EVS30 Symposium
Stuttgart, Germany, October 9 - 11, 2017

Battery Health Monitoring and Degradation Prognosis in Fleet Management Systems

A. Nuhic[1,*], J. Bergdolt[1], B. Spier[1], M. Buchholz[2], K. Dietmayer[2]
[1] *Deutsche ACCUMOTIVE GmbH & Co. KG, Neue Str. 95, D-73230 Kirchheim u. Teck*
[2] *Institute of Measurement, Control, and Microtechnology, Ulm University, D-89069 Ulm*
[*] *Phone: +49 151 586 240 85, E-Mail: adnan.nuhic@daimler.com*

Abstract

Fleet management systems with battery health monitoring capabilities moved now in the foreground more than ever. This paper addresses the development of a novel battery health monitoring algorithm with a degradation prognosis feasibility particularly adapted for usage in fleet management systems. Moreover, the chosen degradation prognosis approach adapts itself continuously on varying environmental conditions or utilization modes by identifying the impact factors which lead to a certain degradation trend. Such findings, when accessible with a fleet management system, offer various possibilities for fleet analysis techniques e.g. to identify an imminent battery failure.

Keywords: lithium battery, battery SoH (State of Health), prediction, data acquisition, fleet

1 Introduction

Original equipment manufacturers (OEMs) are concerned with performance losses as well as battery health and life-time restrictions when bringing the lithium-ion battery technology to the customers. Power and capacity fade are two primary metrics of lithium-ion battery degradation impacting directly the customer range and confidence, especially the latter one. However, battery fade is a complex phenomenon provoked by environmental conditions and utilization modes. In an attempt to capture the real-world utilization impacts on battery degradation, OEMs perform at the first step of an automotive development cycle for lithium-ion batteries as shown in Figure 1, ageing tests on single cells by applying realistic duty cycles on them. However, this duty cycles are vague approximations and do not account for the daily variation and unpredictability of such a system [1]. One way to deal with this uncertainty in life-cycle assessment is to oversize the battery and to narrow its operating window to ensure that the vehicle meets minimum requirements like acceleration possibility and available range at the end of its life-time.

This waste of performance and also expenses can be avoided and still a life-time and cost optimized battery can be developed. There exist many studies that aim to optimize the future (H)EV battery operational utilization to reduce costs and identify preventive maintenance activities. The keyword to achieve this task is an active fleet management – a personalized health monitoring of every battery build in a hybrid or electric vehicle by telematics devices [2], as the final step in the product development cycle of Figure 1. Such a system is usually set up on a telematics data server containing a large amount of raw data packets collected from the vehicles of a fleet, during diagnostic sessions in dealer's workshops, or via telematics devices over-the-air [3]. Therefore, this data is filtered and analysed in various aspects, e.g. to identify an imminent failure of a component. This approach adopted for a lithium-ion battery requires an accurate estimation of degradation parameters and a prognosis for their future deployment on-board of each vehicle.

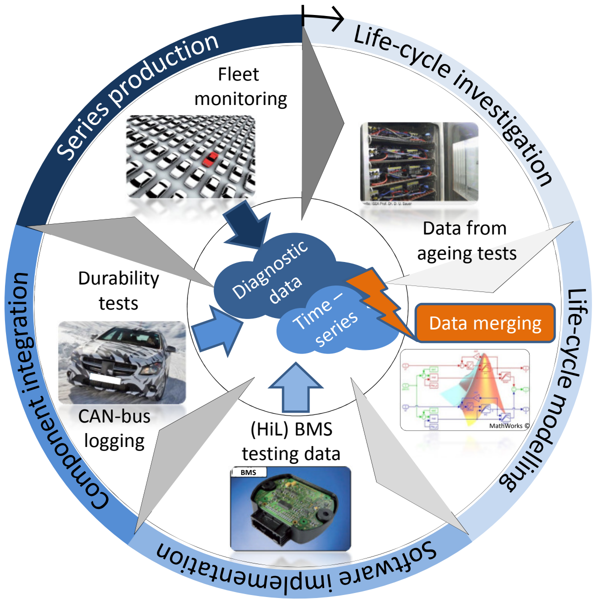

Figure 1: Development cycle for a lithium-ion battery in an automotive environment

The aim of this contribution is the development of a novel battery health monitoring algorithm with a degradation prognosis capability particularly adapted for usage in a fleet management system. Later on, the idea behind such a system is to identify a potential failure of each battery build in a vehicle based on that degradation prognosis data, so preventive maintenance activities could be started. Due to limited capability of transfer devices, the received raw data by a telematics server is limited in its amount - it's of diagnostic data type, as shown in Figure 1, with finite information contained. Therefore, an on-board health monitoring algorithm needs to calculate an individual degradation prognosis model for each battery and provides a limited model parameter set via diagnostic devices to the fleet management system. Furthermore, as battery degradation is unique for each vehicle, because of various environmental conditions and application modes, the real-world utilization needs to be captured in compact form and provided in addition to the degradation model parameters.

2 Life-cycle investigation on cell and battery level

The typical life-cycle investigation on lithium-ion cells usually starts with a large scale of cell degradation tests with a pre-defined duty-cycle. For this work, twelve high-power lithium-ion cells for automotive application in hybrid vehicles were used for the experimental investigation on cell level. Six of them experienced a cyclically repeated real-world current profile at room temperature, with a cycle start at different state of charge (SoC) levels for each pair of three cells. The resulting capacity loss occurred during cycling can be observed in the left graph of Figure 2. By considering these results, the following questions arises immediately - how can one be assured that the tested duty-cycle is representative of real-world loads. To counteract this unilateral life-cycle investigation result, further six cells of the same type with different ageing stages were exposed to much more varying real-world current profiles, temperatures and cycle-levels of SoC which occur at the third and fourth stage of the development cycle in Figure 1. The full description of the performed testing procedure can be found in [4]. Now, the observed capacity loss, shown in the right graph of Figure 2, shows a more differentiated capacity fade over time for each cell depending on their ageing stage at the begin of the test, and the intensity of the actual duty cycle at a certain point of time.

