Long-term Evaluation of the State-of-Health of Traction Lithium-ion Batteries in Operational Buses

Miguel Simão[1], Rune Prytz[2], and Sławomir Nowaczyk[3]

[1,2] *Stratio Automotive, Lisbon, Portugal*
miguelsimao@stratioautomotive.com
rune@stratioautomotive.com

[3] *Halmstad University, Halmstad, Sweden*
slawomir.nowaczyk@hh.se

Abstract

In this paper, we present and evaluate a novel methodology to estimate the usable capacity and state-of-health (SOH) of lithium-ion batteries in battery-electric buses (BEV). This methodology is designed to be applicable to any BEV in normal operation, independently of battery chemistry, and without requiring complex electrochemical models or large data sets. We have tested the proposed methodology on two vehicle fleets with a total of 105 vehicles, for which we have been acquiring data for up to three years. Additionally, we have analysed the operation of the fleets in terms of daily distance driven and the charging strategies chosen by the operators.

The monitored vehicles are part of fleets currently in normal operation in Europe. The data collection is done with a third-party data logger that is connected to the vehicles' Communication Area Network (CAN) buses, and no additional changes were made to the vehicle's hardware or software. The results show that the proposed methodology shows significantly lower variance in SOH estimation than the alternative methodologies. It also shows similar accuracy in the long-term and smaller short-term deviations from the typical capacity fade model.

1. Introduction

Most urban bus fleet operators are planning the replacement of their older Internal Combustion Engine (ICE) buses to zero local emission (ZE) vehicles. At the moment, the most cost-effective ZE vehicle is the battery-electric vehicle (BEV), mainly due to the low operational costs. However, BEVs have a high cost of acquisition and their batteries have a relatively limited lifespan, which means there is a regular replacement cost. The transition from an ICE to a BEV fleet is only profitable if the operational costs are kept low. This requires minimizing downtime and maximizing battery life. Accurate knowledge of the state-of-health (SOH) of the batteries during their life is essential to ensure the vehicle can successfully perform the required routes with minimal downtime. The SOH can also be used to make a projection of when it will reach a threshold at which its operation is no longer reliable, so that a replacement can be planned ahead of time. The ideal solution is minimally intrusive so that any fleet operator can quickly and cheaply deploy it, while also being accurate and reliable enough to ensure usability.

There are several ways of determining the SOH of a battery in the literature. A thorough review is presented in Berecibar et al. (2016). A common way is by direct measurement, where an experimental setup with controlled conditions is used to measure and estimate parameters such as the SOH. Coulomb Counting (CC) is a method that is often used to estimate cell capacity (Ng, Moo, Chen, & Hsieh, 2009; Lipu et al., 2018). In Gismero, Schaltz, and Stroe (2020), the authors proposed a recursive least squares filter to reduce uncertainty in CC-based SOH estimates. More refined approaches include electrochemical impedance spectroscopy (EIS) and incremental capacity analysis (Andre et al., 2011; Pastor-Fernández, Uddin, Chouchelamane, Widanage, & Marco, 2017). Other parameters, such as internal resistance are often estimated with equivalent circuit models (ECM). However, the conditions necessary to implement these methods make them unfeasible for BEV in operation (W. Li et al., 2021; Vichard et al., 2021).

Data-driven approaches are also explored in the literature (Y. Li et al., 2019; W. Li et al., 2021). These approaches

model the SOH using machine learning methods on previously collected data. The limitation of these approaches is that it is very difficult to train a model that will generalize well for most operational conditions. This is especially concerning if a new battery architecture or chemistry that is not in the training data set is found in operation after the model is deployed. Since we do not have a representative data set, nor can we build one at this time, we have not pursued data-driven approaches.

In light of the challenges presented in the state-of-the-art methodologies, we designed a methodology based on CC. This methodology may be less accurate than more refined methodologies found in the literature, but it is faster to implement and should work reliably for any BEV battery in most operational conditions. We introduce some filters and binning on the input data in order to reduce the variance of the SOH estimation. In this paper, we describe the chosen methodology and present results for a fleet of 105 BEV with up to three years of data.

