

A Machine Learning-based Digital Twin for Electric Vehicle Battery Modeling

*Original*

A Machine Learning-based Digital Twin for Electric Vehicle Battery Modeling / Alamin, KHALED SIDAHMED SIDAHMED; Chen, Yukai; Macii, Enrico; Poncino, Massimo; Vinco, Sara. - ELETTRONICO. - (2022), pp. 1-6. (Intervento presentato al convegno IEEE International Conference on Omin-Layer Intelligent Systems (COINS) tenutosi a Barcelona (Spagna) nel 1-3 August 2022) [10.1109/COINS54846.2022.9854960].

*Availability:*

This version is available at: 11583/2968469 since: 2022-09-21T08:24:42Z

*Publisher:*

IEEE

*Published*

DOI:10.1109/COINS54846.2022.9854960

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

IEEE postprint/Author's Accepted Manuscript

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

# A Machine Learning-based Digital Twin for Electric Vehicle Battery Modeling

Khaled Sidahmed Sidahmed Alamin<sup>1</sup>, Yukai Chen<sup>1</sup>, Enrico Macii<sup>2</sup>, Massimo Poncino<sup>1</sup>, Sara Vinco<sup>1</sup>

<sup>\*</sup>Department of Control and Computer Engineering, Politecnico di Torino, Turin, Italy

<sup>†</sup>Interuniversity Department of Regional and Urban Studies and Planning, Politecnico di Torino, Turin, Italy

Email: khaled.alamin@polito.it, yukai.chen@polito.it, enrico.macii@polito.it, massimo.poncino@polito.it, sara.vinco@polito.it

**Abstract**—The widespread adoption of EVs is limited by their reliance on batteries with presently low energy and power densities compared to liquid fuels and are subject to aging and performance deterioration over time. For this reason, monitoring the battery state of charge and state of health during the EV lifetime is a very relevant problem. This work proposes the structure of a battery digital twin designed to reflect battery dynamics at the run time accurately. To ensure a high degree of correctness concerning non-linear phenomena, the digital twin relies on data-driven models trained on traces of battery evolution over time: a state of health model, repeatedly executed to estimate the degradation of maximum battery capacity, and a state of charge model, retrained periodically to reflect the impact of aging. The proposed digital twin structure will be exemplified on a public dataset to motivate its adoption and prove its effectiveness, with a high degree of accuracy and inference and retraining times compatible with onboard execution.

## I. INTRODUCTION

Since General Motors sold the first Electric Vehicle (EV) in 1996, the market of EVs has experienced explosive growth, and in the past five years, global electric car sales have more than quadrupled [1]. This dramatic growth is an effect of both environmental and economic considerations, that lead governments to lower CO<sub>2</sub> emissions standards and to implement increased subsidy schemes for EVs as part of stimulus packages to counter the effects of the pandemic [2]. However, the effectiveness of EVs is limited by their reliance on batteries that have presently low energy and power densities compared to liquid fuels and are subject to aging and performance deterioration over time [3].

To partially overcome these limitations, EVs are usually powered by Lithium-ion batteries that have high power and energy density and long cycle life compared to other chemistry compositions [3]. Nonetheless, battery management is critical to enhancing safety, reliability, and performance. In particular, accurate estimation of battery *State of Charge (SOC)* and *State of Health (SOH)* is crucial to deliver essential information about battery charge and aging level, which can be used to perform maintenance or adapt management strategies to improve performance and reliability (Figure 1) [4].

Unfortunately, both metrics must reflect complex non-linear battery dynamics, and modern Battery Management Systems (BMS) require knowledge of historical data of battery usage that can not be quickly processed onboard in real-time. For

this reason, *battery digital twins* have been introduced, aiming at building an accurate run time model of the SOC and SOH metrics to provide accurate battery capacity degradation assessment and relevant information to the BMS [4]. This is achieved through a combination of Internet of Things communication protocols, cloud/edge computation, and machine learning applications that jointly achieve an exponential increase in data storage and computation capability [4], [5].

