The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2023

Volume 22, Pages 135-141

ICBASET 2023: International Conference on Basic Sciences, Engineering and Technology

State of Health Estimation for Li-Ion Batteries Using Machine Learning Algorithms

Yunus KOC
Istanbul Technical University

Abstract: As an energy storage system, Li-Ion batteries have many applications from mobile devices to vehicles. No matter what application they are used in, Li-Ion batteries lose performance over time, and this negatively affects the user experience in terms of both comfort and safety. For this reason, it is extremely important to estimate state of health (SOH) of Li-Ion batteries and to use the batteries accordingly. In this study, examinations on the SOH estimation of batteries with different machine learning (ML) methods are included using Constant Current (CC) and Constant Voltage (CV) charge-discharge characteristics of the li-Ion batteries. Moreover, how the estimation performance changes by both short-term and long-term features is observed by using mutual information metric. According to results, the highest accuracy on SOH estimation is achieved when long-term features are used with Bayesian Ridge Regression. When the short-term features are used, the accuracy of Bayesian Ridge Regression is dramatically reduced, and Random Forest Regression provides highest performance.

Keywords: Regression, State of health estimation, Machine learning, Feature extraction, Li-ion batteries

Introduction

Lithium-ion batteries are being developed day by day and are important energy storage systems used in many application areas. It is also extremely critical to know how much the performance of batteries changes during use, in other words, how much their capacity is reduced. Battery management systems can optimize battery usage especially through this information and ensure the use of the battery in the longest possible term (Chen et al., 2023). Moreover, SOH is significant parameter showing capacity degradation which is also used for evaluating battery state of safety (Li et al., 2022).

There are lots of data driven methods used for extracting state of health (SOH) of the batteries in literature. Especially machine learning-based approaches have become very popular recently. For instance, in a study using tree-based algorithms deals with predicting lifetime of the li-ion batteries using early cycle data and the study also includes analyze of feature importance using Kendall’s tau and Spearman correlation methods (Çelik et al., 2022). Another one of the SOH estimation studies is the online estimation with DSMTNet, one of the deep learning methods (Wang et al., 2022). They have achieved SOH computing in just 0.14 sec. which is satisfactory result for real time applications. Linear regression analysis with multiple charge and discharge features (Agudelo et al., 2023), Support Vector Regression combining with Voltage-Capacity (VC) model (Zhang et al. 2022), incorporating the DNN into a Kalman filter (Tian et al., 2021), partial analysis of charging curve (Lyu et al., 2021) and wavelet neural networks with genetic algorithm (Chang et al., 2021) were also proposed for SOH estimation of Li-Ion batteries.

In this study, SOH estimation of the batteries is made by using well-known machine learning methods such as Random Forest Regression, Decision Tree Regression, Ridge Regression, Bayesian Ridge Regression, Support Vector Regression and Extreme Gradient Boost Regression. For each model, battery features are extracted in case of long-term and short-term characteristics from CC-CV charging/discharging curves. All details of each step is described in following sections.

Battery Life Cycle Dataset

To perform SOH estimation, or in other words, capacity degradation analysis, a battery data set in which charge/discharge cycles are made under certain conditions is needed. In this study, the HNEI dataset (Hawaii Natural Energy Institute, 2014), in which more than 1000 cycles were made and prepared on 18650 Li-Ion batteries, was used. Details of the dataset are presented in below table (Table 1).

Table 1. HNEI Li-Ion Battery Life Cycle Dataset

Cell & Cycle Parameters	Description
Anode	Graphite
Cathode	NMC-LCO
Capacity	2800 mAh
Form Factor	18650
Temperature	25°C
Max SOC	100
Min SOC	0
Charge Rate	0.5 C (1.4 A)
Discharge Rate	1.5 C (4.2 A)
Max Voltage	4.35 V
Min Voltage	3.0 V
Nominal Voltage	3.8 V

The measured characteristics of the battery during the life cycle are as follow; *Current (A), Min-Max Current (A), Voltage (V), Min-Max Voltage (V), Charge Capacity (Ah), Discharge Capacity (Ah), Charge Energy (Wh), Discharge Energy (Wh).*

Feature Extraction

In this section, the features are observed as *long-term feature* which needs to get complete charge/discharge cycle and *short-term feature* which needs to get a small portion of charge/discharge curve. All features are extracted in CC region of Charge/Discharge Curve. Each of features are described in following sections.

