A State of Health Estimation Method for Lithium-Ion Batteries Based on Improved Particle Filter Considering Capacity Regeneration

Haipeng Pan, Chengte Chen and Minming Gu *

School of Mechanical and Automatic, Zhejiang Sci-Tech University, Hangzhou 310018, China; pan@zstu.edu.cn (H.P.); 201920505015@mails.zstu.edu.cn (C.C.)
* Correspondence: guminming@zstu.edu.cn

Abstract: Accurately estimating the state of health (SOH) of a lithium-ion battery is significant for electronic devices. To solve the nonlinear degradation problem of lithium-ion batteries (LIB) caused by capacity regeneration, this paper proposes a new LIB degradation model and improved particle filter algorithm for LIB SOH estimation. Firstly, the degradation process of LIB is divided into the normal degradation stage and the capacity regeneration stage. A multi-stage prediction model (MPM) based on the calendar time of the LIB is proposed. Furthermore, the genetic algorithm is embedded into the standard particle filter to increase the diversity of particles and improve prediction accuracy. Finally, the method is verified with the LIB dataset provided by the NASA Ames Prognostics Center of Excellence. The experimental results show that the method proposed in this paper can effectively improve the accuracy of capacity prediction.

Keywords: lithium-ion battery; capacity regeneration; capacity estimation; calendar time; improved particle filter

1. Introduction

The LIB with high energy ratio, high cycle life and no memory effect is widely used in ships, automobiles, aerospace, and other industries [1]. However, in the long-term charging-discharging state, the LIB will be aging. Its available capacity will decrease, performance will degrade, and even potential safety hazards will arise.

Therefore, it is essential to predict the available capacity of an LIB [2]. In recent years, methods for predicting the available capacity of an LIB have been mainly divided into data-driven methods and model-based methods. The data-driven method is a method to analyze characteristic data and indicate degrees of performance degradation with the help of intelligent algorithms [3]. A variety of studies in the literature have reported on data-driven methods such as the autoregressive statistical model [4], neural network [5,6], support vector machine [7], and correlation vector machine [8] to predict the available capacity of an LIB and obtain relatively accurate prediction results. However, the data-driven method ignores the internal relationships of the physical process. It uses implicit functions to represent the model, which is not conducive to analyzing the physical relationships between data. In addition, the data-driven method requires a large volume of sample data, but the data on the LIB degradation process are difficult to obtain. These two disadvantages hinder the application of data-driven methods.

Unlike the data-driven method, the model-based method combines mathematical functions and filtering technology to predict performance degradation trends. Zhang et al. [9] proposed a particle filter method based on an exponential degradation model to predict the SOH of LIBs. The method reduced the interference of measurement noise and improved the accuracy of prediction. Su et al. [10] combined a variety of degradation models. They used interactive multi-model particle filtering to predict the available capacity of LIBs, which
improved the fit between the model and the data. Aiming at the low prediction accuracy of standard particle filters, Miao et al. [11] proposed a combination of a particle filter and unscented Kalman filter to predict the SOH of LIBs. This method effectively improved prediction accuracy by performing unscented Kalman filtering on each particle. On this basis, Zhang et al. used the posterior probability density of an unscented Kalman filter to resample particles, which increased particle diversity [12]. However, the above methods only consider the law of the long-term capacity degradation of the LIB after multiple charge–discharge cycles. The impacts of electrochemical reactions on short-term capacity degradation during each charge–discharge cycle of an LIB are not taken into account. Considering the influence of the state of charge (SOC) on health state, Ma et al. [13] performed Kalman filtering and Gauss-Hermite particle filtering on SOC and SOH simultaneously. They used the method of dual time scales to predict the available capacity. Similarly, combining the SOC and SOH, Qiu et al. [14] proposed a fusion method of extended Kalman filtering and particle filtering to estimate the available capacity jointly. To increase the diversity of particles, the cuckoo algorithm was adopted instead of the standard resampling algorithm. In addition to Kalman filtering and particle filtering, Dong et al. [15] used a combination of the Brownian motion degradation model and particle filtering to predict the useful capacity of LIB. On this basis, Zhai et al. [16] proposed a new adaptive Wiener process model, which took into account the sample data collected at different time intervals and combined multiple sample estimation parameters to improve the model’s generalization ability. However, this method only relied on historical data to predict available capacity and did not consider the impact of the current environment on the degradation of LIB. Liu et al. [17] considered LIB operating temperature and discharge depth factors, established a LIB aging model, and used Gaussian process regression to predict available capacity. These methods presented the predicted results in the form of probability density functions. They took into account the randomness of the degradation process to accurately predict the available capacity of LIB to a certain extent. However, none of the above methods considered the capacity regeneration phenomenon of LIB, which would produce a large available capacity estimation error.

