Energy Proceedings
Vol 29, 2022
ISSN 2004-2965

Voltage-temperature Feature-based Capacity Estimation Method for Li-ion Battery Cell Combining Probability Density Function and Random Forests#

Difan Jia[1], Xuqi Fu[2], Zhanyao Ma[3], Xiaowu Zuo[1*]

1 School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, 200093.

2 College of Mechanical and Electrical Engineering, Soochow University, 215137.

3 Glasgow college, University of Electronic Science and Technology of China, 611730.

Corresponding author: Xiaowu Zuo (E-mail: zxw@usst.edu.cn)

ABSTRACT

Accurate capacity estimation is crucial to ensure operational safety of Li-ion battery. In this paper, a novel capacity estimation approach is proposed for Li-ion battery cell. Two voltage-related features on probability density function based incremental capacity curve and average temperature are extracted as healthy indicators. Regression between healthy indicators and capacity is constructed using random forests. Results show that the capacity estimation error could be controlled within 2.5% throughout the whole lifecycle of the battery.

Keywords: Li-ion battery, capacity estimation, data-driven, random forests.

NONMENCLATURE

Abbreviations	
CC	Constant Current
CV	Constant Voltage
IC	Incremental Capacity
OCV	Open Circuit Voltage
PDF	Probability Density Function
RF	Random Forests

1. INTRODUCTION

Accurate capacity estimation is helpful to avoid over-charge or over-discharge of Li-ion battery, thus is crucial for ensuring operational safety [1]. Existing capacity estimation researches can be generally divided into two categories, namely model-based and data-driven based. Model-based methods adopt electro-chemical model or equivalent circuit model to capture the dynamic property of the battery and uses state estimation method like extended Kalman filter or particle filter to realize capacity estimation [2,3]. However, the robustness of the method is relatively weak because unsuitable prior probability distribution or noise variance settings in filtering algorithms will result in divergence of such kind of approach. Data-driven methods directly exploit the relationship between capacity and healthy indicators extracted from recorded data like current, voltage, temperature, SOC, etc. to realize capacity estimation, which is more robust and straightforward due to avoidance of battery modeling [4,5]. IC analysis is one of the most useful data-driven approaches to extract healthy indicators [6]. However, IC analysis is sensitive to measurement noise, which would greatly hinder the identification of the peak on IC curve thus introducing distortion to healthy indicators. In addition, current researches mainly focus on voltage-related healthy indicators on IC curve. However, temperature could also provide valuable information about battery degradation.

In order to mitigate existing research gaps, this paper proposes a novel capacity estimation method which combines both voltage-related and temperature-related features. In order to depress the effect of measurement noise on feature extraction from IC curve, PDF approach is used to construct the IC curve from statistical view. Finally, RF is used to construct the regression model between capacity and healthy indicators.

The reminder of this paper is organized as follows. Section 2 gives a brief introduction of the Oxford battery dataset for verification. Section 3 details the capacity estimation method together with basic knowledge about theoretical basis including PDF, IC and RF used in this paper. Section 4 verifies the effectiveness of the proposed method and Section 5 concludes the whole paper.

2. OXFORD BATTERY DATASET

The Oxford battery dataset contains measurements of battery ageing data from 8 small Li-ion pouch cells manufactured by Kokam with rated capacity 740mAh [7].

The negative electrode material of the pouch cells is graphite and the positive electrode material is a blend of lithium cobalt oxide and lithium nickel cobalt oxide. The cells were all tested in a thermal chamber at 40°C and cycled using a CC-CV charging profile, followed by a discharging process obtained from the urban Artemis profile. Characterization tests including 1C (current=740mA) and pseudo-OCV (current=40mA) charging and discharging tests were conducted every 100 cycles of drive cycles. The recorded data includes voltage, temperature and charge. The calibrated capacity can be derived from the average of accumulated charge in pseudo-OCV charging and discharging tests. In this paper, 1C charging data of cell #3 is used to train the model and 1C charging data of cell #1 is used to verified the effectiveness and generalization ability of the proposed method.

3. CAPACITY ESTIMATION FRAMEWORK

The general capacity estimation framework is as follows. Firstly, the PDF-based IC curve is constructed and the location and amplitude of the second peak are extracted as the voltage-related healthy indicators. Then the average temperature during charging process is calculated as the temperature-related healthy indicator. The RF based regression model, which could combine voltage and temperature related healthy indicators automatically, uses the above three healthy indicators as input and capacity as output. The RF model More details can be found below.