2. METHODOLOGY

2.1. State-of-Health

The SOH is an indicator of the state of the battery with respect to its lifespan. There are several possible definitions; some involving battery parameters such as the internal resistance or parameters extracted from incremental capacity analysis (ICA). We opted to use the definition in Eq. (1), i.e., the fraction between the actual usable capacity and the original capacity. We believe this is the most useful definition for a vehicle operator, as it sets an expectation for the maximum energy (and thus vehicle range) that can be extracted from a battery when it is fully charged.

$$S(t) = \frac{Q(t)}{Q(0)} \quad (1)$$

where $Q(t)$ is the capacity of the battery in Ah at time t. However, it is generally known that the actual capacity also depends on other factors, such as operating temperature and (dis)charge current. These factors will be important to take into account when designing a diagnosis and prognosis model.

Ideally, we would fully charge or discharge a cell under controlled conditions, measure the current over time and then integrate it to estimate the total capacity, as seen in Eq. (2). However, in operation, we cannot assume to have a controlled set of conditions for this estimation. The operation temperature can vary significantly and the discharge current is highly dependent on the usage profile. Furthermore, as a third-party who can only inspect the data available in the Communication Area Network (CAN) bus, we also have limitations in signal resolution and acquisition rate. We are also unable to

control the charge/discharge rates and the depth-of-discharge (DOD) for each cycle. In order to ensure that our estimate is precise, with conditions that do not change significantly over the analysis period, we have chosen to estimate the SOH using data from charging sessions only, which should present similar conditions over time. This should work well for operation modes with slow charging at a depot at the end of the day, but not so well when operation only includes opportunity fast charging throughout the day.

$$Q = \int_{t_0}^{t_1} I(t)dt \quad (2)$$

2.1.1. Calculation Method

The estimation of battery capacity is derived from the CC method commonly used to determine the battery state-of-charge (SOC). Independently, we have reached a method similar to the one presented in Vichard et al. (2021), but we have some additional features in order to improve estimation precision.

In Vichard et al. (2021), the proposed battery capacity estimation is given by:

$$Q_{\max} = \frac{\int_{t_0}^{t_1} I(t)dt}{\text{SOC}(t_1) - \text{SOC}(t_0)} \quad (3)$$

where $I(t)$ is the battery current, t_0 and t_1 are the start and end times for the estimation period. The authors note that this method yields low precision when small SOC variations are used, a limitation that we also experienced. We propose a solution to solve this issue without discarding data from these instances.

We start by extracting charging session segments from the data, i.e., segments where the vehicle is plugged-in and charging. For those segments, we collect the current and SOC signals given by the battery management system (BMS). The charging segments are then cut further into bins of 10% of SOC, e.g., $[0, 10[, [10, 20[, ..., [90, 100]$ %. Generally speaking, the maximum admissible SOC range $[0, 100]\%$ is discretized into N equally sized bins. In terms of notation, we will identify each bin with an indexed variable $b_i(t)$, where the index will be the 0-based position of the bin in the set of bins and t the timestamp of the charging session that originated the data point. For example, $b_1(t)$ corresponds to bin $[10, 20[$ when $N = 10$. Figure 1 shows a few representative examples of how partial charging sessions are discretized. For instance, session 0 represents a charge from 15 to 90%. In that case, bin b_0 represents the $[0, 10[\%$ SOC interval and originates a missing value, as the session does not cover that interval. Bin b_1 ($[10, 20[\%$) is partly covered, as the session starts with the SOC at 15%. Partially covered bins are treated as missing values. Bins b_2 through b_8 are intervals fully-
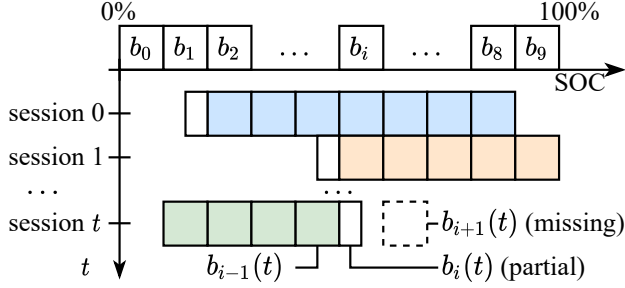
Figure 1. Representation of SOC discretization of partial charging sessions, assuming 10 bins. The filled boxes, e.g. $b_{i-1}(t)$, represent bins that fully cover their SOC interval. White boxes at the beginning or end of a session, e.g. $b_i(t)$, represent a partially covered interval. An example of a bin with a missing charge value is shown as $b_{i+1}(t)$.