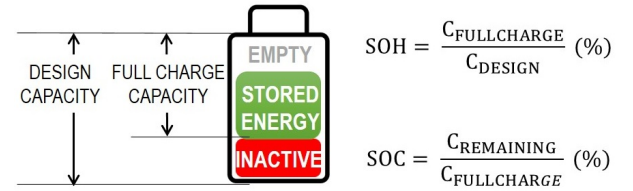


Figure 1: Representation of battery SOH and SOC metrics: SoH quantifies battery capacity degradation and impacts on the estimation of SOC.

The main ingredient of a digital twin is a model of the process/system under monitoring, in this case, the relevant battery SOH and SOC dynamics: the model is used as a reference for the process under analysis to detect any misalignment or unexpected behavior, and to foresee the future evolution. In the literature, many battery models have been proposed, ranging from electrochemical models to circuit-equivalent models [6]. Machine learning techniques are becoming increasingly popular for battery SOC and SOH estimation [7]. Finding a balance between modeling accuracy, model responsiveness, sensitivity to battery evolution, and the architectural constraint imposed by the EV scenario is far from trivial. Additionally, as clarified by Figure 1, SOH and SOC are tightly interleaved: charge and discharge cycles affect battery SOH, which reduces maximum total charge capacity and the corresponding estimation of battery SOC. This implies that the SOH and SOC models can not be built independently and that the SOC dynamics must be aware of the status of the SOH metrics.

This paper discusses the challenges posed by the construction of a battery digital twin, and *proposes a custom architecture of models for SOC and SOH to propose a viable solution for an accurate estimation of battery dynamics and aging effects*. In detail:

- the digital twin provides a model of SOH built on historical data, that is trained once on data of a complete charge cycle;
- the model of the SOC is heavily affected by aging: for this reason, the digital twin features a model of SOC that is retrained periodically on the cloud on data uploaded from the EV. The model is then run either on the cloud (to have higher performance) or on the EV (to provide the accurate estimation to the onboard BMS);
- the paper discusses the edge/cloud architecture, even if discussing the details of the digital twin infrastructure (edge/cloud, protocols, etc.) is out of the scope of this paper and will be part of future work;
- we apply the proposed solution to model to a public dataset provided by NASA Ames Research Centre [8], to motivate its adoption, prove the effectiveness of the proposed solution in modeling Lithium-ion batteries, and provide implementations of the SOC and SOH models.

The paper is organized as follows. Section II provides the necessary background on battery dynamics and digital twins. Section III presents the proposed solution, with a strong focus on the digital twin architecture and on how to reproduce effectively the interconnection between SOH and SOC. Section IV applies the digital twin concept to the NASA Ames Research Centre dataset, and finally Section V draws our concluding remarks.

## II. BACKGROUND

### A. Battery Dynamics

Lithium-ion batteries store and supply electric power based on the movement of Li-ions between anode and cathode. Some irreversible reactions happening during the repeated charges and discharges cycles (such as Li plating and growth of solid electrolyte interface) cause performance degradation over time, generating issues in terms of reduced available capacity and/or increased internal resistance [9].

*State of Charge (SOC)* estimation is a relevant metric that measures the percentage of remaining capacity relative to the maximum capacity of the battery. To provide robust estimates, equivalent circuit models are used that achieve good approximation but require detailed physical knowledge of Li-ion batteries. The complementary approach uses data-based models that reconstruct the evolution of battery SOC given curves of, e.g., current, voltage and temperature. Many solutions have been proposed in the literature, using e.g., support vector machines [10], [11] and neural networks [12]–[15]. The main issue of these approaches is that *they build a once-for-all model of SOC that is applied throughout the battery lifetime, thus missing the impact of aging on the internal evolution of battery*. This is especially critical as one of the major impacts of aging is a reduction in battery available maximum capacity, which impacts battery SOC calculation.

Tracking age-dependent changes in battery dynamics is thus necessary to maintain accurate estimates of the remaining lifetime of the battery. In this perspective, *State of Health (SOH)* is a metric that measures the condition of a battery compared

to its ideal conditions and is typically estimated as the ratio of the aged and rated capacity of the battery. SOH is not an observable quantity and is non-linearly dependent on several factors, including charge and discharge profiles, average SOC, and ambient temperature [16]. Given the difficulty in modeling the underlying electrochemical processes, most approaches estimate SOH models with data-driven methods, ranging from genetic algorithms to learning- and neural network-based algorithms [5], [17].