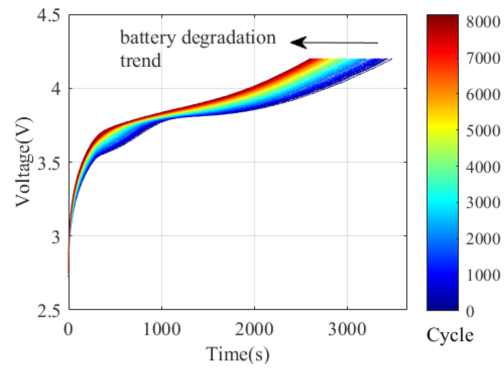
3.1 Voltage-related healthy indicators extraction

The PDF-based IC curve can be constructed using voltage data during the whole charging process with voltage as *x*-axis and PDF as *y*-axis. Intuitively, PDF can be regarded as the normalized frequency. A higher PDF indicates a more frequent appearance. Because PDF methods substituting the division calculation (whose denominator maybe close to zero) in original IC analysis by statistical cumulative counting, its robustness is greatly improved. PDF can be constructed using following kernel density formula [8]:

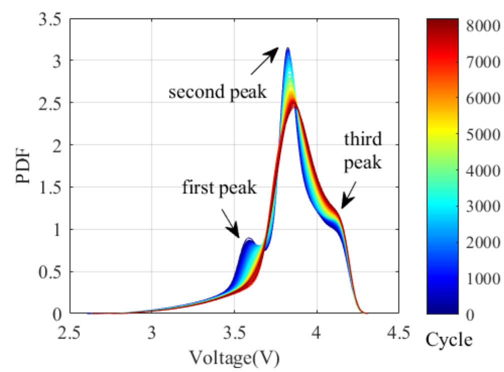
$$\hat{f}_h(x) = \frac{1}{nh}\sum\nolimits_{i=1}^{n} K(\frac{x-x_i}{h}) \tag{1}$$

where $x_1, x_2, \ldots, x_n$ are random samples from an unknown distribution, which are voltage sampling data here. n is the sample size. $K(\cdot)$ is the kernel smoothing function and normal kernel is used in this paper. h is the bandwidth.

Fig.1 demonstrates the voltage sampling data together with corresponding PDF-based IC curve during 1C charging process for cell #3 throughout whole lifecycle. It can be seen from Fig.1(a) that as the battery degrades, the charging process shortens, which indicates the decreasing capacity. From Fig.1(b), it can be found that there are three peaks on the PDF-based IC curve. However, the first peak will disappear when the battery is deeply aged and the third peak is relatively flat and not easy to extract. Thus, only the second peak is used for feature extraction. The second peak appears around 3.85V, which means voltage with around 3.85V appears most frequently. It corresponds to the plateau in original voltage curve. The monotonous property that the second peak gradually shifts rightward and downward as battery degrades provides valuable information for capacity estimation. Thus, the location and amplitude of the second peak are extracted as voltage-related healthy indicators. Fig.2 shows the relationship between voltage-related healthy indicators and capacity.

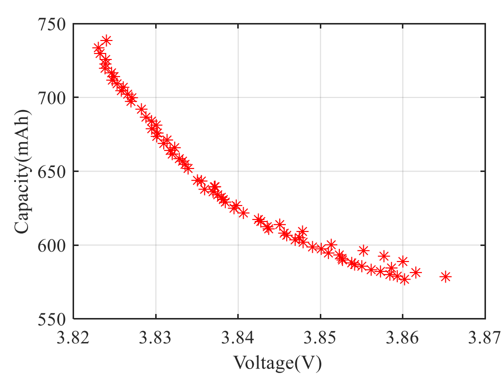


(a) Voltage

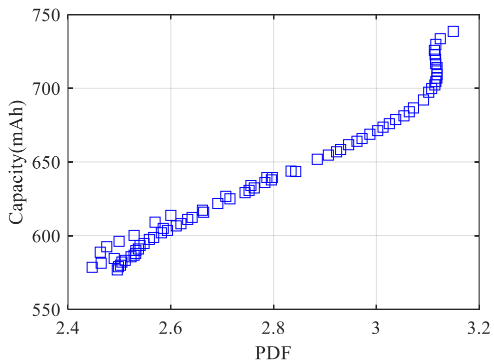


(b) PDF-based IC curve

Fig.1 Voltage sampling data together with corresponding PDF-based IC curve during 1C charging process for cell #3 throughout whole lifecycle