covered in this session, so they originate known/non-missing values and are used for direct estimation of capacity in each SOC bin.

We calculate the charge q_i of each bin and each charging session using Eq. (2). The charge of bin $b_i(t)$ is represented as $q_i(t)$. The signals $q_i(t)$ are then filtered with an exponentially weighted moving average over time. In the case $q_i(t)$ is missing for session t due to partial charging sessions, the values have to be imputed. There are several methods for imputation. In this work, we opted to only experiment with forward-filling, i.e., the missing value is imputed by value of the previous known value $q_i(t - k)$. In practice, we use the value extracted from the last charging session that covered the missing value's bin. This process will not impute missing values for bins that have never been covered with a session. This happens frequently at the beginning of data acquisition for a battery. In this case, since we do not want to use data from the future, we opted by imputing with $\overline{q}(t)$, i.e., the mean value of the non-missing values of q_i within the same charging session t. In order to distinguish non-missing and missing charge values $q_i(t)$, we introduce a symbol ϕ_i that takes the value of zero or one depending on whether q_i is missing or non-missing, respectively.

The final estimated maximum capacity $Q(t)$ as a function of the bin charges q_i is:

$$Q(t) = \sum_{i=0}^{N-1} \phi_i q_i + (1 - \phi_i)\overline{q} \qquad (4)$$

where $\overline{q}$ is the mean value of the non-missing bin charges for charging session t. The maximum capacity estimated at the time t for a partial charging session is the sum of the latest charge values of each bin q_i, with $i \in \{0, 2, ..., 9\}$. The first term of this equation is related to the direct measurement of capacity q_i, that occurs when a partial charging session originates bins with valid values ($\phi_i = 1$) bin. The second term

defines the extrapolation done when a specific SOC bin does not have any valid values in time, on which case imputation with the value $\overline{q}$ is used.

An algorithm based on the direct application of Eq. (3) has an implied assumption of linearity between the charge added during the charging session and the SOC increment. With our approach, the charge q_i of each bin b_i is estimated independently from that of all other bins, as each SOC bin uses independent data. The discretization of the proposed approach can approximate the maximum charge of batteries even with non-linear SOC-charge relationships. As a drawback, this method can only use the bins with data that spans the whole range of the bin. For example, a charging session that starts at 13% and ends at 92% SOC will not use the data in the ranges $[13, 20[\%$ nor $[90, 92]\%$, as they do not cover the whole bin range. We could reduce the impact of this issue by increasing the number of bins. In practice, reducing the bin size increases the estimation variance mainly due to inaccuracies in SOC estimation.

The definition of $Q(t)$ allows us to estimate the SOH using Eq. (1). We will not describe in this paper the best methodology to estimate $Q(0)$ – the maximum capacity at the beginning of life (BOL) –, so we will assume the a value linearly extrapolated from the acquired data.