### B. Digital Twins for Battery Management

A digital twin is a virtual representation that serves as the real-time digital counterpart of a physical object or process. In the case of batteries, digital twins are now being used to complement BMS with high computation power and data storage capability, reliability, and accuracy [18]. Works at state of the art differ in terms of adopted battery model. Circuit-equivalent models (e.g., the Thevenin model) are complicated to characterize and may even require experimental measurements to ensure a reasonable degree of correctness [4]. Algorithmic approaches such as genetic algorithms and particle swarm optimization are characterized by low convergence and are computationally expensive, thus not allowing to provide estimations at run time during EV lifetime [5]. Data analytic strategies (like machine learning, support vector machines, and neural networks) overcome these issues by self-learning parameters from historical data. However, they are trained at the beginning of battery lifetime and thus tend to diverge from the actual aged battery along its lifetime [10]–[15], [19], [20]. In this direction, an interesting strategy has been proposed by [4], which proposes to periodically influence the SOC model with updated SOH information, to improve prediction accuracy. This solution improves prediction accuracy by better reflecting the impact of aging on internal battery dynamics. However, this work adopts the classic H-infinity filter and particle swarm optimization techniques for SOC and SOH estimation, and they require a cumbersome parameter identification process every time SOC and SOH estimation is performed. Moreover, the H-infinity filter for SOC estimation is fixed all the time, only leveraging the SOH value to affect the adaptive process in the filter.

## III. A DIGITAL TWIN FOR BATTERY MONITORING

The proposed digital twin of the EV battery storage system focuses on SOC and SOH estimation as the most significant parameters that affect EV's driving experience and battery life.

### A. Models for Battery SOC and SOH

Figure 2 shows the conceptual diagram of our proposed model. The horizontal arrow represents battery lifetime, from nominal conditions (SOH = 100%) to the point when the battery is considered no longer usable (SOH = 80%). The battery icons represent two different classes of models:

- A model of SOH (on top of the arrow), built once for all on historical data and run throughout the whole lifetime of the battery;

- A model of SOC (below the arrow) that is periodically updated to reflect the impact of SOH on the estimation of full battery charge and thus of SOC. The multiple icons reflect the fact that different SOC models are used by the digital twin to estimate the battery SOC in different battery operation time intervals; the model at  $T_i$  shown in Figure 2 represents the latest retrained SOC model adapted to the battery conditions at the  $i_{th}$  time interval.

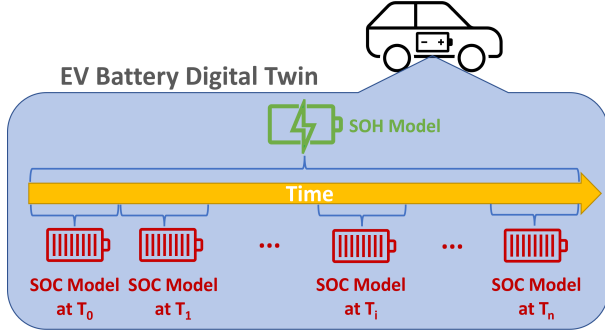


Figure 2: Conceptual diagram of the EV battery digital twin: the arrow is battery lifetime, the icon on top represents the SOH data-driven model, while the icons below the arrow represent the SOC data-driven models that are retrained periodically during battery lifetime.

Both models are data-driven in the sense that they are built based on data traces. It is important to note that any machine learning-based solution can be used to build the battery SOC and SOH models [7]. The battery SOH model is built once for all from historical data of an aging battery of the same type as the one installed on the vehicle. The battery SOC model is initially built on historical data of nominal battery conditions and is then retrained on data sampled from the vehicle during its operation. Two mechanisms can trigger the update of the SOC model: a fixed period or a decrease of SOH by a specific value.

The reason for such a static SOH model combined with a dynamic SOC model design is that the SOH of the battery from 100% to 80% represents the aging of the total available capacitance over the entire operating life cycle of the battery, where the only relationship that changes is that the total available capacitance decreases with the time of use. On the other hand, the SOC model should be constantly updated because the discharge characteristics of the battery are different under different SOH conditions. More specifically, the battery impedance and battery maximum capacity change as SOH decreases; furthermore, the relationship between SOC and battery open-circuit voltage during discharge differs under different SOH conditions [21].