(a) Voltage location of second peak

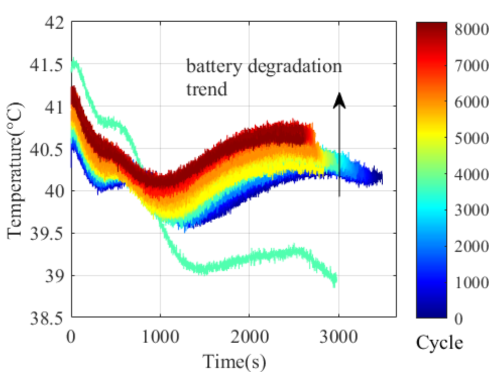


(b) PDF amplitude of second peak

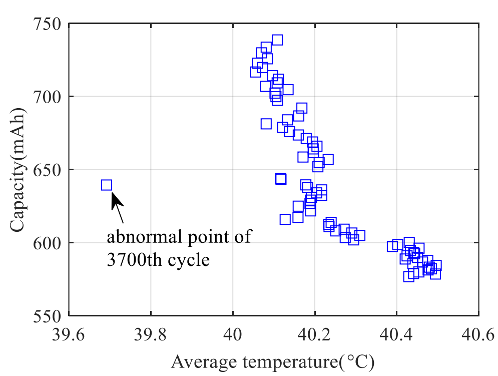
Fig.2 Relationship between voltage-related healthy indicators and capacity

3.2 Temperature-related healthy indicators extraction

Besides voltage, temperature data could also provide valuable information about battery state-of-health. Fig.3 shows the change of temperature for cell #3 during 1C charging process throughout whole lifecycle. It can be seen that the temperature gradually increases as the battery degrades, which can be explained by the increasing ohmic resistance. Except the abnormality of temperature data for the 3700th cycle, all the temperature data shows relatively consistent increasing trend with decreasing capacity. Thus, average temperature is extracted as the third healthy indicator.



(a) Temperature



(b) Average temperature

Fig.3 Change of temperature for cell #3 during 1C charging process throughout whole lifecycle

3.3 Random forests

RF is an integrated machine learning method that combines multiple decision trees to produce repeated prediction results for the same question. For each tree, the RF method performs self-help sampling, so that the calculation of error estimation can be based on the out-of-bag sampling data. When generating a tree, each node of the tree is randomly generated, and the segmentation variables of each node are generated by a small number of randomly selected variables. Finally, a lot of decision trees will be produced, so it is called "random forests". The RF used for regression averages the results of these trees to obtain the predicted value.

The steps of RF regression algorithm are:

(1) Bootstrap is used to extract m self-help sampling sets from N original samples to construct m regression trees. The unselected samples form m out-of-bag data sets.

(2) At each node of each tree, k partition variables ($k<p$) are randomly selected from all p explanatory variables, in which the optimal branch is selected according to the branch optimality criterion.

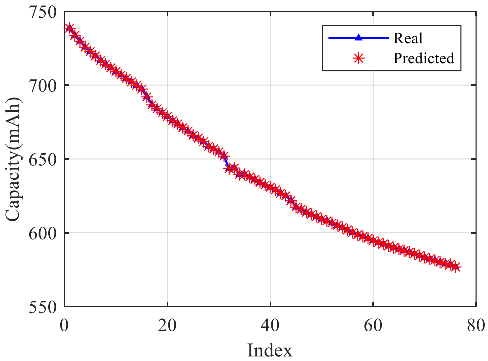
(3) Each regression tree starts top-down recursive branching until the segmentation termination condition is met.

The advantages of RF algorithm include fast learning process and strong robustness to the noise in the data set. There is no need to reserve some additional data for cross validation, and the effect of the algorithm is evaluated by using out-of-bag data. In addition, it is insensitive to multicollinearity.

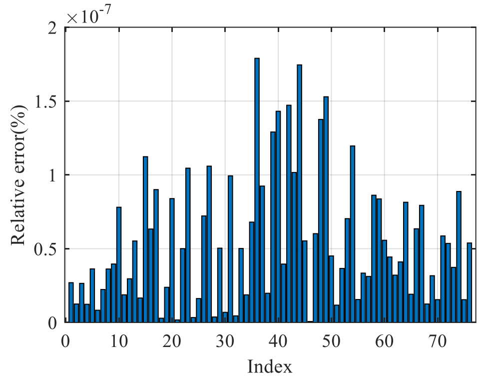
More detailed mathematical deduction of RF can be found in Ref. [9].