As this work introduces a novel battery health monitoring and prognosis algorithm to track the battery degradation for each vehicle in a fleet, a certain type of data, available only by a fleet monitoring system, will be also required for validation purposes. As the world's first mass produced car with a lithium-ion battery [5], the Mercedes-Benz S400 Hybrid – the battery itself is shown left in Figure 3 – is already since eight years in customer hands, one could think that enough fleet data for validation purposes should exist. In fact, that is true, but the existing diagnostic data does not contain relevant information for the developed algorithm to work with properly and needs to be created from the scratch again.

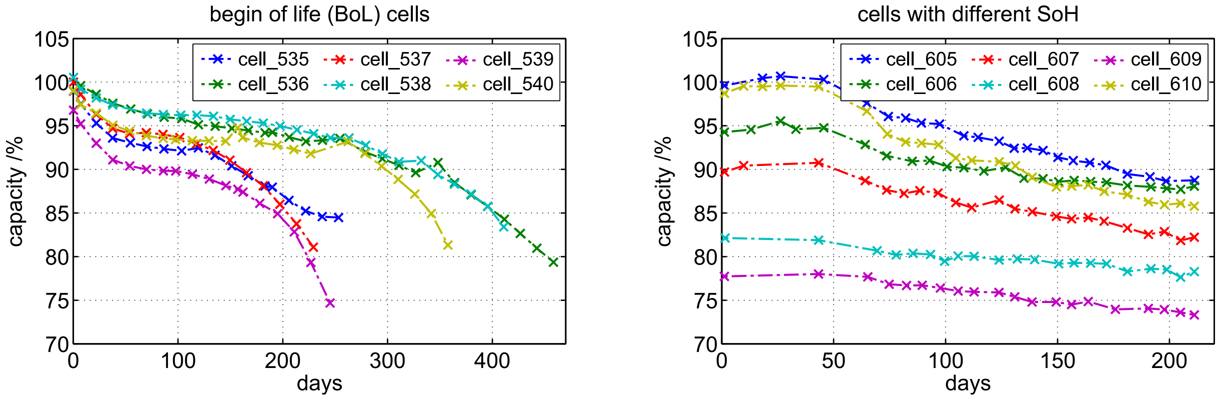

Figure 2: Capacity loss during cycling with a single (left) and various real-world duty cycles on the right graph

For that, recorded real-world duty cycles were disposed to lithium-ion batteries in laboratory conditions over a long period of time as shown on the right side of Figure 3. The batteries experienced identical conditions as if they were integrated in vehicles in terms of ambient temperature and battery cooling. As it is not possible to cyclically investigate enough batteries in laboratory conditions to represent a whole fleet, single cells of each battery will represent single (H)EV vehicles in the further course of this work. In this way, a small „fleet“ of seventy „vehicles“ is available for validation purposes. How the measured time-transient data is converted to a data-type coming from a fleet management system will be shown in the next section.

Figure 3: The Mercedes-Benz S400 Hybrid battery (left) and the same battery type in a durability test on the right

3 Modelling the ageing behaviour of lithium-ion cells

A key step of an automotive development cycle for lithium-ion batteries is the development of an accurate ageing model. Such a model formulates the degradation process as a function of battery operations [6] and is therefore indispensable in the early product development phase. When it is built upon real-world battery operation data, an accurate ageing model performs estimations about degradation behaviour of the utilized lithium-ion cell type, to ensure its life expectancy as a battery component. However, an ageing model – depending on the modelling approach – can also be an important element of the battery management system (BMS). By being part of a continuous health monitoring process, it observes the vital battery parameters like the actual capacity and performs prognosis for their deployment. Furthermore, such a health monitoring process produces also an output relevant for a fleet management system when accessible e.g. during dealer's workshops, or via telematics devices over-the-air. This is only possible if the process output data is of diagnostic type.

The decision to implement an ageing model into a BMS for degradation prognosis purposes depends, beside the chosen modelling approach, mostly on the accuracy of such a model. By being modelled only on life-cycle investigation data, it is hardly to believe that such a model will perform well in all possible operation conditions over lifetime. Therefore, all real-world battery operation data which come up during the software implementation and testing phase of the BMS as well as data from durability tests with whole batteries, should be considered in the modelling of the ageing behaviour. The basics for this approach are introduced in the next subsections.

3.1 Data preprocessing by histograms - load collectives

By considering the different data types generated during the whole battery development cycle from Figure 1, the question comes up how the relevant information for the modelling of the ageing behaviour contained in the data can be merged together. Time-transient data from laboratory investigations on cells and from CAN-Bus data-loggers in (B)EV vehicles need an uniform consideration as well as the diagnostic data received by a fleet management system. An appropriate way to achieve this is to convert time-transient battery operation data into histograms by following clearly defined rules. As this idea is not new, it was already published in [4] and [7], the present work introduces a further modification of it.

Building histograms of time-transient data results in a massive data reduction and, therefore, loss of relevant information. The idea behind using histograms is to put the information contained in the resulting distributions in relationship with the observed capacity fade. The capacity is a relatively slow changing internal battery parameter, calculated e.g. by coulomb-counting and only at certain battery operation events like the charging procedure. This results in relatively long lasting time-series in the interval between two proper calculated capacity values that need to be transformed into histograms. These time-series can contain several billions of data-points, so the classical rules for defining the number of histogram bins would fail resulting in thousands of bins as these rules are dependent on the data length. The interesting histogram central moments like mean, variance and standard deviation can be calculated by a moving window approach directly on times-series data. But this reduces the large amount of data-points only to three relevant values which is to much data reduction. Therefore, a time-series of battery relevant signals like the current, cell-voltages, temperature or SoC measured in a time-interval between two calculated capacity values needs to be separated into segments first before converting them into histograms – or so called load collectives in our case.

Although there exist some really interesting Bayesian approaches for finding change-points in time series for their segmentation [8], we prefer a more „controllable“ approach with the well known rainflow-counting algorithm from machinery. In [4], an application of this algorithm for counting the depth-of-discharges (DoD) in the SoC-signal was shown, in which the found cycles were classified in their mean and amplitude afterwards. Later on, an improvement to this cycle counting was introduced in [7] by explaining each found DoD-cycle in a time-series by its corresponding charge- and discharge-rate. Furthermore, each found DoD-cycle has its corresponding mean cell-voltage and temperature which can also be classified by a histogram. At this point, we introduce our new approach to build the histograms from battery data, the so-called „nested“ load collective which is shown in Figure 4.