#### B. Digital Twin Architecture and Operation

Figure 3 illustrates more details about how to construct our proposed digital twin and its working mechanism. The diagram shows that our proposed digital twin has two main components, the cloud side and the EV edge side:

- The cloud is used to deploy and run the SOH model, and to periodically retrain the SOC model;
- The edge node on the vehicle is used to collect and periodically upload data to the cloud for SOC model retraining, and of running the updated SOC models; this choice allows the vehicle BMS to be constantly aware of battery SOC estimation.

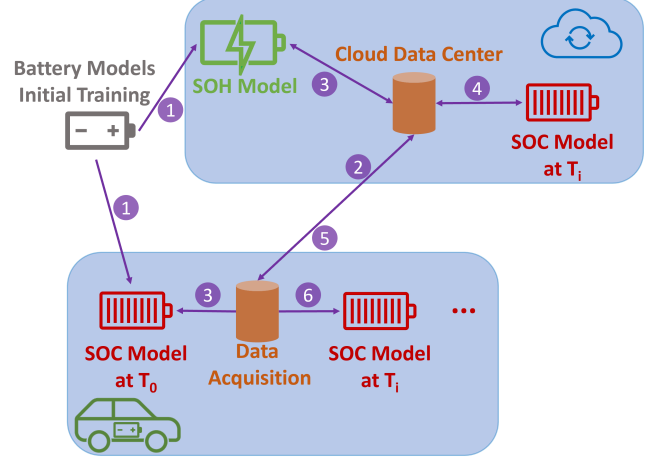


Figure 3: Methodology to implement EV battery digital twin.

The specific steps related to the number shown in the Figure 3 for constructing our proposed digital twin and how it works during operation are described below.

- Step ①: offline training of the SOH model using machine learning techniques on historical measurements of an aged battery; training of the initial SOC model by using a discharge trace of a brand new battery (at nominal battery conditions);
- Step ②: the EV continuously collects online charge and discharge measurements data (i.e., samples of voltage, current, temperature, relative discharge and charge time of the battery) and uploads them to the cloud;
- Step ③: the SOC model is executed on edge in the EV to compute SOC estimation, and the SOH model deployed in the cloud leverages the latest uploaded battery measurements to compute the updated SOH;
- Step ④: when the pre-defined conditions for retraining the SOC model are met (e.g., SOH reduced by 1%, or fixed time interval elapsed), the cloud uses the updated measurement data of the battery to retrain the SOC model;
- Step ⑤: the retrained SOC model is transferred from the cloud to the EV edge node;
- Step ⑥: the EV edge node updates the deployed SOC model and goes back to step ② to execute the following steps repeatedly.

#### IV. EXPERIMENTS

The goal of the proposed solution is not to propose a complete description of the digital twin, but to rather discuss the effectiveness of the proposed SOC-SOH models organization in the context of battery digital twin construction. For this reason, determining the optimal machine learning algorithms

or the optimal cloud/edge architecture is beyond the scope of this work. This section applies the proposed SOC-SOH model structure to estimate its effectiveness in reflecting battery dynamics and the impact of aging.

#### A. Experiments Setup

1) *Dataset Description:* The dataset used in this work is a NASA Ames Prognostics Center of Excellence (PCoE) dataset containing traces of 28 lithium-cobalt-oxide (LCO) 18650 batteries with 2.1Ah nominal capacity and 4.2V rated voltage that were continuously operated using a sequence of charging and discharging currents at different temperatures emulating random real-world battery usage [22]. The dataset contains information about battery current, voltage, temperature, relative time in each reference discharge and charge cycle, and time during the whole measurements are tracked for both reference cycles and randomized cycles (also called Random Walk, RW). Reference charging and discharging cycles were performed periodically to provide reference points for computing the battery's total available capacity, which allows deriving the battery SOH in different moments of battery lifetime.

