

# Capacity estimation based on incremental capacity and Gaussian process regression for retired lithium-ion batteries

Wei Sun, Peng Jin \*, Xiawei Liao, Xiaoming Ma and Bo Yang

School of Environment and Energy, Peking University Shenzhen Graduate School, Shenzhen, China

**Abstract.** Fast capacity estimation for retired batteries is necessary when batteries are recycled for echelon utilization. Here, a fast capacity estimation method is proposed for retired LiFePO<sub>4</sub> battery. First, a full survey of battery pack and cells degradation after a long period of service is studied. Then the filtered ICA is used to study degradation variation phenomenon of retired batteries, the relationship between IC curve feature and remaining capacity was studied. Finally, a fast capacity estimation using incremental capacity and Gaussian process regression is proposed. Our results show high efficiency and accuracy of the proposed method.

## 1 Introduction

Recently, electric vehicles (EVs) have become a significant trend in many countries for the reason of concerns about energy security and environmentally friendly [1]. However, lithium-ion batteries, which is one of the core components in EVs, need to be transposed to guarantee safe operation and sufficient mileage when their capacity drops to 80% [2]. The retired batteries echelon utilization can prolong the useful battery life-span, thus reducing harmful environmental effects and resources conservation [3]. However, for their inconsistency, retired batteries cannot be firsthand recycled for echelon utilization [4], so the capacity estimation and screening methods of retired batteries are of great significance to the echelon utilization of retired batteries and the healthy development of electric vehicles.

In order to analyze degradation mechanism of large scale of cells and estimate their state of health (SOH), incremental capacity analysis (ICA) is employed [5-7]. IC curve cannot only indicate the subtle and slowing-changing electrochemical feature of the electrochemical reactions inside cells, it can also depict current acceptance capability in various potentials. In view of the merit of convenient properties for modeling without specific function, Gaussian process regression has been widely adopted in the field of machine learning and statistics [8].

In this thesis, a rapid capacity estimation is applied for retired LiFePO<sub>4</sub> battery. First, a comprehensive survey of a battery pack aging and cells after a long period of employment is studied.

Then the filtered ICA is used to study degradation variation phenomenon of retired batteries, the relationship between IC curve feature and remaining capacity was analyzed. Finally, a fast capacity estimation method based on incremental capacity and Gaussian process regression is proposed.

## 2 Experiment and test results

### 2.1 Test setup and experimental method

Our retired battery pack is received from an electric bus which is operated for more than three years. As is shown in Fig. 1, the experimental pack consists of 42 series-connected battery modules, 28 battery cells are parallel-connected in each module. Air cooling is adopted for battery heat dissipation in the whole battery pack.

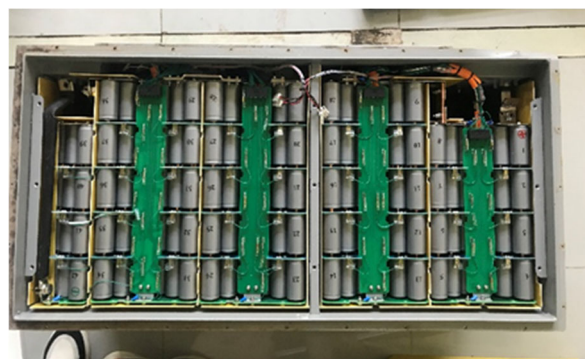


Fig. 1 Experimental retired battery pack

The battery cell is manufactured by OptimumNano Energy CO., LTD of China, and their

\* Corresponding author: [jinpeng@pkusz.edu.cn](mailto:jinpeng@pkusz.edu.cn)

parameters are itemized in Table 1. The 32700-type cell is configured with LiFePO4/graphite as electrode, whose nominal capacity is 5.0 Ah, upper and lower limit voltages are 2.65–3.7V. All of our tests are carried out at room temperature. An 8-channel battery test system (Shenzhen Repower Technology Co.,Ltd ACTS 5V-10A) whose current range is -5A~5A and voltage range is 0~5V(±0.02% at full scale) is used for battery test.