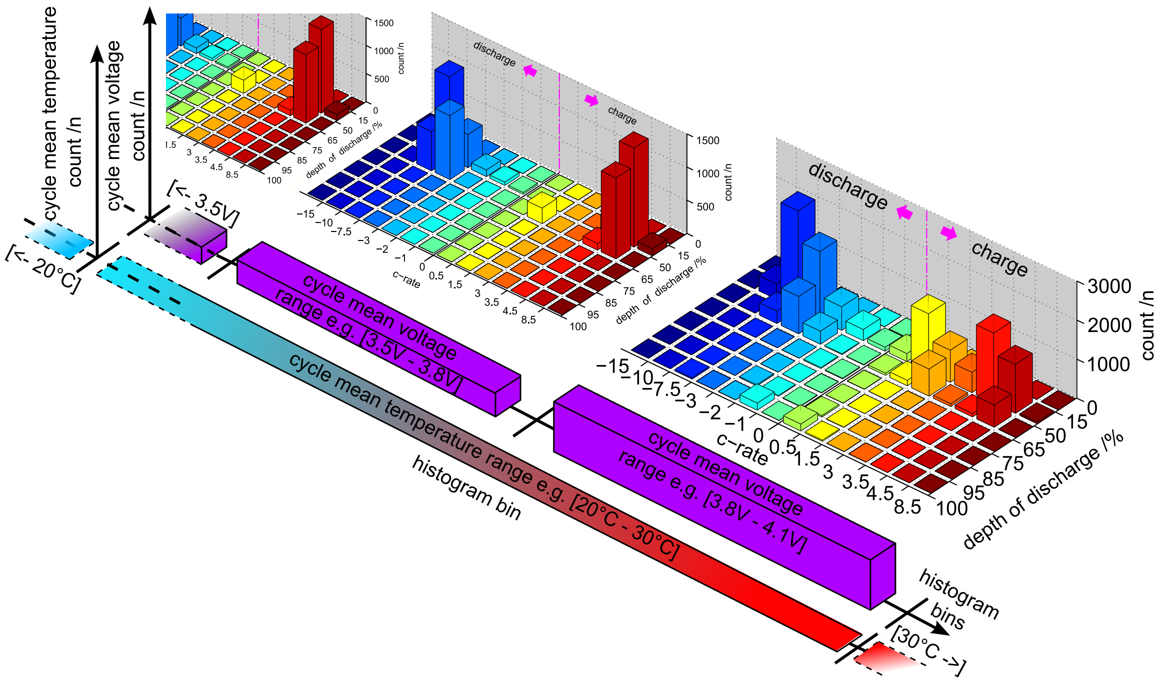

Figure 4: „Nested“ load collectives for battery data segmentation

First, every found DoD-cycle in the SoC-Signal of a lithium-ion battery is classified by its corresponding mean temperature of the cycle in a 1D-histogram. Every bin of this histogram, e.g. in the temperature range [20°C – 30°C], contains a nested 1D-histogram for mean voltage of the found DoD-cycles which are classified afterwards. Finally, every bin of the mean voltage histogram contains a 2D-histogram for classifying the found DoD-cycle by its DoD and the corresponding charge- and discharge-rate. Our expectation of this procedure is to preserve as much relevant information from the time-series data as possible for our model-based approach.

3.2 Data-based modelling

In principle, the modelling of the ageing behaviour in lithium-ion batteries can be classified into theoretical and empirical approaches. While theoretical models are based on the electrochemical physics of ageing mechanisms, like the loss of active lithium-ions and other active materials, empirical models are based on the underlying experimental data from battery operation. As this data is generated during the whole battery development cycle from Figure 1 to an increasing degree, an empirical modelling approach is preferred. Our empirical data-based model was already introduced in [4] and the present work introduces an probabilistic approach in addition to, for comparison purposes afterwards in the result section.

3.2.1 Support Vector Regression

Support vector regression (SVR) has been used in connection with load collectives as the basic data to health-monitoring purposes of lithium-ion batteries, e.g. in [4]. A detailed tutorial on this method can be found in [12], on which this paragraph is based on. Given training data $\{x_i, y_i\}_{i=1}^{N}$, the goal of support vector regression is to find a function $f(x)$ which has at most ε deviation from the values $\{y_i\}_{i=1}^{N}$ and is as flat as possible. First, assume $f(x)$ is linear, hence it has the form $f(x) = \langle \omega, x \rangle + b$. To ensure flatness, one has to minimize $||w||^2$. In addition, constraints have to make sure that the deviation between $f(x)$ and $\{y_i\}_{i=1}^{N}$ is not larger than ε. However, this is often not feasible. Hence, one has to allow for errors and introduce slack variables therefore, here denoted by ξ_i and ξ_i^*. This results in the following optimization problem:

$$\begin{aligned} &\min_{\omega} \frac{1}{2}\|\omega^2\| + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \\ &\text{subject to} \begin{cases} y_i - \|\omega, x_i\| - b \le \varepsilon + \xi_i \\ \|\omega, x_i\| + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0. \end{cases} \end{aligned} \tag{1}$$

The constant C describes the trade-off between the flatness of $f(x)$ and the errors. In many cases, the optimization problem in (1) can be solved more easily in its dual formulation. By solving it, one gets

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b, \tag{2}$$

where α_i and α_i^* stand for the Lagrange multipliers from the Lagrange function for the optimization problem. If either α_i or α_i^* is non-zero, x_i is called a support vector.

So far, $f(x)$ was assumed to be linear. The algorithm can be extended to the non-linear case by mapping x_i into some feature space using a non-linear map Φ and by applying support vector regression afterwards. Unfortunately, this might be computationally infeasible. However, (2) shows that $f(x)$ only depends on the dot product of x_i and x. Hence one only needs to know the dot product between the mappings of x_i and x, not the mapping itself. This dot product, $\langle \Phi(x_i), \Phi(x) \rangle$, corresponds to a kernel function $k(x_i, x)$. Hence $f(x)$ can be expressed as:

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) k(x_i, x) + b \tag{3}$$

The kernel function $k(x_i, x)$ has to satisfy certain conditions, which can be found in [12].