2.1.2. Evaluation

Normally, we would evaluate the proposed method by comparing the SOH estimations to the actual SOH. However, for a fleet in normal operation, we are not able to perform the charge/discharge capacity tests in coherence with those performed by the battery manufacturer. Additionally, the beginning of acquisition (BOA) is not always equal to the battery BOL, which means that the SOH true value will naturally be below 100%. Since we are also not able to measure the true SOH, we cannot quantitatively measure the accuracy of our methodology.

We will focus the evaluation on metrics that fleet managers may also be interested on, such as the precision of our SOH estimate and its ability to detect the typical capacity fade expected in a lithium-ion battery. The precision is an important metric for the fleet manager since in increases their confidence in the estimations, which is important for decision-making. It is measured quantitatively by the standard deviation over a rolling window. On the other hand, the ability to accurately track capacity fade will be a qualitative metric.

2.2. Operation Metrics

As previously mentioned, the charging strategy used in the operation of BEV may have an impact on the accuracy of the SOH estimation. In order to understand the charging strategy, we gather some metrics from the data of each charging

session, such as:

1. Start SOC (%)
2. End SOC (%)
3. Delta SOC (Start SOC − End SOC) (%)
4. Average Power (kW)
5. Odometer value (km)

We then analyse the distributions of these metrics in order to understand if fast-charging is being used during operation, or if overnight charging is dominant. We will also look into correlations between charging strategies and the daily driven distance.

3. RESULTS AND DISCUSSION

3.1. Data Collection

The data is collected on-board the vehicle using Stratio's[1] data logger. The logger is connected the CAN of the vehicle. CAN networks are the *de facto* standard for in vehicle communication and used by the multitude of on-board computers to share data between them. For example, these data can be sensor values or computed actuator positions used for controlling the multitude of actuators within a vehicle. As the complexity of vehicles has increased in the past years, more CAN buses have been added in order to accommodate the increase in used bandwidth. Nowadays, buses and trucks commonly use eight or more CAN lines, often connecting more than 30 on-board computers. The specialization of these networks has lead to several new CAN protocols, as well as some proprietary CAN configurations. These configurations pose a challenging problem when trying to read data from them. Without the proper configuration settings, the data are impossible to decode. The setting can be obtained from the original equipment manufacturer (OEM) or through extensive reverse engineering.

Stratio's data logger can simultaneously connect up to three CAN buses and monitor more than 300 signals. This easily amounts to very large volumes of data and thus high data transfer costs, as the collected data is sent to Stratio's servers through a 4G cellular connection. The device utilizes both down sampling and data compression in order to reduce the transmission costs.

We have collected data in real-time from two vehicle fleets with distinct operators and non-overlapping vehicle models. All vehicles in the fleets are BEV designed for urban operation. The collected date ranges are shown in Table 1. It should be noted that communication can be temporarily interrupted for some vehicles, so there is no guarantee that each vehicle in the fleet has data available in the whole date range. We have collected data from multiple systems within the vehicles, but the signals acquired from the traction battery systems include

Table 1. Description of the vehicle fleets data collection date ranges.

Fleet	Start Date	End Date
FLT0	2021-06-01	2021-11-15
FLT1	2019-01-01	2021-11-15

Table 2. Description of the vehicle data acquisition parameters.

Signal	Fleet 0		Fleet 1	
	Period	Resolution	Period	Resolution
Odometer	10s	1km	10s	1km
Current	1s	4A	1s	1A
Voltage	1s	3.8V	1s	3.4V
SOC	10s	1.2%	10s	0.4%

current, voltage, BMS SOC and mean cell temperature. The acquisition parameters for each signal are described in Table 2.

3.2. Operation Analysis

We start by doing an analysis of how the vehicles are being operated in our fleets. In Figure 2, we plot the distributions of the starting SOC and delta SOC for charging sessions, grouped by the fleets described in Table 1. It should be noted that in these box plots, the whiskers represent 10 and 90% quantiles, while the colored box represents the quantile Q1, median and quantile Q3, in order.