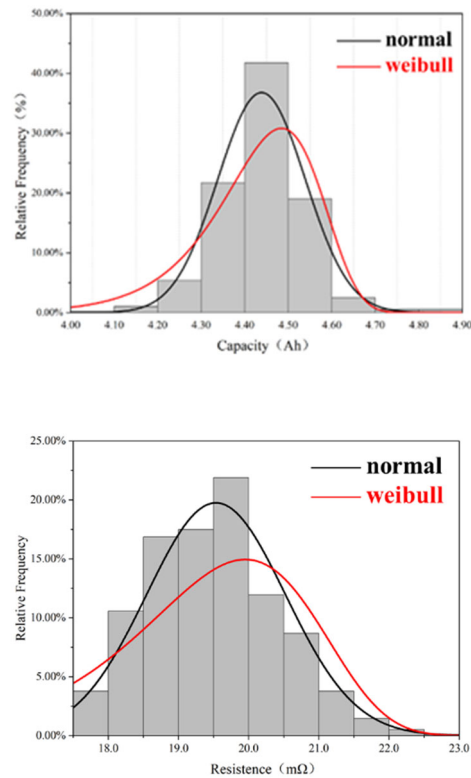
**Table 1** Parameters for the LiFePO4 cell.

Items	Parameters
Manufacturer	OptimumNano
Group pattern	28 parallel-42 series
Cathode material	LiFePO4
Anode material	Graphite
Nominal capacity	5.0Ah
Nominal voltage	3.2V
End-of-charge voltage	2.65V
End-of-discharge voltage	3.7V

In order to obtain the remaining capacity, we apply the charging current rate of 1/3 C, i.e., 1.66 A, and test the internal resistance by hybrid pulse power characterization (HPPC) method.

## 2.2 Test results

Capacity and resistance of all 1176 retired battery cell is surveyed by battery test system, Fig.2 shows the distribution histogram of these batteries. As shown in Fig.2, the capacity data of 1176 single cells in the entire battery pack shows that most of the cells have a capacity between 4.2Ah and 4.6Ah and obey a normal distribution. 90.2% of the cells still maintain the rated capacity. 80%, 98.7% batteries maintain 70% of the rated capacity. In the histograms, the distribution of battery capacity is left-skewed with  $s=-0.755$  and the distribution of battery resistance is  $s=-1.18$ .



**Fig. 2.** Battery cell capacity and resistance bargraph and probability plots

For both battery capacity and resistance probability plots, some cells are outside the confidence intervals. However, normal probability has a better fitting quality than that of Weibull probability, for the reason that in the low resistance range of the probability map, there are some points far away from the lower confidence interval. Therefore, the normal distribution is more suitable than the Weibull distribution for the batteries that we study with parallel connections.

## 3 Methodology

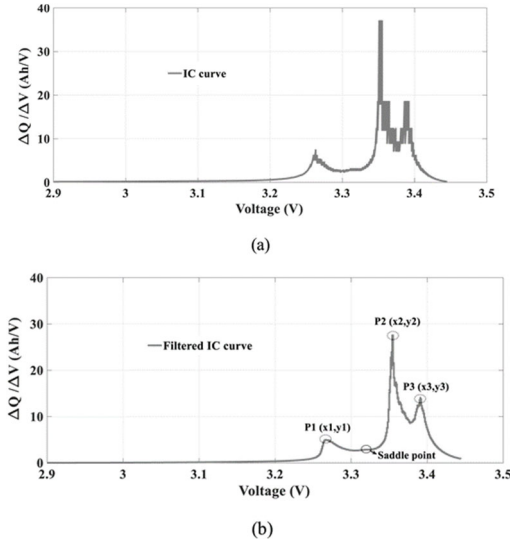
### 3.1 Incremental capacity analysis

The incremental capacity is obtained from a full charging/discharging cycle, where we get the changing rate of capacity with the voltage as the vertical coordinate, and the changing voltage as the horizontal ordinate, which is shown as following:

$$\frac{dQ(t)}{dV(t)} \approx \frac{\Delta Q(k)}{\Delta V(k)} = \frac{Q(k)-Q(k-N)}{V(k)-V(k-N)} \quad (1)$$

In Eq. 1,  $\Delta Q(k)$  and  $\Delta V(k)$  are the rate of electric quantity and terminal voltage during charging/discharging, respectively. As a result,

$\frac{\Delta Q(k)}{\Delta V(k)}$  can be described as  $\frac{dQ(t)}{dV(t)}$  when the sample interval  $N$  is short enough.



**Fig. 3.** (a) Initial IC curve; (b) Filtered IC curve

As a novel technique to analysis aging phenomenon and survey state of health (SOH) of lithium battery, ICA is applied in fast estimation in our essay. In order to extract capacitance features of incremental capacity, preprocessing needs to be conducted previously. Figure.3 illustrates the initial IC curve contaminated by noises and filtered IC curve using Kalman filter in the 1/3C current rate. Finally we need to record the coordinates of the characteristic peak(P1, P2 and P3) which are highly related to the attenuation of battery capacity.