3.2.2 Relevance Vector Machine for regression

Relevance vector machines (RVM) have been used for forecasts related to lithium-ion batteries, for instance in [9] and [10]. This introduction follows them as well as [11]. In this work, a RVM is adopted for use with histogram – load collective data for the first time. Suppose we are given measurements adopted to use with histograms from the dependent variable $y = (y_1, ..., y_N)$ and from inputs $x = (x_1, ..., x_N)$. Their relationship is described by a model $f(x)$ and error ϵ. This can be expressed as

$$y_i = f(x_i) + \epsilon_i,\ i = 1, ..., N. \tag{4}$$

The errors ϵ_i are assumed to be independent, normally distributed random variables with mean zero and variance σ^2. Hence the joint probability density function of the errors is given by

$$p(\epsilon) = \prod_{i=1}^{N} \mathcal{N}\left(\epsilon_i | 0, \sigma^2\right). \tag{5}$$

The model $f(x_i)$ is a linear combination of basis functions $\phi_j(x_i)$. Hence, $f(x_i) = \sum_{j=1}^{M} w_j \phi_j(x_i) = \phi(x_i) w$, w being a vector of weights and $\phi(x_i) = (\phi_1(x_i), ..., \phi_M(x_i))$. In matrix notation, equation (4) can be written as

$$y = \Phi w + \epsilon, \tag{6}$$

where Φ is a matrix containing $\phi(x_i), i = 1, ..., N$. The aim is to find values for w such that $f(x)$ generalises well to new data and w is sparse, which means that most elements of w are equal to zero. Sparsity controls the complexity of the model and avoids overfitting. Furthermore, it is useful if memory resources are limited as in the case of a battery management system. The training vectors x_i, which are associated with non-zero elements of w, are called relevance vectors.

Using the normality assumption in (5), one can derive the likelihood function

$$p\left(y|w, \sigma^2\right) = (2\pi)^{-\frac{N}{2}} \sigma^{-N} exp\left(-\frac{||y - \Phi w||^2}{2\sigma^2}\right). \tag{7}$$

One approach might be to maximize the likelihood equation with respect to w. However, this can result in overfitting because the complexity of the model is not controlled. Controlling complexity means to control the growth of the weights in w. In a bayesian setting, this is done by using a prior distribution on w. It is specified by $p(w|\alpha) = \prod_{i=1}^{N} \mathcal{N}\left(w_i|0, \alpha_i^{-1}\right)$. $\alpha_i, i = 1, ..., N$ stand for hyperparameters which have their own prior distribution given by $p(\alpha) = \prod_{i=1}^{N} \Gamma(\alpha_i|a, b)$, where a and b stand for the parameters of the gamma distribution. In addition, there is a prior distribution for the inverse variance σ^{-2}. In accordance with the literature, we set $\beta := \sigma^{-2}$. Its prior distribution is specified as $p(\beta) = \Gamma(\beta|c, d)$, where c and d are the parameters of the gamma distribution. Using the likelihood, the prior distributions and Bayes' theorem one can calculate $p\left(w|y, \alpha, \sigma^2\right)$, as it is done in [9]. This distribution can be used to compute the predictive distribution $p\left(y_*|y, \alpha, \sigma^2\right)$ as in [11]. It is given by

$$p\left(y_*|y, \alpha, \sigma^2\right) = \int p\left(y_*|w, \sigma^2\right) p\left(w|y, \alpha, \sigma^2\right) \mathrm{d}w. \tag{8}$$

This results in a normal distribution. Its mean can be seen as the predicted value and its variance describes the uncertainty of the prediction. A more detailed explanation, especially on how to choose the values of α and σ^2 in (8), is given in [11].

3.3 Incremental approach for on-board implementation

Batch versions of SVR and RVM, which are introduced in the two subsection above, have to be retrained from scratch whenever a new data point is added to the training set. This is not feasible for on-board applications, because the whole data set used for batch training would have to be stored on board. Hence an incremental approach is used. This incremental approach is based on the idea of [13] for support vector machines, which was extended to relevance vector machines in [10]. The procedure is the following: whenever a new data point comes in and the last prediction error was higher than a certain threshold, a new training set is created consisting of the relevance vectors and the new data point. The relevance vector machine is subsequently retrained using the new training set. With this approach, the data points from the former training set which were no relevance vectors are deleted. This leads to low memory requirements, because only the relevance vectors and the new data point have to be stored. However, one should make aware that potential future relevance vectors are also deleted. The procedure for support vector regression is similar, with the only difference being that support vectors are used instead of relevance vectors.

3.4 Parameter tuning and feature selection

For support vector regression, the hyper-parameters C, ε and a kernel parameter (in our case for a Gaussian kernel) have to be chosen. For the relevance vector machine, a parameter of the basis function has to be selected. Furthermore, a feature selection has to be conducted, too. Both tasks are done using repeated grid-search cross-validation for variable selection and parameter tuning as described in [14]. The feature selection method used in the repeated grid-search cross-validation is RRELIEFF from [15]. The selection of hyper-parameters and variables takes place in the batch case. They are not changed during incremental learning because of the computational restrictions of the battery management system.

3.5 Degradation prognosis approach

A simple degradation prognosis approach feasible for on-board implementation was already introduced in [4]. As the future battery operation is not known, one option is to multiply the already stored values in the load collectives by a factor to achieve a battery load extrapolation. This extrapolated load collectives serve as the input for an incremental SVM or RVM implemented on-board, to estimate the capacity value in the near feature. Due to limited computational capabilities on-board, more sophisticated approaches in modelling load variability especially for machinery [16] are reserved only for off-board use.

At this point, we introduce an utilization of these methods to extrapolate the described load collectives of a battery for the first time. By examining the 2D-histogram from Figure 4, it stands out that this discrete distribution has many empty bins in some regions. It is expected that many of these bins would have counts in them if the battery would operate for a longer time. An expected load collective for any desired number of counted DoD-cycles from Figure 4 could be constructed by randomly placing cycles into the histogram. Therefore, the discrete points of a histogram need to be converted into a continuous probability density. Each cycle is placed then into the histogram with the choice of bins weighted by the probability density for each bin so that the bins with higher probability densities have more cycles. Kernel estimators [16] provide a convenient way to estimate the probability density for extrapolation purposes. Such an extrapolation of the 2D-histogram from Figure 4, by employing kernel methods, is shown on the right side of Figure 5 below.