Both of these metrics show very different distributions. In regards to the start SOC of the charging sessions, FLT0 shows a median of 73%, while FLT1 has a median of 48%. There is no intersection between fleets of the intervals between Q1 and Q3. On the other hand, the delta SOC median is 14% for FLT0 and 52% for FLT1. There is also no intersection of the intervals between Q1 and Q3 for this metric.

The data show that FLT0 charges mostly during operation with short stops for fast-charging. The SOC rarely drops below 50% and the typical stop adds less than 20% of charge. On the contrary, FLT1 uses exclusively slow-charging at the depot after a full operation day. This is confirmed by the significantly lower starting SOC and larger SOC increases of the charging sessions. The vehicles typically charge when the SOC reaches below 50% SOC and then are charged until full. The fact that FLT0 has half of their charges with SOC deltas below 15% means that likely less than half of the acquired data is being used for the SOH estimation following our methodology.

We also explored the relationship between the fleet charging strategy, average charging power and the daily distance driven, represented in Figures 3a and 3b. FLT1 shows a median daily distance driven of 170km, with 75% of days
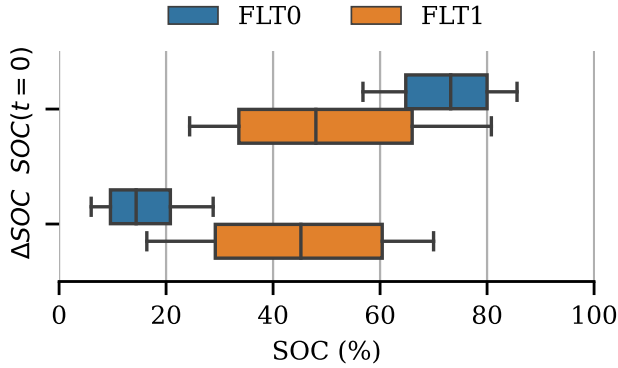
Figure 2. Box plot of two charging sessions parameters: starting SOC, $SOC(t = 0)$, and the total SOC increase, ΔSOC. The whiskers define the 10 and 90% quantiles. Colors encode the distinct vehicle fleets.
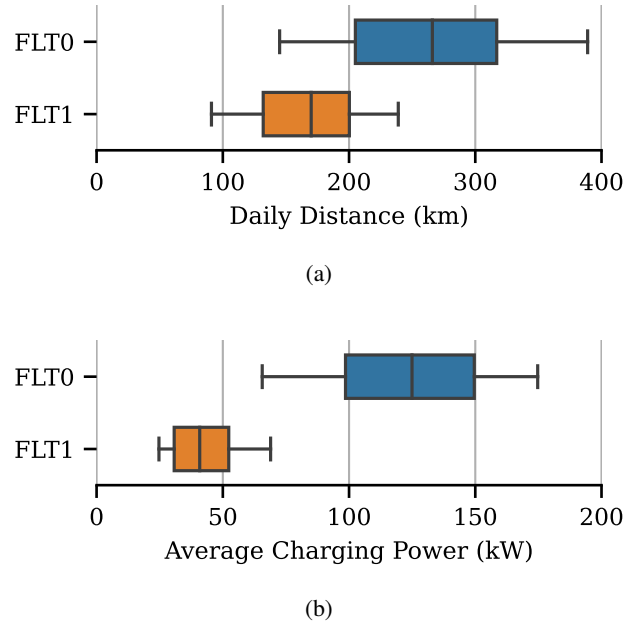

(a)

(b)

Figure 3. Visualization of daily (a) travelled distance and (b) session average charging power, separated by the two vehicle fleets. The whiskers define the 10% and 90% quantiles.

showing less than 200km. For these distances, which can be covered with a single charge, overnight charging is ideal, as the operator can also take advantage of off-peak electricity rates. FLT0 shows a median daily distance of 266km with 75% of days below 317km, and some vehicles going over 400km. This is beyond the operational limit of today's BEVs and therefore requires fast-charging during the day. This specific fleet uses a network of fast-chargers deployed by the fleet operator. The use of fast-charging is confirmed by the visualization of charging power in Figure 3b. While FLT0 shows a median power of 125kW peaking on 250kW, FLT1 is limited to a median of 41kW and peak of 125kW.