### 3.2 Gaussian process regression

Gaussian process regression (GPR) is a neoteric regression approach in machine learning, so we explain how apply the algorithm to estimate attenuation of battery capacity.

First, in our training dataset  $D=(X, y)$ , the input variable and output variable are indicated as  $X$  and  $y$  accordingly. As the formula  $X=\{x_1^D, x_2^D, x_3^D, \dots, x_N^D\}$  shows, there are  $D$ -dimensional  $N$  input vectors in variable  $X$ , which follows multivariable distribution with mean function and covariance function of  $m(x)$  and  $k_f(x, x')$ , respectively. We express the properties of  $f(x)$  as following:

$$\begin{cases} m(x) = E(f(x)) \\ k_f(x, x') = E[(m(x) - f(x))(m(x') - f(x')))] \end{cases} \quad (2)$$

For the GPR algorithm, in order to express the prior assumptions for the properties of latent function, calculating the the kernel function  $k_f(x, x')$  is necessary. We use the squared exponential covariance (SE) to complete this step, as shown in Eq. (3):

$$k_f(x, x') = \sigma_f^2 \exp\left(-\frac{(x-x')^2}{2l^2}\right) \quad (3)$$

Considering there are  $D$ -dimensional  $N$  input vectors in variable  $X$ , we can converse Eq. (3) as follows:

$$k_f(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{d=1}^D \frac{(x_i^d - x_j^d)^2}{l_d^2}\right) \quad (4)$$

Here, the noise covariance matrix  $\sigma_n^2 I$  is appended into Eq. (4) as below:

$$k(x_i, x_j) = k_f(x_i, x_j) + \sigma_n^2 I \quad (5)$$

Considering the input variable  $X$  and the properties of the latent function  $f$ , the distribution of  $y$  is denoted as follows:

$$p(y|X) = \int p(y|f, X)p(f|X)df = N(0, K + \sigma_n^2 I) \quad (6)$$

Then, we denote the marginal log-likelihood of  $y$  as:

$$\log p(y|X, \Theta) = -\frac{1}{2}y^T (K + \sigma_n^2 I)^{-1} y - \frac{1}{2} \ln |K + \sigma_n^2 I| - \frac{n}{2} \log 2\pi \quad (7)$$

Here, for the marginal log-likelihood function, we employ the gradient method to optimize the hyper-parameters, and the calculation process of the gradient descending is:

$$\begin{cases} \frac{\partial \log p(y|X, \Theta)}{\partial \theta_i} = -\frac{1}{2} \text{tr} \left( \varphi^{-1} \frac{\partial \varphi}{\partial \theta_i} \right) + \frac{1}{2} y^T \varphi^{-1} \frac{\partial \varphi}{\partial \theta_i} \varphi^{-1} y \\ \varphi = K + \sigma_n^2 I \end{cases} \quad (8)$$

Here, in training dataset, we employ a small representative subset for the sparse approximation propose to handle GPR. The joint prior distribution of the training output  $y$  and the test output  $y^*$  is deduced and expressed as [8]:

$$p(y, y^*|X, x^*, \Theta) = N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \varphi & k^* \\ (k^*)^T & k^{**} + \sigma_n^2 \end{bmatrix} \right) \quad (9)$$

Based on Eq. (10), the posterior distribution can be derived as:

$$p(y^*|x, y, x^*) = N(y^*|\bar{y}^*, \text{cov}(y^*)) \quad (10)$$

Then, in the predictive distribution, we calculate the mean  $\bar{y}^*$  and covariance  $\text{cov}(y^*)$ :

$$\bar{y}^* = (k^*)^T \varphi^{-1} y \quad (11)$$

$$\text{cov}(y^*) = \sigma_n^2 + k^{**} - (k^*)^T \varphi^{-1} k^* \quad (12)$$

Finally, a full GPR algorithm modeling is done.

### 4 Validation and discussion

In view of the proposed GPR battery screening model mentioned above, 5 LFP retired batteries data of 190 cycle are employed to verify the validity of battery estimation. In our experiment, we regard the first 80% data of charging/discharging period as the training datasets and the rest 20% data as testing part.

The results for battery No. 61, 62, 63, 64 and 65 are shown in Table. 2.