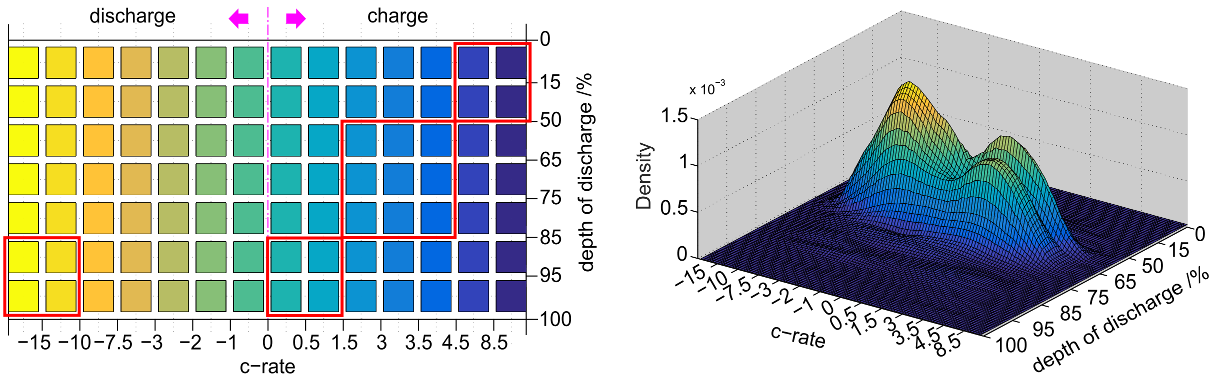

Figure 5: Defining „damage“ regions by red squares in the load collectives for kernel density estimation

Furthermore, the variability in histograms can be modelled individually by defining „damage“ regions in the load collective directly as shown in Figure 5 on the left. These regions can also have a physical background like it is well-known that high loading currents at low temperatures can have influence on the ageing due to lithium plating [5] or a high discharging current affects ageing, too. By choosing different bandwidths for the kernel in certain regions of the load collective, more variability can be achieved in them resulting in a histogram extrapolation that affects the ageing most.

4 Model implementation & usage by fleet management system

Finally, after explaining the basics for our battery degradation modelling and prognosis approach, the model implementation and usage by a fleet management system can be introduced in Figure 6. We distinguish between a model implementation on-board the BMS (arrows at the top) and a model usage off-board (bottom arrows), respectively on workstations with installed MATLAB© software. The single steps from the battery development cycle of Figure 1 are shown vertically from the left to the right of the picture. The initial degradation model is firstly build by a batch-modelling process based on experimental data from life-cycle investigations on single cells, as introduced in Figure 2. The conversion of time-series data from the laboratory is performed by load collectives to obtain the desired distributions of all battery relevant signals according to subsection 3.1. This is performed every time when new data is available in the following product development steps, e.g. during BMS testing or endurance tests of vehicles. Thus, it is obvious why the circumstance with building load collectives is indispensable. If diagnostic data from a certain fleet, which is accessible via diagnostic sessions in dealer's workshops or via telematic devices i.a. according to [3], contain the identical load collectives as they were used in batch-modelling process from the begin on, the developed degradation model can actively support the fleet management in health monitoring and prognosis.

Furthermore, the initial batch model is iteratively updated by incoming data during further battery development steps and gathers on this way additional „knowledge" about the different degradation behaviours from varying operating conditions, e.g. during endurance tests in hot or cold countries. At the end, when the series production starts, a final batch model is set up which is able to perform degradation prognosis by our approach from subsection 3.5 for each vehicle based on its set of load collectives or even a whole fleet by e.g. averaging their load collectives. Nevertheless, the batch-model updates do not stop here. As part of an active fleet management, the batch-model is updated iteratively by incoming load collectives which may contain further relevant information on battery degradation. The goal behind this procedure is to identify the ageing impacting factors continuously by performing an intensive feature selection on the whole data-set, reaching from the life-cycle investigation to the actual point in time. These information is very precious in the planning of life-cycle tests on cells for the next battery product generation.

In parallel to the batch-model, an incremental variant of the initial model can be implemented into the BMS during the software implementation phase together with the new „nested" load collectives. This type of model updates itself incrementally each time a new capacity value is calculated on-board, which happens often during the development phase e.g. by testing vehicles on the roller test bench. After that, an already updated incremental model is ready for series production where further adaptation to the individual degradation of a certain battery takes place. As described in subsection 3.5, this model performs a simple degradation prognosis for the capacity value on-board the vehicle which is a valuable input for an e.g. variable battery operation window.