Long-Term Features

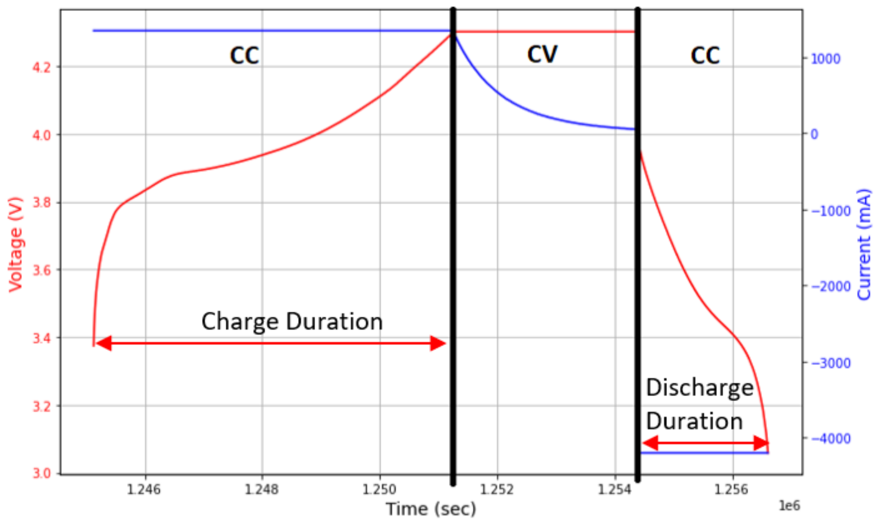

Figure 1. CC and CV regions of complete charge/discharge cycle

1) Charge/Discharge Duration

Charge and discharge duration are extracted from the duration of charge and discharge states during the CC region for each cycle.

2) Charge/Discharge Current Count

Current accumulation of CC charge and discharge states is another key feature regarding to overall capacity. Because these features are extracted in CC domain, it is easy to handle current count operation using below formula for each charge and discharge states.

$$Charge\ Count = Charge\ Duration \text{ x Charge Current (CC)} \tag{1}$$

$$Disharge\ Count = Discharge\ Duration \text{ x Discharge Current (CC)} \tag{2}$$

3) Charge/Discharge Voltage Integration

Because the voltage-time curve of the battery changes by long cycles, the area under the curve is also significant clue about state of the health analysis. On this issue the area can be calculated as follow:

$$Charge\ Voltage\ Integration = \int_{T_1}^{T_2} V_T^{Charge}\, dt \tag{3}$$

$$Discharge\ Voltage\ Integration = \int_{T_1}^{T_2} V_T^{Discharge}\, dt \tag{4}$$

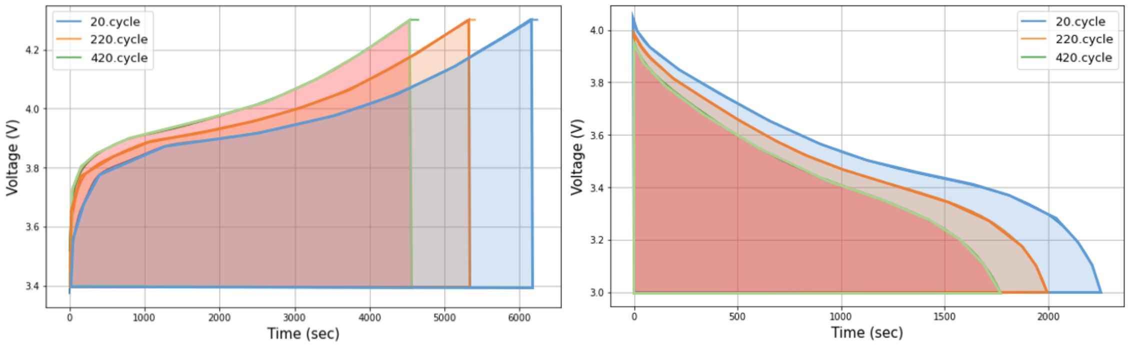

Figure 2. Area under the voltage-time charge/discharge curve for different cycles

In equation (3) and (4), T_1 and T_2 correspond to start and end time of the charge and discharge states, V_T shows terminal voltage of the battery, respectively.