2) *Data Processing:* In this work, we selected 10 out of the 28 battery cells in the dataset, discarding the cells with malformed data (e.g., too low operating temperature or nominal capacity). We keep the reference discharge data in the measurements and remove the random walk portions. For example, after removing all the random walk cycles, RW9 data contains 79 reference discharge cycles, and RW13 has 22 reference discharge cycles, spanning the entire lifetime of battery measurements. We adopt all reference discharge data to construct our dataset in this work because we can use it to build both the SOH and SOC models and retrain the SOC model when SOH decreases.

Within each reference discharge cycle, we conduct the Coulomb Counting method to calculate the residual capacity corresponding to each time step in the discharge phase and compute the total available discharge capacity at the end of the discharge cycle. Then we convert the residual capacity during one reference discharge to the corresponding value of SOC and use each reference discharge's total available capacity to derive the SOH. Notice that the SOH value in one discharge cycle keeps constant; it only updates from discharge cycle to discharge cycle.

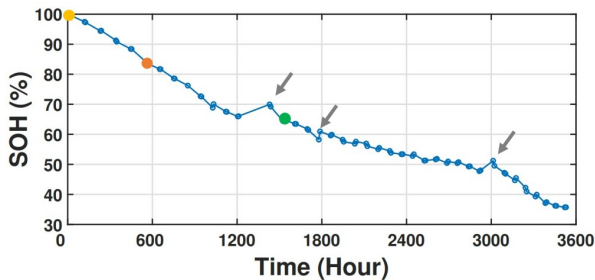


Figure 4: SOH degradation profile extracted from the RW9 dataset. Arrows points where the monotonically decreasing nature of SOH is not reflected due to measurement errors.

#### 3) Motivation of The Adopted Digital Twin Strategy:

Figure 4 indicates the extracted SOH decreasing profile during the whole lifetime of battery cell RW9, and each dot on the curve corresponds to a value of SOH computer from a reference discharge for RW9 provided in the dataset.

Each dot in Figure 4 corresponds to a different voltage discharge curve, and different SOH values imply significant differences in the discharge curve. Figure 5 exemplifies the discharge voltage curve for the colored points in Figure 4 (corresponding to SOH 100%, 85% and 65%). The curves reflect that when SOH decreases, the overall discharge time is shortened due to the reduction of the total available capacity of the battery. Note that there is also a change in the downward shape of the overall voltage discharge profile. All these changes support the idea that it is necessary to retrain the SOC model, and that a SOC model can not be accurate throughout the whole battery lifetime.

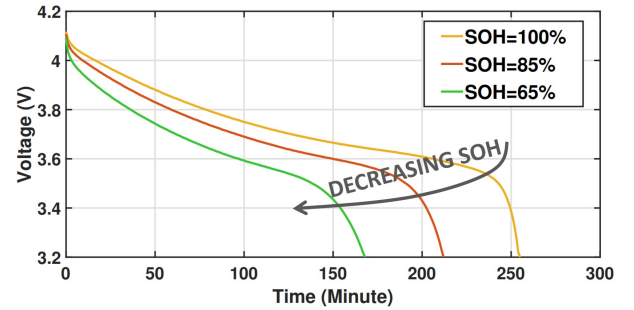


Figure 5: Voltage profile of reference discharge at different SOH values: a lower value of SOH leads to a reduction of overall discharge time (due to the reduction of the total available capacity) and to a change in the downward shape of the overall voltage discharge profile.

4) *Data Cleaning:* After computing the SOH value for each reference discharge cycle, it was possible to identify some irregular cycles break in the monotonic decreasing nature of SOH due to the measurement errors (e.g., the points highlighted by the arrows in Figure 4). We removed the erroneous samples from the data traces for all discharge cycles. For instance, over 79 reference discharge cycles, RW9 has nine samples with these erroneous behaviors, and they have been removed from the SOH profile to build the final dataset.

Ultimately, our dataset consists of the following columns, current, voltage, temperature, relative time, SOC, and SOH; the first four are used as inputs to train the models in our proposed digital twin, and the last two are used as outputs of the models.