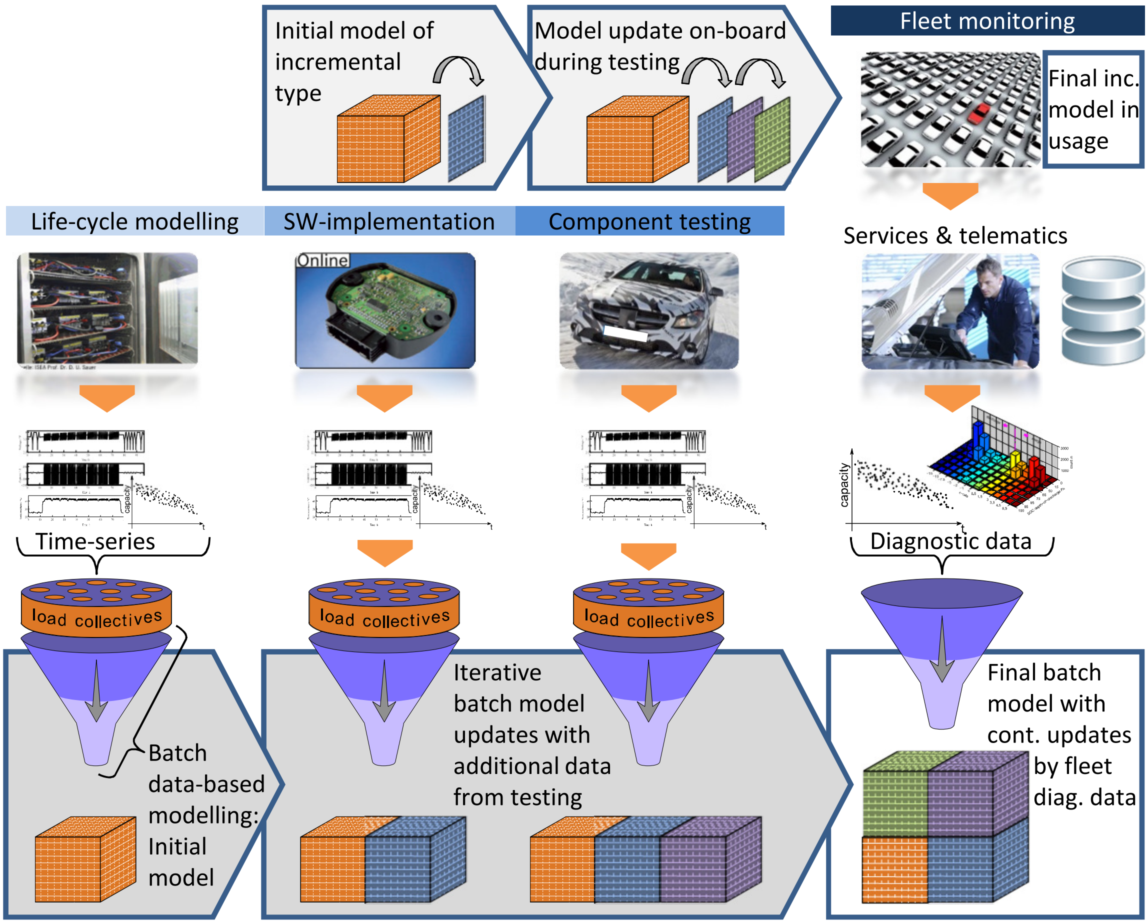

Figure 6: Model usage on-board and by a fleet management system respectively, for health monitoring and degradation prognosis of lithium-ion batteries

5 Results & discussion

In the past, a lot of battery health monitoring approaches suffered of the unavailability of applicable data for validation. Most of the accomplished investigations are based on very uniform tests, batteries are cycled only at one or a few values of SoC, temperature, and/or DoDs, so that the obtained approaches are at least not validated for an application in real-world dynamical situations. As already introduced, for our approach there exist a portfolio of measurements reaching from life-cycle investigations on lithium-ion cells in laboratory conditions to measurements from endurance tests on battery level which should guarantee a thoroughly validation of the developed models.

5.1 Batch model

The aim of the batch modelling process in case of a SVR approach is to find a function which is simultaneously as flat as possible and which approximates the capacity degradation trend most. By following the steps from Section 3 and starting with a data set of all available input-output pairs of a process, e.g. the ageing behaviour of cells in a life cycle investigation, this process results in the generation of an initial model from Figure 6 which contains only the support vectors. Additionally, the RVM approach results also in establishing an initial model by considering only the life cycle investigation data. This model contains the relevance vectors needed for the predictive distribution function from Equation (8). Figure 7 shows the results of adopting both approaches only on life cycle data (left graph) and on the same data set, but iteratively updated by measurements from vehicle durability testing on the right.

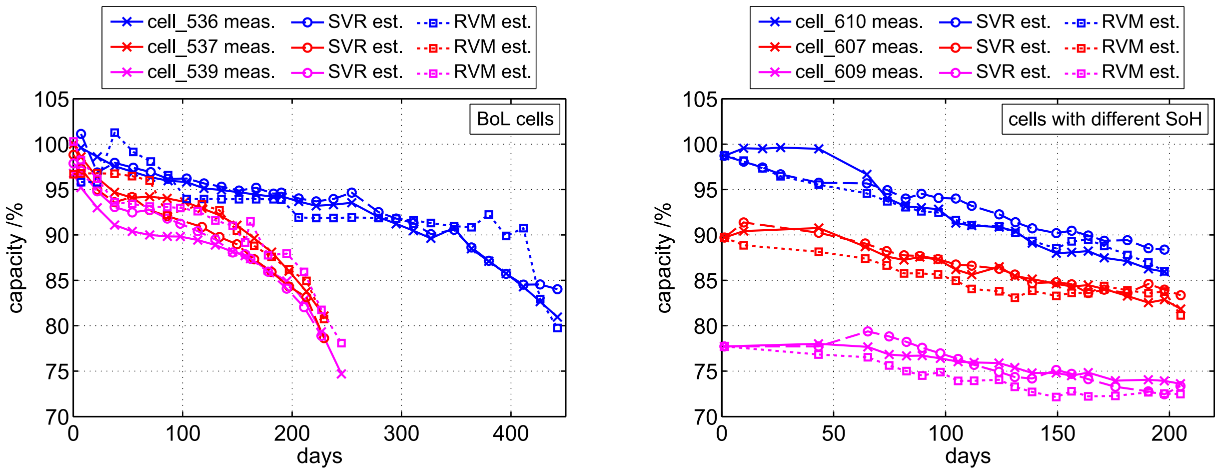

Figure 7: Batch model results validated on data only from life cycle measurements with BoL cells (left) and on a updated data-set by measurements from vehicle durability tests on the right graph

5.2 Incremental model

This work introduces also an incremental type of the data-based modelling approaches from Subsection 3.2, too. The initial batch models of SVR and RVM type from Figure 6 are both capable for an on-board implementation to the BMS which was already shown in [4].

If these initial batch models are only based on data from life cycle investigations without an on-board incremental update enhancement, the prognosis of the capacity degradation is not satisfactory when applied to data from vehicle durability tests as the Figure 8 presents. The reason for that could be a vague duty cycle used in the laboratory investigations which does not represent very well the variability of a real-world battery utilization. Therefore, the support- or relevance-vectors of the batch model, stored in the BMS, need an incremental update by new incoming input-output data pairs comprising always a set of load collectives as input and an actual capacity value as the output of a data point. In doing so, an on-board incrementally updated batch model, either of SVR or RVM type, is able to track the capacity degradation trend very precisely as Figure 9 illustrates. Furthermore, on this way updated data-based models are in the best position to perform an accurate prognosis of the capacity degradation at every point in time as they incorporate the „knowledge“ about the degradation behaviour from the life-cycle investigations on.

By considering the model results from Figure 9, it is obvious that both approaches, either the SVR or the RVM, perform well in terms of tracking the capacity degradation trend when finding new appropriate support or relevance vectors for the incremental model.