3.3. State-of-Health

In this section, we present the results of the evaluation strategy described in Section 2.1.2. We have evaluated the performance of the methodology proposed in this paper described by Eq. (4), hereinafter referred to as *Proposed Algorithm*. For comparison, we have also evaluated the methodology described by Eq. (3) in Vichard et al. (2021), which we call *Algorithm 1*.

The first evaluation metric is the standard deviation of the SOH predictions. We calculate it on 30-day rolling windows per vehicle and average it for the entire fleet. We aggregate this metric across the two methodologies and fleets analysed. We have also removed charges below 10% for *Algorithm 1*, in order to reduce the estimation noise. The results are shown in the top plot of Figure 4.

The plot shows no interesting features or trends in the noise over time in any case. For either fleet, the proposed methodology has significantly lower noise ($p \ll 0.01$). There is a significant difference in precision between fleets with both methodologies. We believe this is mostly due to the different operation factors that we presented in the previous section.

Fleet FLT1 uses mostly fast-charging at a rate that is limited by the battery SOC at the time. Therefore, the battery current variance is proportional to the variance observed in the SOC level. Fleet FLT0 does not as high a current variance as FLT1 observes due to the use of overnight charging. A second source of uncertainty may be sensor accuracy (SOC and current), as different fleets contain vehicles from different OEMs. In this study, we will not test these hypotheses.

In regards to the accuracy qualification of the SOH curves, our best efforts are limited to verifying the capacity fade rates seen in the fleets. To estimate the fade rate, we applied a linear regression model on the SOH metric using the odometer value as the independent variable. The results show strong variance within the fleets, but mean capacity fade rates of 2.4 % per 100.000km for *Proposed Algorithm* and 7.1% for *Algorithm 1*. Either value is possible for lithium-ion batteries depending on design and operation parameters. For visualization, we show the fleet-averaged SOH with a rolling window of 30 days in the bottom plot of Figure 4. The plot shows the negative trend of SOH that characterizes the process of capacity fade across all fleets and methodologies. Comparing the two methodologies on the longest operating fleet FLT1, we see that our methodology shows significantly smaller deviations from the linear decay. Although some cell processes, such as capacity regeneration, may temporarily change the maximum available capacity, they are unlikely to create the deviations we see with *Algorithm 1*. Therefore, we believe that the proposed methodology may be more accurate.
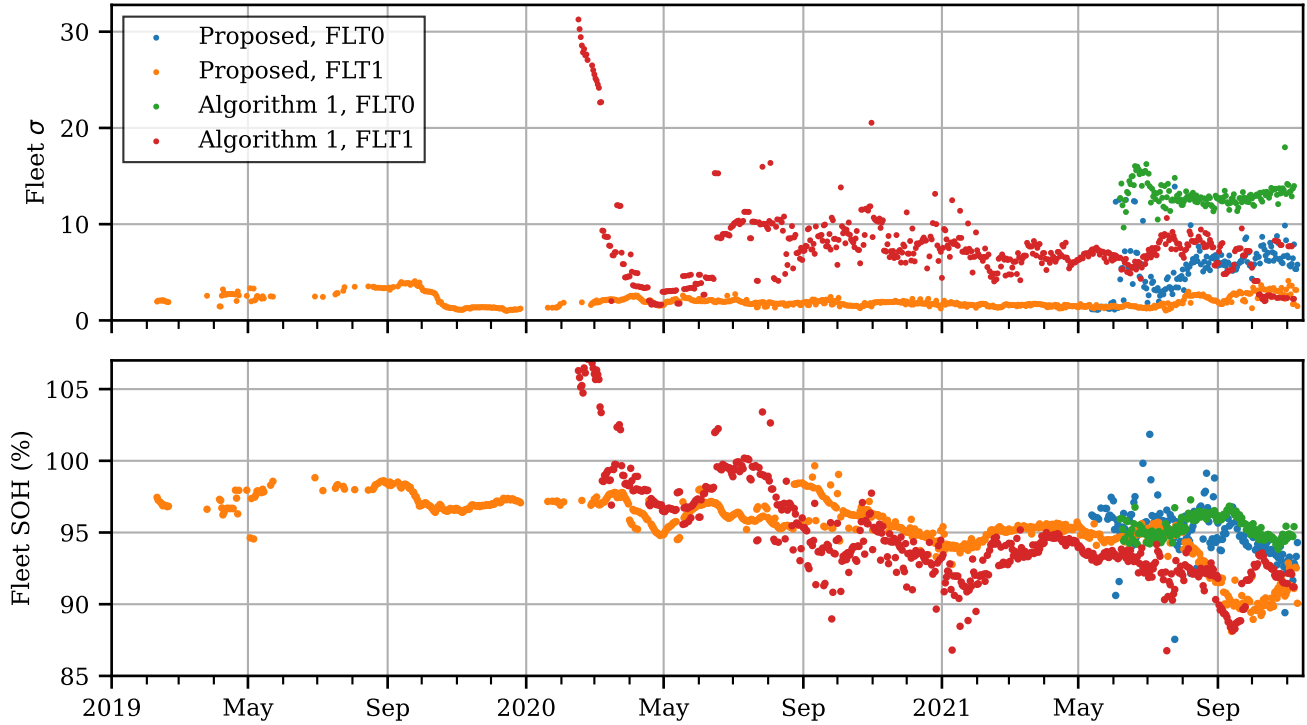
Figure 4. Top figure: 30-day rolling average of the SOH metric standard deviation. Bottom figure: 30-day rolling average of the SOH metric. Both plots contain data split by methodology and vehicle fleet. The *Proposed* methodology refers to the one proposed in this paper, described by Eq. (4), and *Algorithm 1* refers to the methodology presented in Vichard et al. (2021).