Capacity regeneration refers to the phenomenon that the available capacity of LIB suddenly increases after a complete charge–discharge and a long period of storage [18]. Capacity regeneration will have a great impact on the available capacity, which is becoming a research hotspot. Wang et al. used the total variation filtering method to detect the capacity regeneration point of historical data [19]. However, this method could not predict the capacity regeneration point. Xu et al. proposed a method for SOH prediction based on the Wiener process for capacity regeneration [20,21], which accurately predicted the LIB capacity regeneration point. However, the premise of this method was to obtain all the calendar time data of the LIB.

Aiming at the impact of capacity regeneration on available capacity estimation and reducing the available capacity prediction error, this article will focus on the following two aspects. Firstly, the degradation process of LIB is divided into the normal degradation stage and capacity regeneration stage, and an MPM for the SOH is established. Then, the particle filter algorithm based on a genetic algorithm is used to modify the prediction results of the model to ensure the accuracy of the available capacity prediction. Finally, this paper will use the LIB dataset provided by NASA Ames Prognostics Center of Excellence to verify the effectiveness of this method.

2. Lithium-Ion Battery State of Health Model

In this paper, the data on No. 5, No. 6, No. 7, and No. 18 batteries in the NASA LIB dataset were selected. The dataset came from a batch of second-generation 18650 size LIB with positive electrode consisting of lithium cobalt oxide (LiCoO₂) and negative electrode of lithiated carbon (LiC) [22,23]. In this dataset, the aging test results were obtained by a set of four LIB operating in three different operating modes (charge, discharge, and impedance) at room temperature. In this process, the LIBs were charged in 1.5 A constant current (CC) mode until the battery voltage reached 4.2 V and were then charged in constant voltage
(CV) mode until the charging current dropped to 20 mA. After that, the No. 5, No. 6, No. 7, and No. 18 batteries were discharged at 2 A constant current (CC) level until their battery voltage dropped to 2.7 V, 2.5 V, 2.2 V, and 2.5 V, respectively. The experiment stopped when a battery reached the end of life (EOL) standard.

LIB aging is usually caused by the growth of solid electrolyte interface (SEI) film [24]. This process will reduce the available capacity of an LIB [25]. However, in the aging process of an LIB, the available capacity is not monotonously decreasing, but undergoing a process of sudden increase and slow decrease alternately. Therefore, this paper divides the LIB degradation process into a normal degradation stage and a capacity regeneration stage.

The SOH of an LIB is defined as follows.

\[ S = \frac{C_{\text{cur}}}{C_{\text{ini}}} \]  

(1)

where \( S \) represents the SOH, \( C_{\text{cur}} \) represents the current available capacity, and \( C_{\text{ini}} \) represents the initial capacity.

Therefore, the capacity degradation rate is expressed as follows.

\[ y = (1 - S) \times 100\% = \frac{C_{\text{ini}} - C_{\text{cur}}}{C_{\text{ini}}} \times 100\% \]  

(2)

where \( y \) represents the capacity degradation rate.

In Figure 1, the abscissa is the number of LIB charge–discharge cycles. The ordinate is the capacity degradation rate. Obviously, the degradation process of an LIB is a combination of multiple curves rather than a complete curve. Assuming that capacity regeneration is eliminated, the degradation process of an LIB can be regarded as several normal degradation stages.