#### B. Machine Learning Techniques Used in This Work

In this work, to perform the SOH and SOC estimation, we have experimented with different machine learning-based models. Choosing the optimal model configuration (models types and architectures) is beyond the scope of this work; the focus of this work is to clarify the architecture of our proposed digital twin and its operation mechanism, and in our future work, we will explore the optimal models' configuration based



on different design requirements. Machine learning methods adopted in this work include:

a) *Random Forest (RF)*: A RF comprises many decision trees: when performing the classification task, new input samples enter, and each decision tree in the forest classifies them separately; the classification results chosen by most decision trees is taken as the final result [23].

b) *Light Gradient Boosting (LGB)*: LGB is a gradient learning system based on decision trees and the boosting concept [24]. It employs a leaf-wise growth approach with depth limitations and uses histogram-based techniques to speed up the training process and decrease memory usage.

c) *Deep Neural Network (DNN)*: The DNN adopted in our work is a feed-forward artificial neural network with more than one layer of hidden units between its inputs and outputs layers. Each hidden unit typically employs the logistic function to map its total input from the layer below to the scalar state that it sends to the layer above [25]. The neural network architecture used in this experiment is based on input data connected with seven hidden layers and the output layer with a linear activation (since it is a regression problem) function.

During the training phase of all the methods mentioned above, the verification method uses k-fold cross-validation. The SOH and SOC estimation results are compared based on different models' accuracy, training time, and inference time.

### C. SOH Estimation

1) *Feature Selection*: There is a strong relationship between the relative time discharge time and the corresponding voltage, current, and temperature values, but most of the existing work in the literature ignores this feature [7]. In this work, we found that the inclusion of the relative time feature significantly improved the accuracy of the models: e.g., RF accuracy (RMSE) of SOH estimation reaches 1.773% when considering the relative time feature, while without the relative time, it increases to 116.20%.

2) *SOH Models Comparison*: Table I clearly shows that the RF has the best performance with an RMSE = 1.773%, followed by a slightly worse accuracy of LGB. The performance of the DNN is worse, with an RMSE = 7.114%. The low accuracy of DNN is due to the limited size of our training dataset: machine learning generation techniques to do data augmentation for training the deep learning models may improve the accuracy. However, our experimental results demonstrate that simple machine learning techniques, e.g., RF and LGB, are completely capable of estimating the SOH.

Table I: Different SOH Models Performance

Model	T + I (s)	RMSE(%)	MAE(%)
<b>RF</b>	13.704 + 0.470	1.773	0.603
<b>LGB</b>	10.629 + 0.977	2.310	0.791
<b>DNN</b>	399.582 + 1.439	7.114	1.696

Concerning time consumption (**T** represents training time, and **I** indicates inference time), the LGB method took the shortest time to train the model with only about 11 seconds, followed by the RF, that took approximately 14 seconds.

Although the SOH model in our proposed digital twin is pre-trained offline and training time is not a decisive factor, too long training time can delay the deployment of the digital twin. For the comparison of SOH models in terms of time consumption, inference time is more critical than training time because this is computed in real-time in the cloud, but inference time of the three models is comparable (around 1s).

### D. SOC Estimation

1) *Feature Selection*: The same features used to train the SOH have been used to train the SOC models after being scaled using Min-Max normalization.

2) *SOC Models Comparison*: The SOC models investigated in this work show a similar performance as SOH models. Table II and III show the performance of SOC models trained at the beginning (SOH=100%) and when the SOH decreases to 70% to conduct the SOC estimation at the same SOH level. RF and LGB show a high degree of accuracy, followed by DNN. Furthermore, the training time of DNN is relatively too long compared to the other two (approx. two orders of magnitude), and such a training time will weaken the overall performance of the digital twin, considering that the SOC model in our proposed digital twin architecture is constantly retrained and updated. Thus, the DNN method is proved to be less suitable than the others for the digital twin context.

Table II: SOC model performance at SOH = 100%

Model	T + I (s)	RMSE(%)	MAE(%)	Max Err.(%)
<b>RF</b>	0.825 + 0.034	0.038	0.132	1.442
<b>LGB</b>	0.140 + 0.005	0.024	0.125	1.010
<b>DNN</b>	67.002 + 0.138	0.609	1.783	2.478

Table III: SOC model performance at SOH = 70%

Model	T + I (s)	RMSE(%)	MAE(%)	Max Err.(%)
<b>RF</b>	0.622 + 0.017	0.809	0.624	3.873
<b>LGB</b>	0.097 + 0.005	0.599	0.549	2.993
<b>DNN</b>	29.97 + 0.068	2.775	1.380	5.071

3) *SOC Model Estimation at Different SOH Cases*: In this section, we illustrate why the SOC model needs to be retrained periodically by comparing the results of SOC estimation for the latest trained SOC model and the outdated SOC models.