4. VERIFICATION OF THE PROPOSED METHOD

The proposed method is trained by the data of cell #3 and then the trained model is verified using the data of cell #1. This practice ensures that the data of cell #1 is totally unseen for the trained model during training process, which is used for fair generalization ability assessment during test process. The estimation performance on training dataset of cell #1 and testing dataset of cell #3 are shown in Fig.4 and Fig.5 respectively. It can be seen that thanks to the powerful regression capability of RF, the estimation error on training dataset is practically neglectable. The estimation absolute relative error on data of cell #1 is controlled within 2.5%, which verifies the effectiveness and generalization of the proposed method.

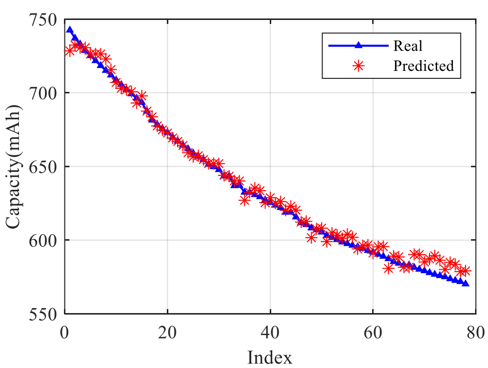


(a) Real and predicted capacity

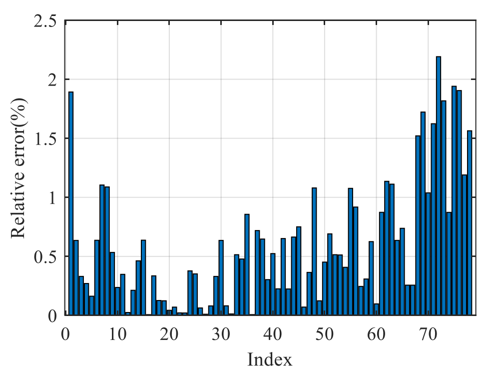


(b) Relative error

Fig.4 Estimation performance of the proposed model on training dataset, namely cell #3



(a) Real and predicted capacity



(b) Relative error

Fig.5 Estimation performance of the proposed model on test dataset, namely cell #1

5. CONCLUSION

In this paper, a novel capacity estimation method is proposed. Two voltage-related healthy indicators and one temperature-related healthy indicator are extracted from voltage and temperature data. Regression mapping from healthy indicators to capacity is constructed using RF. Result shows that the proposed method could control the capacity estimation error below 2.5%.

REFERENCE

[1] Nana Zhou, Xianhua Zhao, Bing Han, et al. (2022). A novel quick and robust capacity estimation method for Li-ion battery cell combining information energy and singular value decomposition. Journal of Energy Storage, DOI: 10.1016/j.est.2022.104263.

[2] Michel, P. H. , & Heiries, V. . (2015). An Adaptive Sigma Point Kalman Filter Hybridized by Support Vector Machine Algorithm for Battery SoC and SoH Estimation. Vehicular Technology Conference. IEEE.

[3] Topan, P. A. , Ramadan, M. N. , Fathoni, G. , Cahyadi, A. I. , & Wahyunggoro, O. . (2017). State of Charge (SOC) and State of Health (SOH) estimation on lithium polymer battery via Kalman filter. 2016 2nd International Conference on Science and Technology-Computer (ICST). IEEE.

[4] Xia, Z. , & Qahouq, J. . (2021). Lithium-ion battery ageing behavior pattern characterization and state-of-health estimation using data-driven method. IEEE Access, PP(99), 1-1.

[5] Jm, A. , Lei, C. , Dis, D. , Jm, A. , Gl, E. , & Rt, D. . An optimized ensemble learning framework for lithium-ion battery state of health estimation in energy storage system. Energy, 206.

[6] Agudelo, B. O. , Zamboni, W. , & Monmasson, E. . (2021). Application domain extension of incremental capacity-based battery soh indicators. Energy, 234(4), 121224.

[7] Crb, A. , Mrr, B. , Mtb, C. , Pgb, B. , & Dah, A. . (2017). Degradation diagnostics for lithium ion cells. Journal of Power Sources, 341, 373-386.

[8] Feng, X. , Li, J. , Ouyang, M. , Lu, L. , Li, J. , & He, X. . (2013). Using probability density function to evaluate the state of health of lithium-ion batteries. Journal of Power Sources, 232(18), 209-218.

[9] Breiman, L. . (2001). Random forest. Machine Learning, 45, 5-32.