4) Kurtosis and Skewness

Before the calculation of kurtosis and skewness feature, it is needed to calculate difference of charge and discharge voltage curve. On this issue, it is handled by subtracting discharge voltage values from reversed form of charge curve as below figure. By this way, skewness and kurtosis features can be calculated using third and fourth order central moment of the voltage difference respectively as below (Cheng et al., 2021).

$$Skewness = \frac{1}{n} \sum\nolimits_{i=1}^{n} \left[\left(\frac{X_i - \mu}{\sigma} \right)^3 \right] \tag{5}$$

$$Kurtosis = \frac{1}{n} \sum\nolimits_{i=1}^{n} \left[\left(\frac{X_i - \mu}{\sigma} \right)^4 \right] \tag{6}$$

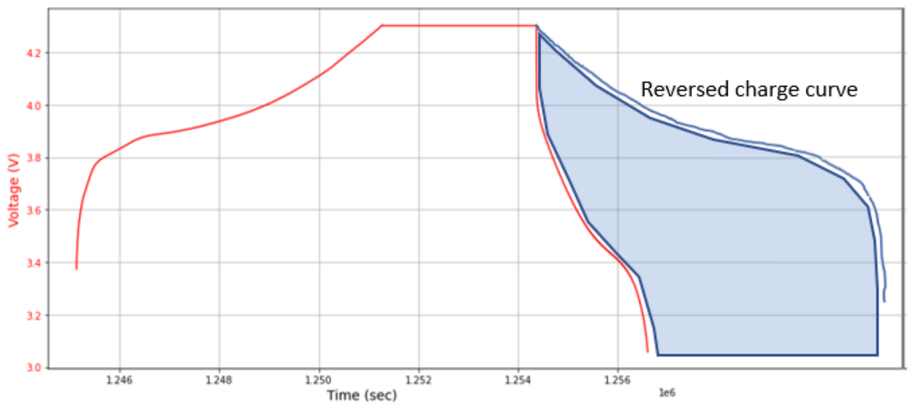

Figure 3. Voltage differences between charge and discharge states

In equation (5) and (6); n is the number of samples, X_i is the value of the i th sample of voltage difference, μ and σ are the mean and variance of the voltage difference sequence, respectively.

Short-Term Features

While analyzing the short-term features, the nominal voltage value of the battery is considered as a base voltage. To partial analyze of charge/discharge curve, threshold voltage which maximize the time skew of voltage curve between first and last cycle is calculated by below formula:

$$V_{threshold}^{Charge} = \arg\max_{V}\left(\left|T_{first\ cycle}^{V_{charge}} - T_{last\ cycle}^{V_{charge}}\right|\right) \tag{7}$$

$$V_{threshold}^{Disharge} = \arg\max_{V}\left(\left|T_{first\ cycle}^{V_{discharge}} - T_{last\ cycle}^{V_{discharge}}\right|\right) \tag{8}$$

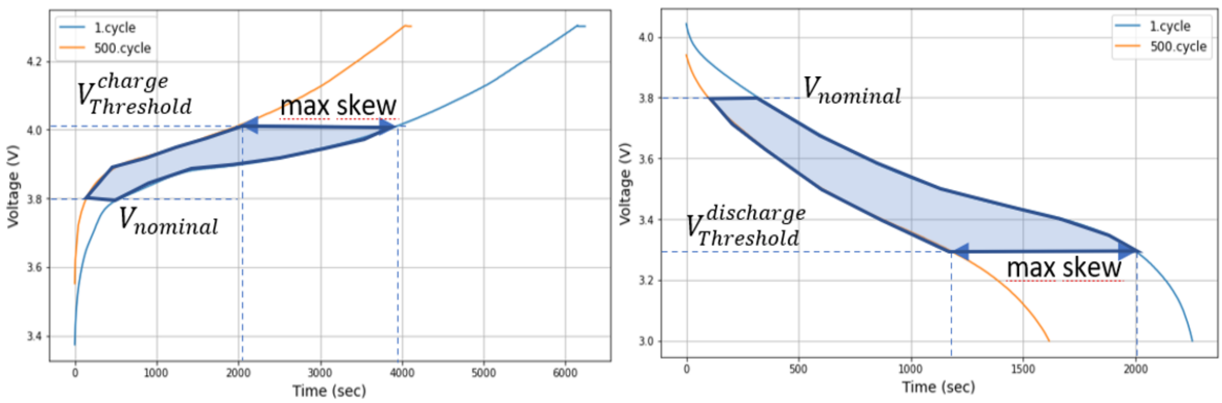

Figure 4. Partial analysis of charge/discharge curves based on nominal and threshold voltages

1) Charge/Discharge Nominal Duration

As described in Figure 4, charge and discharge duration features are calculated regarding to the area limited by nominal and threshold voltages. Thus, nominal charge and discharge durations are equal to duration from $V_{threshold}^{Charge}$ to $V\_nominal$ and and $V_{threshold}^{discharge}$ to $V\_nominal$ points respectively. By this way, there is no need to wait complete cycle.