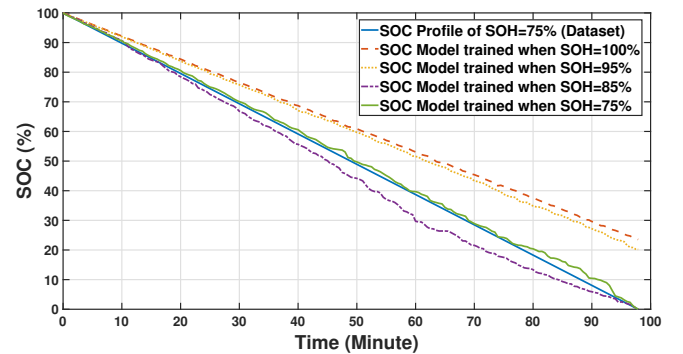


Figure 6: SOC discharge profiles comparison among different SOC Random Forest models trained in different SOH cases.

Figure 6 summarizes the difference in SOC values estimated using the latest SOC RF model trained based on discharge data

with SOH equal to 75% and outdated SOC models trained based on SOH equal to 100%, 95%, 85%, and 75% when SOC estimation is performed at the time of SOH = 75%. The solid blue dot line is the actual SOC discharge profile, and only the green circle line is closest to it (RMSE of these two lines is only 0.809%, and MAE is 0.624%), which is the value obtained from the SOC model trained using the discharge data at SOH=75%. The SOC values obtained with the outdated SOC models are far from the correct values, and the older the model, the larger the error. This proves that it is necessary to retrain and update the SOC model periodically.

The first row in the Table IV shows the errors of the SOC estimation at SOH=75% when using the SOC model trained with SOH=100%, corresponding to the difference between the orange star line and blue dot line shown in the Figure 6. The other two models also show significant errors due to the use of too outdated SOC models. In particular, the MAE of DNN increases to 25.315%. Considering that the range of SOH values is 0% to 100%, and the normal SOH operating range of the battery is only 80% to 100%, an error of more than 20% is entirely unacceptable.

Table IV: SOC model trained at SOH = 100% used at SOH=75%

Model	RMSE (%)	MAE (%)	Max Error (%)
RF	184.450	11.731	23.516
LGB	188.917	11.858	23.666
DNN	859.818	25.315	54.067

## V. CONCLUSIONS

This work proposed a digital twin architecture to estimate battery SOH and SOC during the operation of an EV. The digital twin includes a model of SOH, built once for all on historical data, and a model of SOC that is periodically retrained to reflect the internal aging dynamics of the EV battery. The work focused on proving the impact of the proposed models on the relevance of preserving the relationship between SOC and SOH, and of adjusting models to the evolving conditions of a specific EV. The proposed solution has been exemplified on a public dataset to highlight the motivation experimentally and to prove its effectiveness in terms of extensibility to a variety of machine learning data-driven models and compatibility with onboard execution during EV operation (maximum inference times in the order of 1s, and retraining time for the SOC model lower than 1s). Future work will discuss the infrastructure supporting these models, including edge and cloud collaborative framework considerations.

