped equivalent circuit models of a lithium battery for state of charge prediction*International Journal of Energy Research* **2019**
42. Brown, R. G.; et al.; Introduction to Random Signals and Applied Kalman Filtering: with MATLAB® Exercises*John Wiley and Sons, Inc.* **2012**
43. Haykin, S.; et al.; Kalman Filtering and Neural Networks*John Wiley and Sons, Inc.* **2001**

 doi:10.20944/preprints202305.0621.v1

44. Li, Yang; et al.; A physics-based distributed-parameter equivalent circuit model for lithium-ion batteries*Electrochimica Acta* **2019** *299* 451-469
45. Goodfellow, Ian; et al.; Deep Learning **2016**
46. Yang, Bo; et al.; Classification, summarization and perspectives on state-of-charge estimation of lithium-ion batteries used in electric vehicles: A critical comprehensive survey*Journal of Energy Storage* **2021**