![Figure 1. Capacity degradation rate curve of No. 5 battery.](image)

This paper adopts the power function fitting method to establish a mathematical model of the relationship between the number of cycles and the capacity degradation rate in the normal degradation stage.

\[ y(t) = at^b \]  

(3)

where \( a, b \) represent degradation parameters and \( t \) is the number of cycles after the capacity regeneration stage.

In Figure 2, the abscissa is the number of charge–discharge cycles. The ordinate is the available capacity value. The bold curve is the capacity regeneration stage, which makes the degradation process of the LIB nonlinear. If this stage is ignored, the modeling accuracy will be reduced, and the available capacity prediction error for the LIB will be increased.

In Figure 3, the abscissa is the calendar time. The ordinate is the capacity degradation rate. Capacity regeneration is often caused by calendar time exceeding its threshold. [20]
According to this law, a prediction algorithm based on capacity regeneration is proposed. The algorithm is shown in Figure 4.

![Figure 2](image-url)

**Figure 2.** Degradation process curve of No. 5 battery.

![Figure 3](image-url)

**Figure 3.** Capacity degradation rate curve of No. 5 battery by calendar time.

![Figure 4](image-url)

**Figure 4.** Capacity regeneration detection algorithm flow chart.

\[ y_{RN} \]

\[ \Delta t_{C,N} = t_{C,N+1} - t_{C,N} \]

\[ \Delta t_{C,Ri}, i = 0 \]

\[ \Delta t_{C,Ri} > \Delta t_{C,Rh} \]

\[ \Delta t_{C,Rh} = t_{C,Rh+1} - t_{C,Rh} \]

\[ \Delta y_{RN} = y_{N,i} - y_{N+1} \]

Yes \quad i = i + 1

No

\[ y_{RN} \]
$N$ represents the number of correlated cycles, $y$ represents the capacity degradation rate, $t_C$ represents the calendar time, $\Delta t_{C,th}$ represents the calendar time threshold of capacity regeneration, $\Delta t_{C,Re}$ represents the calendar time during capacity regeneration stage, and $\Delta y_{Re}$ represents the capacity regeneration value.

The relationship between LIB capacity regeneration value and the calendar time is shown in Figure 5.

Figure 5. Relationship between capacity regeneration value and calendar time.

In Figure 5, the abscissa is the calendar time and the ordinate is the regeneration capacity value. Although the capacity regeneration value is positively correlated with calendar time, there is a non-linear relationship between the two. To express the relationship between the two more accurately, we use a power function curve to approximate the relationship between the capacity regeneration value and the calendar time.

$$y_{Re}(t_C) = a_C(t_C - t_{C,th})^{b_C}$$

where $y_{Re}$ represents the capacity regeneration value, $a_C, b_C$ represent parameters, $t_C$ represents the calendar time, and $t_{C,th}$ represents the calendar time threshold of capacity regeneration.

In summary, the model proposed in this paper is as follows:

$$y(k) = y(k-1) + a \left( t_k^b - t_{k-1}^b \right) - \lambda a_C (t_k - t_{C,th})^{b_C} + \omega(k)$$

$$\lambda = \begin{cases} 0 & t_{C,k} < t_{C,th} \\ 1 & t_{C,k} \geq t_{C,th} \end{cases}$$

where $k$ represents the number of correlated cycles, $\lambda$ represents the capacity regeneration parameter, $t_k$ represents the number of cycles after the capacity regeneration stage, $y$ represents the predicted capacity value, and $\omega(k)$ represents the systematic error, which obeys a normal distribution with a mean value of 0 and a variance of $Q_\omega$.

Because there are some unknown parameters in the normal degradation model and capacity regeneration model, we use the least squares method to estimate these unknown parameters. The least squares method is a mathematical optimization technique that seeks the best parameter matching of data by minimizing the sum of squares of errors. The least squares method parameter estimation is specifically divided into the following two steps.