2) Charge/Discharge Nominal Voltage Integration

Charge and discharge nominal voltage integration features are calculated by same way in equation (3) and (4) based on the area limited by threshold voltage and nominal voltage as illustrated in Figure 4.

3) Ohmic Response

Ohmic response can be defined as resistive effect of internal chemical to drawn current. It is also known from the (Koç et al., 2022) that as batteries age, their internal resistance increases accordingly. İncreased resistance also causes higher voltage to drop under the load. Hence, the increase in voltage drop across the cycle is caused by increased internal resistance or ohmic response, which is directly related to the health of the battery. That is why, the ohmic response is considered as a key feature and calculated by ohm law as below.

$$Ohmic\ Response = \frac{V_{drop}}{I_{discharge}} = \frac{V_{charge}^{end} - V_{discharge}^{beginning}}{I_{discharge}} \quad (9)$$

In equation (9), V_{drop} refers to the voltage drop experienced at the first moment of discharge. In this study, V_{charge}^{end} corresponds to 4.35 V, $I_{discharge}$ is 1.5C which equals to 4.2 A and $V_{discharge}^{beginning}$ decreases due to ohmic response.

State of Health Estimation

Feature Importance

As already described in previous sections, the features were classified as short and long term. While the short-term features seem more useful for time critical application, they need to ensure satisfactory results in machine learning method as they can. Long-term features are also expected to yield good results in terms of model accuracy. To understand the importance of all features, mutual information (Koç et al., 2021) scores are calculated as illustrated in below table.

Table 2. Mutual information scores of features (left side: long-term, right side: short-term)

Feature Name	Score	Feature Name	Score
discharge duration	4.0954	discharge nominal duration	3.1166
discharge current count	4.0934	charge nominal duration	3.0606
charge duration	3.6494	ohmic response	2.5726
charge current count	3.6467	charge nominal voltage integration	2.4771
charge voltage integration	2.4124	discharge nominal voltage integration	2.4126
discharge voltage integration	2.3427		
kurtosis	1.9288		
skewness	1.8855		

As shown in Table 2, long-term features have the highest scores with the disadvantage of taking a long time. On the other hand, short-term features provide superiority better scores compared to some of long-term features.

SOH Analyze Using Machine Learning Methods

Because the dataset has continuous numerical data, regression methods are preferred to handle SOH estimation. On this issue, Random Forest Regressor, Decision Tree Regressor, Ridge Regressor, Bayesian Ridge Regressor, Support Vector Regressor and Extreme Gradient Boost Regressor methods are used with default parameters defined in sklearn library. Before performing these methods, feature values were normalized using min-max normalization to map different ranges of the features into [0-1]. Each machine learning methods were trained and tested with both short-term and long-term features separately. Train and test sizes was defined as %75 and %25, respectively.

Results and Discussion

After training and testing steps, each method performances were obtained as below. All train and test results were calculated by using cross validation technique with 5 splits and 5 repeats to ensure more precise results.

Table 3. Performance Results (Left Side: Long-Term Features, Right Side: Short-Term Features)

Model	Train Accuracy (%)	Test Accuracy (%)	Model	Train Accuracy (%)	Test Accuracy (%)
Random Forest Regressor	98.3	92.6	**Random Forest Regressor**	98.9	**95.3**
Decision Tree Regressor	100	92.3	Decision Tree Regressor	100	93.0
Ridge Regressor	97.1	96.2	Ridge Regressor	80.6	80.1
Bayesian Ridge Regressor	99.3	**98.2**	Bayesian Ridge Regressor	86.6	81.1
Support Vector Regressor	94.9	89.3	Support Vector Regressor	82.9	79.0
Extreme Gradient Boost Regressor	100	95.3	Extreme Gradient Boost Regressor	100	93.1

As shown from the Table 3, tree-based methods like Random Forest Regression and Decision Tree Regression provides higher test accuracy even they used short-term features which yield small amount of data sample. When the long-term features are used in the model, the multicollinearity between features and targets can be higher compared to correlation with short-term features. Because Ridge Regression models are more sensitive to multicollinearity between independent variables, the performance dramatically reduced when the short-term features which provide less observation are used. On the other hand, Random Forest Regression can handle multiple ensemble methods and randomly samples the data during the train and test steps. By this way, Random Forest Regression can adapt itself during the training with short-term features and yields better results compared to long-term feature cases.