**Table 2** Capacity estimation results vs. real capacity of the screening method

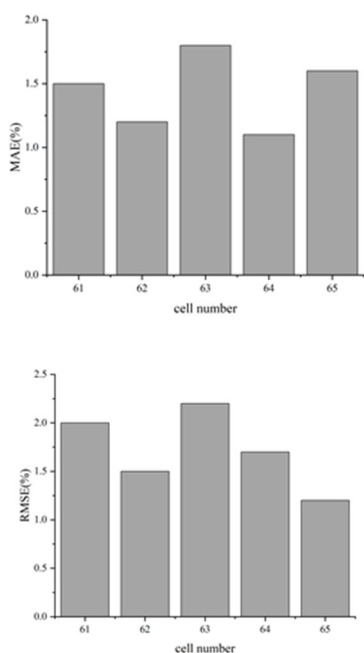
Cell number	Estimation results (Ah)	Real capacity (Ah)	error (%)
C61	3.94	4.04	1.76
C62	3.69	3.64	1.41
C63	4.92	4.79	2.59
C64	4.15	4.01	1.53
C65	3.91	3.96	1.93

As illustrated in Table 2, error of the estimation of battery No. 61, 62, 63, 64 and 65 are 1.76%, 1.41%, 2.59%, 1.53%, 1.93%, respectively. As a result, the battery SOH estimation of GPR shows strong performance. In addition, we also employ the mean absolute error (MAE) and the root mean square error (RMSE) to analyze the performance of the model quantitatively, whose definition is shown as:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}_i^*| \quad (18)$$

$$RMSE = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y}_i^*)^2\right)} \quad (19)$$

Where  $\bar{y}_i^*$  is the estimated value, the  $y$  is the real value, and it is worth noting that  $N$  represents how many times battery cycles, so MAE and RMSE could be used to assess the algorithm modelling of different cells in view of absolute sign.



**Fig.4.** The error analysis of battery estimation based GPR

MAE reflects the mean absolute error between real value and estimated value in our cells SOH estimation, we calculate MAE and RMSE of battery No. 61, 62, 63, 64 and 65. As Fig.4 shows, battery No.63 has the maximum MAE with 1.8% which is not overrun 2%, while battery NO.64 has the minimum MAE with 1.1%. In term of RMSE, all 4 LFP cells are nearly 1.7%, battery NO.63 has the maximum RMSE with 2.2%, and battery NO.65 has the minimum RMSE with 1.4%. As a result, our proposed SOH fast estimation based on GPR indicates high accuracy and reliability.

## 5 Conclusion

This paper presents an efficient capacity estimation based on ICA and Gaussian process regression for retired lithium-ion batteries. The main contributions are summarized as follows: (1) a detailed survey of a retired battery pack and cells was conducted, normal distribution is more suitable than the Weibull distribution for the batteries that we study with parallel connections. (2) The GPR is employed to establish SOH estimation model and how the algorithm works is demonstrated. (3) Our novel method is verified by LFP retired cells charging and discharging datasets. Verification result with a maximum MAE and RMSE of 1.8%, 2.2% shows strong performance.

## Acknowledgments

This work was financially supported by Shenzhen Science Research Project under Grant GJHZ20180928162210431.

## References

1. Debnath, U.K., I. Ahmad and D. Habibi, Quantifying economic benefits of second life batteries of gridable vehicles in the smart grid. *International Journal of Electrical Power & Energy Systems*, 2014. 63: p. 577-587.
2. Cusenza, M.A., et al., Reuse of electric vehicle batteries in buildings: An integrated load match analysis and life cycle assessment approach. *Energy and Buildings*, 2019. 186: p. 339-354.
3. Canals Casals, L., M. Barbero and C. Corchero, Reused second life batteries for aggregated demand response services. *Journal of Cleaner Production*, 2019. 212: p. 99-108.
4. Zhou, L., et al., A study on parameter variation effects on battery packs for electric vehicles. *Journal of Power Sources*, 2017. 364: p. 242-252.

5. Tang, X., et al., A fast estimation algorithm for lithium-ion battery state of health. *Journal of Power Sources*, 2018. 396: p. 453-458.
6. Xiaoyu, L., et al., State of health estimation for Li-ion battery using incremental capacity analysis and Gaussian process regression. *Energy*, 2020. 190: p. 116-467
7. Weng, C., et al., State-of-health monitoring of lithium-ion battery modules and packs via incremental capacity peak tracking. *Applied Energy*, 2016. 180: p. 360-368.
8. Zhongkai, Z., et al., An efficient screening method for retired lithium ion batteries based on support vector machine. *Journal of Clean Production*, 2020. 267: p. 121-882