4. CONCLUSION

We presented in this paper a novel methodology to address the problem of determining the SOH of a traction battery within an electric vehicle. Although there are several methodologies already described in the literature, many of them are very difficult to implement in an active fleet. The proposed methodology was implemented using a third-party connected data logger in urban vehicle fleets in normal operation. This type of solution is valued by fleet operators who are interested in monitoring the BEV in their fleets with a minimally intrusive solution.

We have tested our methodology on a data set that we have acquired from 105 vehicles covering a time-span of up to three years. The results show that the solution can be integrated in a battery health management product. Moreover, our estimations show lower standard deviation while maintaining or slightly improving accuracy when compared to other methodologies in the literature. On a longer time scale, the estimated SOH over time appears to correctly follow the expected capacity fade of lithium-ion batteries.

Future work will involve further reducing variance of the SOH estimation by improving the SOC signal given by the vehicle's BMS. There is also the opportunity to study the im-

pact of other imputation methods for the missing data caused by partial charge sessions. We will also keep acquiring data for these fleets in an effort to develop new prognostics models for the battery based on the SOH metric.

REFERENCES

Andre, D., Meiler, M., Steiner, K., Wimmer, C., Soczka-Guth, T., & Sauer, D. (2011). Characterization of high-power lithium-ion batteries by electrochemical impedance spectroscopy. i. experimental investigation. *Journal of Power Sources*, *196*(12), 5334-5341. doi: https://doi.org/10.1016/j.jpowsour.2010.12.102

Berecibar, M., Gandiaga, I., Villarreal, I., Omar, N., Mierlo, J. V., & Bossche, P. V. D. (2016, 4). Critical review of state of health estimation methods of li-ion batteries for real applications. *Renewable and Sustainable Energy Reviews*, *56*, 572-587. doi:

10.1016/J.RSER.2015.11.042

Gismero, A., Schaltz, E., & Stroe, D.-I. (2020). Recursive state of charge and state of health estimation method for lithium-ion batteries based on coulomb counting and open circuit voltage. *Energies*, *13*(7). doi: 10.3390/en13071811

Li, W., Sengupta, N., Dechent, P., Howey, D., Annaswamy, A., & Sauer, D. U. (2021, 1). Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. *Journal of Power Sources*, *482*, 228863. doi: 10.1016/J.JPOWSOUR.2020.228863

Li, Y., Liu, K., Foley, A. M., Zülke, A., Berecibar, M., Nanini-Maury, E., ... Hoster, H. E. (2019). Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and Sustainable Energy Reviews*, *113*, 109254. doi: https://doi.org/10.1016/j.rser.2019.109254

Lipu, M. H., Hannan, M., Hussain, A., Hoque, M., Ker, P. J., Saad, M., & Ayob, A. (2018). A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations. *Journal of Cleaner Production*, *205*, 115-133. doi: https://doi.org/10.1016/j.jclepro.2018.09.065

Ng, K. S., Moo, C.-S., Chen, Y.-P., & Hsieh, Y.-C. (2009). Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Applied Energy*, *86*(9), 1506–1511. doi: https://doi.org/10.1016/j.apenergy.2008.11.021

Pastor-Fernández, C., Uddin, K., Chouchelamane, G. H., Widanage, W. D., & Marco, J. (2017). A comparison between electrochemical impedance spectroscopy and incremental capacity-differential voltage as li-ion diagnostic techniques to identify and quantify the effects of degradation modes within battery management systems. *Journal of Power Sources*, *360*, 301-318. doi: https://doi.org/10.1016/j.jpowsour.2017.03.042

Vichard, L., Ravey, A., Venet, P., Harel, F., Pelissier, S., & Hissel, D. (2021, 6). A method to estimate battery soh indicators based on vehicle operating data only. *Energy*, *225*, 120235. doi: 10.1016/J.ENERGY.2021.120235

BIOGRAPHIES

Miguel Simão received the M.Sc. degree from the University of Coimbra, Portugal, in 2014, and the joint Ph.D. degree from University of Coimbra and École National Supérieure d'Arts et Métiers, Paris, France, in 2014 and 2018 respectively, both degrees in mechanical engineering with a focus on machine learning applied to robotics. He is a Data Scientist at Stratio Automotive, Portugal, since 2019, working on research and development of prognostics and diagnostics models. He was also a visiting Researcher at the Jet Propulsion Laboratory, NASA, California Institute of Technology, Pasadena, CA, USA, in 2015. His research interests are in the application of machine learning to novel problems, previously in robotics, and nowadays to vehicle prognostics and electric vehicles.

Rune Prytz is the Head of Research at Stratio. Rune is an automotive expert with an extensive background in vehicle data analysis, from his 12 years old experience at Volvo as a Research Engineer and Specialist. Rune is responsible for all the strategic and technical planning of Stratio's AI research roadmap for vehicle and system failure detection, and he is part of the strategic business planning team. His work in the field is considered a reference and has been cited over 250 times.

Sławomir Nowaczyk is Professor in Machine Learning, working at Center for Applied Intelligent Systems Research, Halmstad University, Sweden. He has received his MSc degree from Poznan University of Technology in 2002 and his PhD degree from the Lund University of Technology in 2008. During the last decade his research focused on knowledge representation, data mining and self-organising systems, especially in large and distributed data streams, including unsupervised modelling. He is a board member for the Swedish AI Society, and a research leader for the School of Information Technology at Halmstad University. Sławomir has led multiple research projects related to applying Artificial Intelligence and Machine Learning in many different domains, such as transport and automotive, energy, smart cities as well as healthcare. In most cases, this research was done in collaboration with industry and public administration organisations – inspired by practical challenges and leading to tangible results and deployed solutions.