Step 1: Assuming that the capacity regeneration model is estimated with $N$ capacity regeneration data $y_1, y_2, y_3 \ldots y_n$. Take the logarithm of the capacity regeneration model and calculate the square difference.

$$F = \sum_{i=1}^{N} \left( \ln y_i - \ln a_c - b_c \ln(t_{c,i} - t_{c,th}) \right)^2$$
Step 2: Derivation of the function.

\[
d\frac{F}{da_c} = \sum_{i=1}^{n} \frac{2}{a_c} (\ln y_i - \ln a_c - b_c \ln (t_{ei} - t_{ejh}))
\]

(7)

\[
d\frac{F}{db_c} = \sum_{i=1}^{n} 2 (\ln (t_{ei} - t_{ejh})) (\ln y_i - \ln a_c - b_c \ln (t_{ei} - t_{ejh}))
\]

(8)

Make the derivation result of the above formula equal to 0, and the derivation of \(b_c\) is as follows.

\[
b_c = \frac{n \sum_{i=1}^{n} \ln y_i \ln (t_{ei} - t_{ejh}) - \sum_{i=1}^{n} (\ln (t_{ei} - t_{ejh}) \sum_{j=1}^{n} \ln y_j)}{n \sum_{i=1}^{n} (\ln (t_{ei} - t_{ejh}))^2 - \sum_{i=1}^{n} (\ln (t_{ei} - t_{ejh}) \sum_{j=1}^{n} \ln (t_{ei} - t_{ejh}))}
\]

(9)

Then, by substituting the obtained parameter \(b_c\) into (7), parameter \(a_c\) is estimated.

3. Model Modification

Because the LIB degradation model is not completely accurate when estimating the available capacity, we need to combine the measured values and model prediction values to correct the prediction results and thus reduce subsequent estimation errors. The method flow chart is shown in Figure 6, and the specific method is introduced as follows.

![Figure 6. Flow chart of model correction method.](image)

In statistics, particle filtering is called the sequential Monte Carlo method, which combines the Bayesian principle and importance sampling [26]. It uses sample form instead of function form to describe the probability distribution of nonlinear stochastic systems [27].

Firstly, the state-space model is established. The state equation is as Equation (5), and the output equation is as follows.

\[
z(k) = y(k) + v(k)
\]

(10)
where $k$ represents the number of related cycles, $y$ represents the predicted capacity degradation rate, $v(k)$ obeys a normal distribution with a mean value of 0 and a variance of $Q_v$, representing the measurement error, and $z(k)$ represents the actual measured capacity degradation rate during $k$ cycles.

The standard particle filter algorithm is shown in the following steps.

Step 1: Initialization.
When $i = 0$, $N$ particles $\{ \hat{z}_k \}^N_{k=1}$ are randomly generated from the prior normal distribution $N(z_0, H_0)$, and the corresponding weight $\{ \omega_k \}^N_{k=1}$ is generated for each particle.

Step 2: Prediction
According to $y_{k|i-1} \sim q(y_{k|i-1} | y_{k|i-1}, z_i)$, a priori estimate $y_{k|i-1}^k$ can be derived from (5).

$$y_{k|i-1}^k = y_{k|i-1} + a(t_b^i - t_{i-1}^b) - \lambda aC_{C,C} + \omega_i$$

where $k = 1 : N$.

Step 3: Weight update
The particle weight is calculated by the following formula.

$$\omega_i^k = 1/\sqrt{2\pi Q_v} \exp \left( - \frac{(z_i - y_{k|i-1}^k)^2}{2Q_v} \right)$$

Normalize the weights.

$$\omega_i^k = \omega_i^k / \sum_{i=1}^{N} \omega_i^k$$

Step 4: Standard resampling
Generate a random number between 0 and 1, and copy the particles according to the particle weight interval. Then, a new weight $\omega_i^k$ is generated for each particle. Finally, the state estimate is obtained based on the particles and their weights.

$$y_{i|i} = \sum_{i=1}^{N} \omega_i^k y_{k|i-1}^k$$

Standard resampling will lead to particle dilution and reduce the prediction accuracy. This paper proposes a resampling algorithm based on a genetic algorithm to increase the diversity of current particles.

The specific steps of resampling algorithms based on the genetic algorithm are as follows.

Step I: Coding
The particle weight $\omega_i^k$ obtained in step 3 is binary coded.