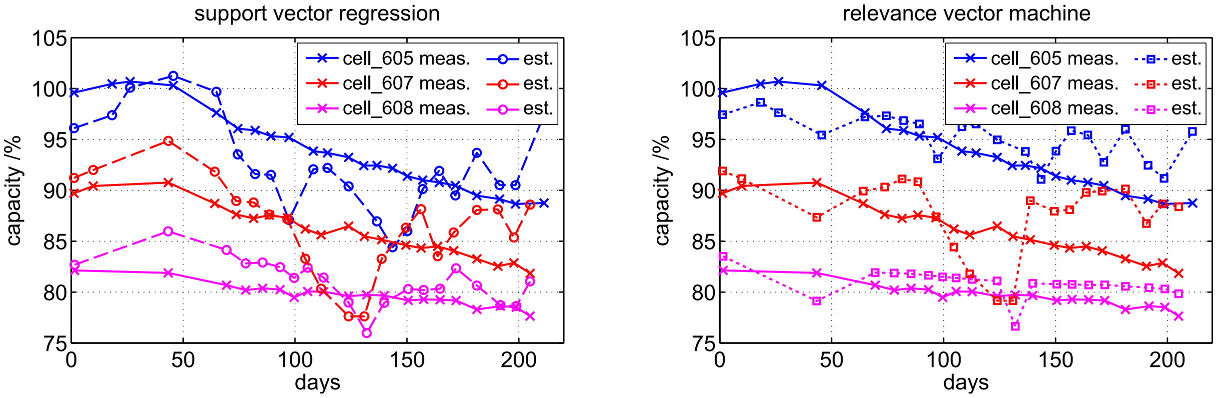

Figure 8: Results from a SVR and RVM initial batch model from Figure 6 without an incremental update enhancement on data from vehicle durability tests

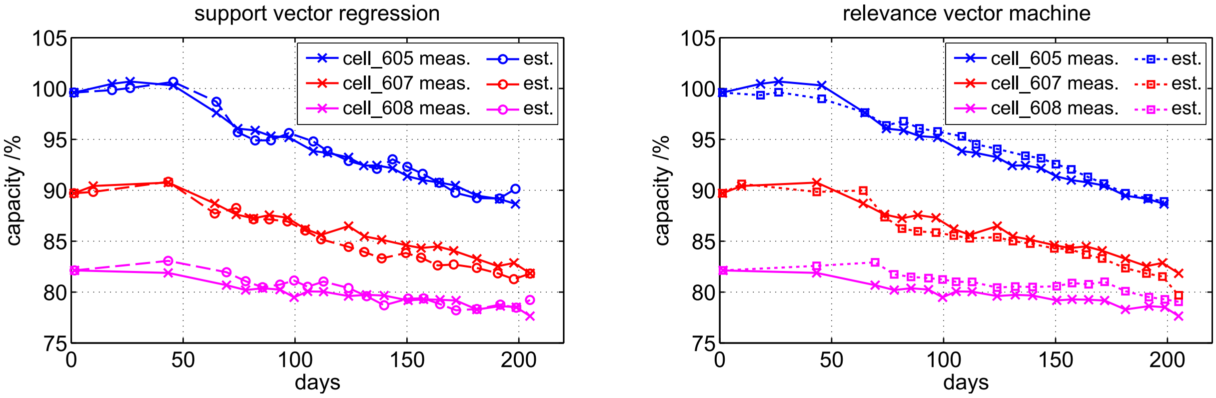

Figure 9: Incremental model results when the initial batch model is updated iteratively on the same data-set containing cells of different SoH

However, the RVM has one crucial advantage over the SVR – its model output provides also an probability distribution for the predicted capacity values. Later on, this confidence interval contains with 95% probability the true capacity value. Therefore, for further validation only this model type is taken into close consideration.

5.3 On-board and fleet management prognosis results

Finally, the prognosis capability of the on-board approach as well as of the fleet management system can be validated on battery data from durability tests of Figure 3. Until now, this data is unseen by the two model approaches, neither by the final RVM batch model from Figure 6 nor the incremental RVM model type, and therefore, represents a future real use-case together with the series production start. Due to clarity reasons, only two cells from a batch of 70 cells were chosen randomly for validation purposes. The prognosis results for the incremental model type are depicted on the left of the Figure 10.

By starting with a quite wrong estimated capacity for both „vehicles" it can be noticed that the incremental model adapts itself to the degradation trend more and more by hitting it after 100 days of battery utilization. At this point, a first capacity prognosis for the next slightly more than 50 days can be started by generating input data for the incremental RVM through repeating the very last available set of load collectives from the BMS memory multiple times. This results for both „vehicles" in a predicted capacity degradation, which is pretty close to the measured capacity trend. It is remarkable that, straight after the second incremental update, all the measured capacity values lie inside the confidence interval which should in fact contain the true capacity value with a 95% probability.

The prognosis results of the fleet management system are presented on the right of Figure 10. For the first 100 days, the incremental model update is performed on-board and this model, together with all load collectives of this period, are available for the fleet management system according to Figure 6. Now, the received load collectives can be extrapolated by adopting the kernel method from Subsection 3.5. For that, the load collective counts present on the hundredth day can be treated as a whole during extrapolation, resulting in a 100 days long prognosis window whereas a higher resolution for the capacity degradation trend would be preferred. Therefore, the load collective counts are divided into smaller „slices“ for extrapolation, resulting in the shown trend of Figure 10. Such a fast capacity degradation in a durability tests is probably realistic by considering Figure 2. The rather wide confidence intervals should be investigated more in detail in future work.