Conclusion

In this study, different machine learning methods are used to estimate SOH parameter of the battery with long-term and short-term features. According to the results, Bayesian Ridge Regression ensures highest estimation performance as %98.2 on long-term features and it can be preferred in accuracy critical applications. On the other hand, if time critical applications need to be handled, short-term features can also be used with Random Forest Regression method. For the future work, more efficient short-term features can be defined even they also ensures higher accuracy.

Acknowledgement

* This article was presented as oral presentation at the International Conference on Basic Sciences, Engineering and Technology (www.icbaset.net) held in Marmaris/Turkey on April 27-30, 2023.

*I would like to thank Token Financial Technologies Inc. for providing a variety of business tools and academic opportunities that encourage me to follow new researches.

References

Agudelo, B. O., Zamboni, W., Postiglione, F., & Monmasson, E. (2023). Battery state-of-health estimation based on multiple charge and discharge features. *Energy, 263*, 125637.

Celik B., Sandt R., dos Santos L.C.P., & Spatschek R. (2022) Prediction of battery cycle life using early-cycle data, machine learning and data management. *Batteries 2022. 8(12)*, 266.

Chang, C., Wang, Q., Jiang, J., & Wu, T. (2021). Lithium-ion battery state of health estimation using the incremental capacity and wavelet neural networks with genetic algorithm. *Journal of Energy Storage, 38*, 102570.

Chen, M., Ma, G., Liu, W., Zeng, N., & Luo, X. (2023). An overview of data-driven battery health estimation technology for battery management system. *Neurocomputing, 532*(1), 152-169.

Cheng, D., Sha, W., Wang, L., Tang, S., Aijun, M., Chen, Y., Wang, H., Lou, P., Lu, S., & Cao, Y. (2021). Solid-state lithium battery cycle life prediction using machine learning. *Applied Science, 11*, 4671.

Hawaii Natural Energy Institute (2014). *HNEI [Dataset]*. https://database.batteryarchive.org/public/dashboards

Koç, Y., & Çetin, M. (2021) Comparative evaluation of machine learning algorithms with parameter optimization and feature elimination for fraud detection. *International Conference on Electrical, Computer and Energy Technologies (ICECET),* 1-6.

Koç, Y., Doğru U. E., & Bilir, R. A. (2022). Evaluation of internal resistance methods for tracking battery state of health. *2022 4th Global Power, Energy and Communication Conference (GPECOM)*, 112-116.

Li, Y., Wang, W., Yang, X. G., Zuo, F., Liu, S., & Lin, C. (2022). A smart Li-ion battery with self-sensing capabilities for enhanced life and safety. *Journal of Power Sources*, *546*, 231705.

Lyu, Z., Gao, R., Li, X. (2021). A partial charging curve-based data-fusion-model method for capacity estimation of Li-Ion battery. *Journal of Power Sources, 483*, 229131.

Tian, J., Xiong, R., Shen, W., & Lu, J. (2021). State-of-charge estimation of LiFePO4 batteries in electric vehicles: A deep-learning enabled approach. *Applied Energy, 291*, 116812.

Wang, H., Li, J., Liu, X., Rao, J., Fan, Y., & Tan, X. (2022). Online state of health estimation for lithium-ion batteries based on a dual self-attention multivariate time series prediction network. *Energy Reports, 8*, 8953-8964

Zhang, Y., Liu, Y., Wang, J., Zhang, T. (2022). State-of-health estimation for lithium-ion batteries by combining model-based incremental capacity analysis with support vector regression. *Energy, 239*, 121986.

Author Information

Yunus Koç
Istanbul Technical University
Istanbul, Turkey
kocyun@itu.edu.tr

To cite this article:

Koc, Y. (2023). State of health estimation for Li-ion batteries using machine learning algorithms. *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 22,* 135-141.