## REFERENCES

- [1] K. Swindells, "The data behind the electric car market's explosive growth," 2021. [Online]. Available: <https://elitetraveler.com>
- [2] EV-Volumes, "Global EV sales for 2021," 2022. [Online]. Available: <https://www.ev-volumes.com>
- [3] International Energy Agency, "Technology roadmap – electric and plug-in hybrid vehicles," 2009. [Online]. Available: <https://www.iea.org>
- [4] W. Li, M. Rentemeister, J. Badeda, D. Jost, D. Schulte, and D. U. Sauer, "Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation," *Journal of Energy Storage*, vol. 30, p. 101557, 2020.
- [5] Z. Chen, C. C. Mi, Y. Fu, J. Xu, and X. Gong, "Online battery state of health estimation based on genetic algorithm for electric and hybrid vehicle applications," *Journal of Power Sources*, vol. 240, pp. 184–192, 2013.
- [6] Y. Chen, E. Macii, and M. Poncino, "A circuit-equivalent battery model accounting for the dependency on load frequency," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017*. IEEE, 2017, pp. 1177–1182.
- [7] C. Vidal, P. Malysz, P. Kollmeyer, and A. Emadi, "Machine learning applied to electrified vehicle battery state of charge and state of health estimation: State-of-the-art," *IEEE Access*, pp. 52 796–52 814, 2020.
- [8] B. Saha and K. Goebel, "Battery data set," Tech. Rep., 2007. [Online]. Available: <http://ti.arc.nasa.gov/project/prognostic-data-repository>
- [9] H. Zhang, Q. Miao, X. Zhang, and Z. Liu, "An improved unscented particle filter approach for lithium-ion battery remaining useful life prediction," *Microelectronics Reliability*, vol. 81, pp. 288–298, 2018.
- [10] J. C. Álvarez Antón, P. J. García Nieto, C. Blanco Viejo, and J. A. Vilán Vilán, "Support vector machines used to estimate the battery state of charge," *IEEE Transactions on Power Electronics*, vol. 28, no. 12, pp. 5919–5926, 2013.
- [11] L. Yao, Z. Fang, Y. Xiao, J. Hou, and Z. Fu, "An intelligent fault diagnosis method for lithium battery systems based on grid search support vector machine," *Energy*, vol. 214, p. 118866, 2021.
- [12] W.-Y. Chang, "Estimation of the state of charge for a LFP battery using a hybrid method that combines a RBF neural network, an OLS algorithm and AGA," *International Journal of Electrical Power & Energy Systems*, vol. 53, pp. 603–611, 2013.
- [13] W. He, N. Williard, C. Chen, and M. Pecht, "State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter-based error cancellation," *International Journal of Electrical Power & Energy Systems*, vol. 62, pp. 783–791, 2014.
- [14] L. Xu, J. Wang, and Q. Chen, "Kalman filtering state of charge estimation for battery management system based on a stochastic fuzzy neural network battery model," *Energy Conversion and Management*, vol. 53, no. 1, pp. 33–39, 2012.
- [15] X. Tang, K. Liu, X. Wang, F. Gao, J. Macro, and W. D. Widanage, "Model migration neural network for predicting battery aging trajectories," *IEEE Transactions on Transportation Electrification*, vol. 6, no. 2, pp. 363–374, 2020.
- [16] E. Chemali, P. J. Kollmeyer, M. Preindl, Y. Fahmy, and A. Emadi, "A convolutional neural network approach for estimation of Li-Ion battery state of health from charge profiles," *Energies*, vol. 15, no. 3, 2022.
- [17] X. Sui, S. He, S. B. Vilsen, J. Meng, R. Teodorescu, and D.-I. Stroe, "A review of non-probabilistic machine learning-based state of health estimation techniques for lithium-ion battery," *Applied Energy*, vol. 300, p. 117346, 2021.
- [18] B. Wu, W. D. Widanage, S. Yang, and X. Liu, "Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems," *Energy and AI*, p. 100016, 2020.
- [19] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6730–6739, 2018.
- [20] X. Qu, Y. Song, D. Liu, X. Cui, and Y. Peng, "Lithium-ion battery performance degradation evaluation in dynamic operating conditions based on a digital twin model," *Microelectronics Reliability*, vol. 114, p. 113857, 2020.
- [21] J.-H. Lee and I.-S. Lee, "Lithium battery soh monitoring and an soc estimation algorithm based on the soh result," *MDPI Energies*, vol. 14, no. 15, 2021.
- [22] B. Bole, C. S. Kulkarni, and M. Daigle, "Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomized use," in *Annual Conference of the PHM Society*, vol. 6, no. 1, 2014.
- [23] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [24] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, vol. 30, 2017.
- [25] G. Hinton, L. Deng, D. Yu, G. E. Dahl, and eta, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.