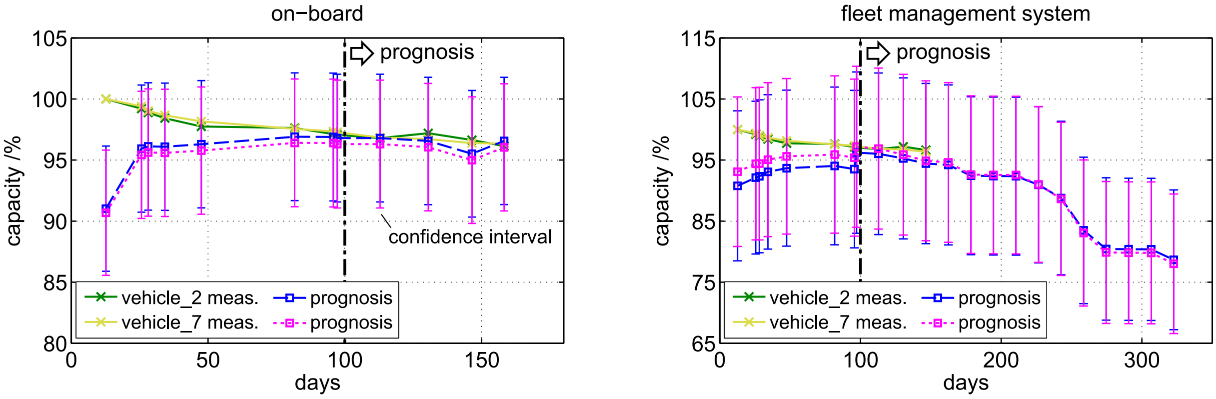

Figure 10: RVM prognosis results from the on-board approach on the left graph and from the fleet management system, respectively, for a longer time period on the right

References

[1] E. Mallia et.al., *Using EV Telematics to Monitor Real-World Battery Health for EV Owners and Fleet Operators*, EVS29 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, 2016.

[2] A. Barré et.al., *Battery Capacity Estimation and Health Management of an Electric Vehicle Fleet*, IEEE Vehicle Power and Propulsion Conference (VPPC), DOI:10.1109/VPPC.2014.7007025, 2014.

[3] F. Hermansson et.al., *Experiences from Data Collection in Living Labs-The Swedish National Database for Car Movements*, EVS29 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, 2016.

[4] A. Nuhic et.al., *Health Diagnosis and Remaining Useful Life Prognostics of Lithium-Ion Batteries Using Data-Driven Methods*, Journal of Power Sources, DOI:dx.doi.org/10.1016/j.jpowsour.2012.11.146, 239(2013), 680-688.

[5] A. Lamm et.al., *Lithium-Ionen-Batterie Erster Serieneinsatz im S400 Hybrid*, Automobiltechnische Zeitung, DOI:10.1007/BF03222086, 111(2009), 490-499.

[6] B. Xu et.al., *Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment*, IEEE Transactions on Smart Grid, DOI:10.1109/TSG.2016.2578950, 2016.

[7] A. Nuhic et.al., *A Health-Monitoring and Life-Prediction Approach for Lithium-Ion Batteries Based on Fatigue Analysis*, Advanced Automotive Battery Conference Europe 2013 , 2013.

[8] J. J. Oliver et.al., *Bayesian Approaches to Segmenting a Simple Time Series*, DOI:doi.org/10.4225/03/5934f8a05b1c7 , 1997.

[9] D. Wang et.al., *Prognostics of Lithium-Ion Batteries Based on Relevance Vectors and a Conditional Three-Parameter Capacity Degradation Model*, Journal of Power Sources, DOI:dx.doi.org/10.1016/j.jpowsour.2013.03.129, 239(2013), 253-264.

[10] D. Liu et.al., *Lithium-Ion Battery Remaining Useful Life Estimation with an Optimized Relevance Vector Machine Algorithm with Incremental Learning*, Measurement, DOI:dx.doi.org/10.1016/j.measurement.2014.11.031, 63(2015), 143-151.

[11] M. E. Tipping, *Sparse Bayesian Learning and the Relevance Vector Machine*, Journal of Machine Learning, DOI:dx.doi.org/10.1162/15324430152748236, 2001, 211-244.

[12] A.-J. Smola et.al., *Statistics and Computing*, 2004, 199-222.

[13] N. Syed et.al., *Incremental Learning with Support Vector Machines*, In Proceedings of the Workshop on Support Vector Machines at the International Joint Conference on Articial Intelligence (IJCAI-99), 1999.

[14] D. Krstajic et.al., *Cross-Validation Pitfalls when Selecting and Assessing Regression and Classification Models*, Journal of Cheminformatics, 2014.

[15] M. Robnik-Sikonja et.al., *An Adaptation of Relief for Attribute Estimation in Regression*, In Proceedings of the Fourteenth International Conference on Machine Learning (ICML '97), 1997, 296-304.

[16] D. Socie et.al., *Modelling Variability in Service Loading Spectra*, Journal of ASTM International, DOI:doi.org/10.1520/JAI11561, 2004, 1-12.

Authors

Dipl.-Ing. Adnan Nuhic studied Electrical Engineering at the University of Hannover, Germany, and joined the Daimler AG in 2011 as a PhD-student. His research activity is carried out within the data-based modelling of lithium-ion battery health as well as the estimation of remaining useful life in on-board applications. In 2015 he moved to the Deutsche ACCUmotive GmbH and Co Kg as a Senior Software Developer responsible for application software functionalities like power calculation and prediction.

Jonas Bergdolt is a statistics student at Humboldt University Berlin, Germany. At Deutsche ACCUmotive GmbH and Co Kg, he writes his master thesis about life-time prediction for lithium-ion batteries. His research interests lie in the areas of machine learning, Bayesian statistics and robust estimation.

Dr. rer. nat. Bernd Spier studied Physical Chemistry and gained his PhD on phase transitions of liquid crystals in electric fields at the University of Paderborn, Germany. He joined Daimler Benz Research in 1991 and worked on development of electrochemical storage systems for aerospace and electro-mobility. After a three years stay in the joint Hybrid Development Center of GM, BMW and DaimlerChrysler in Troy, Michigan, USA he moved to the Deutsche ACCUmotive GmbH and Co Kg in 2009. Currently, he is responsible for the development of the battery management application software and the thermo-electrical Simulation of high-voltage battery systems.

Dr.-Ing. Michael Buchholz received his diploma degree in Electrical Engineering and Information Technology as well as his Ph.D. degree from University of Karlsruhe/Karlsruhe Institute of Technology in 2004 and 2010, respectively. Since 2009, he has been serving as a group leader and lecturer at the Institute of Measurement, Control, and Microtechnology at Ulm University. His main research interests comprise system identification and diagnosis, electric mobility, automated and connected driving, and mechatronical systems.

Prof. Dr.-Ing. Klaus Dietmayer is director of the Institute of Measurement, Control, and Microtechnology at Ulm University and head of driveU. His main research interests are information fusion, classification methods, stochastic filters and tracking, signal processing, and dynamical modeling and control within automotive applications.