Step II: Selection operator
The selection operation can avoid the loss of good genes and improve the global convergence and computational efficiency. This paper chooses the roulette selection method to generate a random number $f_i$ between 0 and 1. When $f_i$ falls in the particle interval, the particle will be copied as the parent particle.

Step III: Crossover operator
The crossover operation refers to randomly pairing the parent particles. The two particles exchange part of their genes in a single-point or multi-point manner to form two new offspring particles. In this paper, the crossover method is the single point crossover, as shown in Figure 7. By performing this operation, the search ability of the genetic algorithm can be greatly improved.
to generate a random number.

In genetic algorithms and at the same time ensure the diversity of particles [28]. The mutation probability \( V_m \) determines the probability of a certain particle participating in the mutation. Once the mutated gene position is determined, these genes will change, and the remaining genes will remain unchanged.

**Step V: Decoding**

Perform a decimal decoding operation on the parent particle and the child particle to obtain the weight \( \omega_i \). Finally, the state estimation is obtained based on the particle and its weight, such as Equation (14).

**4. Experimental Results and Discussion**

Firstly, this article sets the calendar time threshold for capacity regeneration to 9.72 h and the failure of an LIB is defined as a capacity degradation rate of 24%. Figure 5 shows that when the calendar time of the LIB increases, the capacity regeneration value increases significantly. According to the least squares parameter estimation method, the parameters of the normal degradation model and the capacity regeneration model are estimated, and the results are as follows.

In Table 1, \( a, b \) are parameters of the normal degradation model, and \( a_c, b_c \) are parameters of the capacity regeneration model. According to the estimated parameters, the normal degradation model curve and the capacity regeneration model curve can be obtained, as shown in Figure 8.

**Table 1. Multi-stage prediction model parameters.**

<table>
<thead>
<tr>
<th>Model Parameter</th>
<th>( a )</th>
<th>( b )</th>
<th>( a_c )</th>
<th>( b_c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.0230</td>
<td>0.4472</td>
<td>( 1.3967 \times 10^{-4} )</td>
<td>0.4392</td>
</tr>
</tbody>
</table>

Obviously, in Figure 8a, the normal degradation model curve is very close to the true value, which reflects the superiority of the model in Figure 8b, because capacity regeneration is a complex chemical reaction and there are errors in capacity measurement. There is a certain error between the capacity regeneration model and the true value. Using the obtained model parameters and the current calendar time, we can estimate the next capacity degradation rate. In this paper, the data on No. 6, No. 7, and No. 18 batteries were used to identify the parameters of a capacity regeneration model to predict the capacity degradation rate of the No. 5 battery. To verify the superiority of the method, the improved particle filtering method based on an empirical model (EM-IPF) and the particle filtering method based on a multi-stage prediction model (MPM-IPF) were compared with the proposed improved particle filtering method based on a multi-stage prediction model (MPM-IPF).
Perform a decimal decoding operation on the parent particle and the ... The blue curve represents the EM-IPF method, the black curve the MPM-PF method, and the red curve the MPM-IPF method proposed in this paper. Obviously, the three methods have approximate prediction results. However, the EM-IPF method cannot predict the capacity regeneration stage, so the capacity degradation will not drop suddenly in the capacity regeneration stage. The MPM-PF method can predict the capacity regeneration stage. Still, the prediction error is relatively large due to the low accuracy of the standard particle filter. The MPM-IPF method proposed in this paper has a better tracking and prediction effect with the true value in the normal degradation stage. Additionally, it can predict the true value more accurately during the capacity regeneration stage, which proves the superiority of the proposed prediction method.

![Figure 8](image)

**Figure 8.** (a) Normal degradation model curve; (b) capacity regeneration model curve.

Figure 9 shows three prediction curves. The blue curve represents the EM-IPF method, the black curve the MPM-PF method, and the red curve the MPM-IPF method proposed in this paper. Obviously, the three methods have approximate prediction results. However, the EM-IPF method cannot predict the capacity regeneration stage, so the capacity degradation will not drop suddenly in the capacity regeneration stage. The MPM-PF method can predict the capacity regeneration stage. Still, the prediction error is relatively large due to the low accuracy of the standard particle filter. The MPM-IPF method proposed in this paper has a better tracking and prediction effect with the true value in the normal degradation stage. Additionally, it can predict the true value more accurately during the capacity regeneration stage, which proves the superiority of the proposed prediction method.

![Figure 9](image)

**Figure 9.** Comparison chart of No. 5 battery capacity prediction.

To compare the performance of the proposed MPM-PF with other algorithms (such as EM-IPF and MPM-PF), the error results of MPM-PF, EM-IPF, and MPM-PF were compared. Figure 10 describes the MAE and RMSE results of the above three methods for the capacity prediction error of the No. 5 battery. Table 2 shows the detailed capacity prediction error results of different prediction methods at 30, 60, and 90 cycles. The three error curves shown in Figure 10 increase sharply in the capacity regeneration stage due to the deviation between the predicted value and the true value. However, obviously, in the capacity regeneration stage, the MAE and RMSE values of the MPM-IPF method proposed in this paper are significantly smaller than those of the EM-IPF and MPM-PF methods. This demonstrates the prediction accuracy of the MPM-IPF method. Moreover, during the
degradation process of an LIB, the MAE and RMSE values of the method proposed in this paper are lower than those of the comparison methods, and neither exceeds 0.6%. Once again, this proves the superiority of the MPM-IPF method.

Figure 10. Comparison chart of prediction error of No. 5 battery capacity: (a) MAE; (b) RMSE.

Table 2. Comparison of error results of three algorithms based on No. 5 battery data.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Cycle</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPM-IPF</td>
<td>30</td>
<td>0.0024</td>
<td>0.0032</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.0026</td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.0029</td>
<td>0.0049</td>
</tr>
<tr>
<td>EM-IPF</td>
<td>30</td>
<td>0.0043</td>
<td>0.0077</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.0041</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.0042</td>
<td>0.0083</td>
</tr>
<tr>
<td>MPM-PF</td>
<td>30</td>
<td>0.0026</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.0031</td>
<td>0.0043</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.0033</td>
<td>0.0053</td>
</tr>
</tbody>
</table>

5. Conclusions

Capacity regeneration will increase the available capacity prediction error for an LIB. The LIB degradation model considering capacity regeneration is more consistent with the degradation process of an LIB. An improved particle filtering algorithm based on a multi-stage prediction model was proposed using data for the 18650 lithium battery. Compared with previous research, this method considers the influence of calendar time on capacity regeneration, accurately predicts the available capacity of LIB by taking calendar time as input, provides early warning of LIB failure, and improves the reliability of the energy storage system. This research lays a foundation for obtaining SOH and the remaining useful life of LIB.

Author Contributions: Conceptualization and methodology, H.P.; validation and writing—original draft preparation, C.C.; writing—review and editing and supervision, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Zhejiang Science Technology Department Public Service Technology Research Project, grant number LGG19F030002 and the Zhejiang Science Technology Department Public Service Technology Research Project, grant number LGG21F030015.

Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in NASA PCoE Datasets.

Acknowledgments: The authors would like to thank the National Aeronautics and Space Administration (NASA)’s Ames Prognostics Center of Excellence for providing the reliability testing data on lithium-ion batteries.

Conflicts of Interest: The authors declare no conflict of interest.

References

7. Xue, Z.; Zhang, Y.; Cheng, C.; Ma, G. Remaining useful life prediction of lithium-ion batteries with adaptive unscented Kalman filter and optimized support vector regression. *Neurocomputing* 2020, 376, 95–102. [CrossRef]
14. Qiu, X.; Wu, W.; Wang, S. Remaining useful life prediction of lithium-ion battery based on improved cuckoo search particle filter and a novel state of charge estimation method. *J. Power Sources* 2020, 450, 227700. [CrossRef]
20. Xu, X.; Yu, C.; Tang, S.; Sun, X.; Si, X.; Wu, L. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Wiener Processes with Considering the Relaxation Effect. * Energies* 2019, 12, 1685. [CrossRef]


